

DEPARTMENT OF MANAGEMENT

AFDELING FOR VIRKSOMHEDSLEDELSE

Working Paper 2005-5

The Effect of Long Memory in Volatility on Stock
Market Fluctuations

Bent Jesper Christensen and Morten Ørregaard Nielsen

UNIVERSITY OF AARHUS • DENMARK

ISSN 1398-6228

The Effect of Long Memory in Volatility on Stock Market Fluctuations*

Bent Jesper Christensen
University of Aarhus

Morten Ørregaard Nielsen[†]
Cornell University

November 15, 2005

Abstract

Recent empirical evidence demonstrates the presence of an important long memory component in realized asset return volatility. We specify and estimate multivariate models for the joint dynamics of stock returns and volatility that allow for long memory in volatility without imposing this property on returns. Asset pricing theory imposes testable cross-equation restrictions on the system that are not rejected in our preferred specifications, which include a strong financial leverage effect. We show that the impact of volatility shocks on stock prices is small and short-lived, in spite of a positive risk-return trade-off and long memory in volatility.

JEL Classifications: C32, G12.

Keywords: Financial leverage, long memory, realized volatility, risk-return trade-off, stochastic volatility, stock prices, VARMA models, VIX implied volatility.

*We are grateful to three anonymous referees for useful and constructive comments, to Tim Bollerslev for supplying the realized volatility data used in our empirical work, to Rolf Karlsen for research assistance, and to the Center for Analytical Finance (CAF) for financial support. The second author is also grateful for financial support from the Danish Social Science Research Council (grant no 24-02-0181).

[†]Corresponding author. Please address correspondence to: Morten Ørregaard Nielsen, Department of Economics, Cornell University, 482 Uris hall, Ithaca, NY 14853, USA; phone: +1 607 255 6338; fax: +1 607 255 2818; email: mon2@cornell.edu

1 Introduction

The definition, computation and analysis of realized volatility in financial return series has attracted considerable interest in recent literature, e.g. French, Schwert & Stambaugh (1987), Schwert (1989), Andersen & Bollerslev (1998), and Barndorff-Nielsen & Shephard (2002). Essentially, integrated instantaneous variance is estimated consistently by its sample analogue based on high-frequency return observations. This approach allows gathering much more detailed information on the properties of financial market volatility than previously. A striking finding in the recent empirical literature is that realized volatility exhibits long memory. This finding is consistent across a number of studies, and financial theory may accommodate long memory in volatility as well, see e.g. Robinson (1991), Crato & de Lima (1994), Baillie, Bollerslev & Mikkelsen (1996), Ding & Granger (1996), Breidt, Crato & de Lima (1998), Comte & Renault (1998), Robinson (2001), and Andersen, Bollerslev, Diebold & Labys (2003) and the references therein.¹ However, so far no study has related the long memory property of volatility to the level of asset prices themselves. The time series behavior of volatility may be expected to matter for asset prices, since volatility enters the risk premium. In this paper, we establish theoretically and empirically the consequences of long memory in volatility for asset prices.

An important ingredient in the relation between the level of asset prices and the memory properties of volatility is the fundamental risk-return trade-off, i.e., the contemporaneous relation between the conditional mean and variance of returns, which has been the subject of extensive research. Early theoretical and empirical contributions on the functional form of this relation were due to Merton (1973, 1980). The risk-return trade-off has been studied in the time-varying volatility case using GARCH-type models by Engle, Lilien & Robins (1987), Bollerslev, Engle & Wooldridge (1988), Chou (1988), Campbell & Hentschel (1992), Chou, Engle & Kane (1992), Backus & Gregory (1993), and others. A two-factor model including a hedging component was considered by Scruggs (1998), who found that the risk-return component indicated a positive trade-off. Recent work in asset pricing has examined cross-sectional risk premia induced by covariance between innovations in volatility and stock returns. This literature finds negative premia, e.g. Ang, Hodrick, Xing & Zhang (2005). The idea is that

¹Alternative models have been proposed to capture the persistence properties of volatility, e.g. nonlinear models as in Granger & Ding (1996) and Diebold & Inoue (2001), but the long memory model remains the most popular both in theoretical and empirical studies.

since innovations in volatility are higher during recessions, stocks which co-vary with volatility are stocks that pay off in bad states, and these stocks should require a smaller risk premium. Thus, there is mixed evidence on the sign of the risk-return relation in the literature. A negative relation is also found in some GARCH-type models, but Harrison & Zhang (1999) use a seminonparametric density approach and show that GARCH models may be misspecified and lead to an erroneous indication of a negative risk-return relation, and they conclude that the true relation is positive. Recently, a positive risk-return trade-off has been indicated by several studies using efficient high-frequency data volatility measurement, e.g. Ghysels, Santa-Clara & Valkanov (2005), who use weighted rolling sample windows in the variance measurements. Similar conclusions are reached by Brandt & Kang (2004) using a latent VAR methodology. For a survey of these and related studies, see Lettau & Ludvigson (2004).

Common to the risk-return trade-off studies is that they do not incorporate the long memory characteristic of volatility, which has been established in the realized volatility literature. In the latter, realized volatility is shown to be well described by a fractionally integrated or $I(d)$ process, with long memory parameter d in the vicinity of $0.3 - 0.4$, e.g. Andersen, Bollerslev, Diebold & Ebens (2001) and Andersen et al. (2003), whereas the risk-return literature takes volatility to be $I(0)$, as in standard GARCH models.

Among the most important uses of models for expected returns is the discounting of streams of expected future cash flows, and thus the calculation of asset prices. While an important ingredient, the risk-return trade-off is not the only determinant of asset values. Poterba & Summers (1986) elaborate on the asset valuation aspect and derive the way in which both the risk-return trade-off and serial correlation in volatility jointly determine the level of stock prices. Their results show that a stronger positive risk-return relation and higher serial correlation in volatility both contribute towards numerically higher (namely, more negative) elasticity of stock prices with respect to volatility. They argue that effects of shocks to volatility decay rapidly and hence only affect required returns and thereby asset prices over short intervals. Thus, it is difficult to ascribe long term movements in the stock market to volatility changes.

Obviously, this argument was made before the advent of long memory models in financial economics, and is called into question in view of the amassing recent literature on long memory in financial volatility series. The new evidence calls for a reassessment of the relation between the risk-return trade-off, serial dependence in volatility, and the level of asset prices. The

received evidence on the positive risk-return trade-off and manifest long memory in volatility seem to suggest possible long-lasting effects of volatility shocks on required returns, and hence on the level of the stock market. A high elasticity (in absolute terms) of stock prices with respect to volatility would be expected. In the present paper, we examine this possibility. We establish the relation between the risk-return trade-off, serial dependence in volatility and the elasticity of asset values with respect to volatility in the presence of long memory in volatility, and we estimate all three empirically. Our volatility measurements are efficient high-frequency realized return volatilities as well as model free implied volatilities from option prices. Our results confirm those from recent literature on both the positive risk-return trade-off and the long memory property of volatility. On the other hand, we show that the consequences for asset values are not as anticipated, based on the foregoing interpretation. Thus, the effect of volatility shocks on stock prices is modest and short-lived, in our empirically supported model, in spite of the positive risk-return trade-off and long memory in volatility.

For the joint analysis of asset returns and realized volatility, we introduce a general simultaneous bivariate system of the form

$$A(L)X_t = B(L)u_t, \tag{1}$$

where X_t is a vector containing the data on return and volatility in period t , $A(L)$ and $B(L)$ are lag polynomials, and u_t are the system errors. We derive and test parametric restrictions on the general vector autoregressive moving average (VARMA) system (1) implied by asset pricing theory. The restrictions are derived under alternative specifications for both the short and long memory properties of the volatility process and the functional form of the central risk-return trade-off. An additional important and well documented empirical feature in financial time series is the financial leverage effect, see e.g. Black (1976), Engle & Ng (1993), and Yu (2005). The standard argument from Black (1976) is that bad news decreases the stock price and hence increases the debt-to-equity ratio (i.e. financial leverage), making the stock riskier and increasing future expected volatility. This generates a negative relationship between volatility and past returns, and our framework accommodates this as well. Our empirical application to Standard & Poor's 500 stock market index returns and high-frequency realized volatilities as well as VIX implied volatilities over a 15 year period confirms that volatility exhibits long memory, but equity returns do not, and the risk-return trade-off is significantly positive. Fur-

thermore, the cross-equation restrictions implied by asset pricing theory on the joint model of returns and volatility in (1) are not rejected by the data in some of our specifications, and point to an important leverage effect.

We generalize the Poterba & Summers (1986) analysis of the elasticity of asset values with respect to volatility shocks to allow for both short and long memory in volatility. Briefly reviewing their treatment of the short memory case, the starting point is the standard requirement that stock prices P_t satisfy

$$\frac{E_t(P_{t+1}) - P_t}{P_t} + \frac{D_t}{P_t} = r_f + \alpha_t, \quad (2)$$

where r_f is the riskless rate of interest, assumed constant, α_t the equity risk premium, and D_t the dividend paid at time t . The Merton (1973, 1980) linear relation between the equity risk premium and volatility² σ_t^2 of equity returns is employed, viz.

$$\alpha_t = \gamma \sigma_t^2. \quad (3)$$

Thus, the sign and magnitude of the risk-return trade-off is given by the parameter γ . The specific short memory model used is a stationary AR(1) specification for monthly stock market variances,

$$\sigma_{t+1}^2 = \rho_0 + \rho_1 \sigma_t^2 + \mu_{t+1}. \quad (4)$$

Under these assumptions, the elasticity of the stock market level with respect to volatility is given by

$$\frac{d \log P_t}{d \log \sigma_t^2} = \frac{-\bar{\alpha}}{1 + r_f + \bar{\alpha} - \rho_1(1 + g)}, \quad (5)$$

where $\bar{\alpha}$ is the mean value of α_t and g is the constant growth rate of expected dividends, so that $E_t(D_{t+j}) = (1+g)^j D_t$. Expression (5) is evaluated at the mean values of the risk premium and the dividend yield. For plausible values of the parameters the elasticity varies wildly with ρ_1 , the short memory parameter of the volatility process. For example, a doubling of volatility from its average level reduces the value of the market by only 0.6% if $\rho_1 = 0$, but by 38.7% if $\rho_1 = .99$. For the empirical estimates of ρ_1 , the drop is of the order 1.4% – 2.2%, and in alternative AR(12) and IMA(1,3) models up to 22.5%.

Expression (5) is a convenient representation of the fact that the risk-return trade-off and serial dependence in volatility have separate impacts on asset prices. An increase in either

²Henceforth, we use the terms volatility and variance synonymously.

γ (and hence $\bar{\alpha}$, through (3)) or ρ_1 increases the magnitude of the (negative) elasticity. The recent empirical findings of a significantly positive risk-return trade-off ($\gamma > 0$ in the present context) and long memory in volatility might now be expected to translate into an asset price elasticity of great magnitude. However, the derivation of (5) was based on AR(1) volatility. Hence, motivated by the recent empirical literature, we generalize the framework to allow for both short and long memory in volatility. In particular, we allow for long memory of the fractional integration or $I(d)$ type. As the IMA(1,3) model implies nonstationarity of the standard $I(1)$ unit root type, while the AR(1) model (of class $I(0)$) approaches the unit root model as ρ_1 tends to one, it might be expected that fractional integration of volatility would lead to elasticity estimates in the same range as the empirical estimates quoted above, between 1.4% and 22%. However, perhaps surprisingly, our theoretical and empirical results point towards a true elasticity in the presence of long memory in volatility probably not exceeding 1%.

It is important to note the structure of the model, in particular that σ_t^2 is the conditional variance process, as is common in long memory stochastic volatility models (see e.g. Breidt et al. (1998) and Robinson (2001)), even though it enters the mean specification (2) explicitly. Thus, writing r_t for the return over period t ,

$$r_t = \frac{P_{t+1} + D_t - P_t}{P_t},$$

returns are conditionally heteroskedastic with a time-varying risk premium,

$$r_t = r_f + \alpha_t + \sigma_t e_t, \tag{6}$$

where e_t is an \mathcal{F}_t -martingale difference in the mean and the variance, i.e., $E(e_t | \mathcal{F}_{t-1}) = 0$ and $E(e_t^2 | \mathcal{F}_{t-1}) = 1$, and it is assumed that σ_t is measurable with respect to \mathcal{F}_{t-1} , the information set at time $t - 1$. Throughout the paper, we consider risk premium specifications of the type

$$\alpha_t = g(\sigma_t^2, \sigma_{t-1}^2, \dots), \tag{7}$$

where $g(\cdot)$ is a well-behaved function, and since σ_t^2 is measurable with respect to \mathcal{F}_{t-1} , so is the risk premium α_t . It follows that

$$\begin{aligned} E(r_t | \mathcal{F}_{t-1}) &= \alpha_t + \sigma_t E(e_t | \mathcal{F}_{t-1}) = \alpha_t, \\ E(r_t^2 | \mathcal{F}_{t-1}) &= \alpha_t^2 + \sigma_t^2 E(e_t^2 | \mathcal{F}_{t-1}) = \alpha_t^2 + \sigma_t^2, \end{aligned}$$

and the conditional variance of returns is

$$\text{Var}(r_t | \mathcal{F}_{t-1}) = E(r_t^2 | \mathcal{F}_{t-1}) - (E(r_t | \mathcal{F}_{t-1}))^2 = \sigma_t^2.$$

Hence, σ_t^2 is indeed the conditional variance of the returns. This is a conditional heteroskedasticity-in-mean type specification similar to Engle et al. (1987) and Bollerslev et al. (1988).

In Section 2 below, we derive the relevant elasticity of the market level with respect to volatility under long memory in σ_t^2 . We show that the elasticity is smaller in magnitude than earlier estimates, and much more stable under variations in the long memory parameter than in the short memory case. We also consider a new specification of the equity risk premium consistent with the stylized fact that equity returns do not exhibit long memory, as the simple specification (3) would require under long memory in σ_t^2 , and the relevant elasticity remains small with this specification. In Section 3, we introduce our new bivariate modelling strategy, where volatility and equity returns are modelled jointly in the VARMA system (1). We derive testable cross-equation restrictions on the system from asset pricing theory. Section 4 presents our empirical analysis, using monthly stock market returns along with high-frequency realized volatilities and model free VIX implied volatilities. We find evidence in favor of our VARMA specification for volatilities and returns, with a positive risk-return relation and long memory in volatility as well as a strong leverage effect, consistent with recent literature. However, in spite of the latter characteristics, our theoretical and empirical results show that the impact of volatility shocks on stock prices is small, with an elasticity probably not exceeding 1%, and short-lived, dying out after about six months. Section 5 concludes.

2 Consequences of Long Memory in Volatility

To model the long memory in volatility, we consider the class of autoregressive fractionally integrated moving average (ARFIMA) processes introduced by Granger & Joyeux (1980) and Hosking (1981). For excellent surveys on long memory processes and fractional models, see Robinson (1994, 2003) and Baillie (1996), and for a textbook treatment, see e.g. Beran (1994). A process is labelled an ARFIMA(p, d, q) process if its d 'th difference is a stationary and invertible ARMA(p, q) process. Here, d may be any real number and if $-1/2 < d < 1/2$ the process is stationary and invertible which we assume throughout. For a precise statement, x_t

is an ARFIMA(p, d, q) process if

$$\phi(L) \Delta^d x_t = \theta(L) \mu_t, \quad (8)$$

where $\phi(L) = 1 - \sum_{i=1}^p \phi_i L^i$ and $\theta(L) = 1 + \sum_{i=1}^q \theta_i L^i$ are polynomials of order p and q in the lag operator L ($Lx_t = x_{t-1}$), with roots strictly outside the unit circle, μ_t is a martingale difference sequence with conditional variance ω_μ^2 , and the fractional filter $\Delta^d = (1 - L)^d$ is defined by its binomial expansion

$$(1 - L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} L^j, \quad \Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt. \quad (9)$$

The parameter d determines the (long) memory of the process. If $d > -1/2$ the process is invertible and possesses a linear (Wold) representation, and if $d < 1/2$ it is covariance stationary. If $d = 0$ the spectral density is bounded at the origin and the process has only weak dependence (short memory). Furthermore, if $d < 0$ the process is said to be anti-persistent, and has mostly negative autocorrelations, but if $d > 0$ the process is said to have long memory, since the autocorrelations die out at a hyperbolic rate (and are not summable), in contrast to the much faster exponential rate of decline of the autocorrelations in the short memory case. Throughout this paper, we shall be concerned mainly with the stationary long memory case $0 \leq d < 1/2$. This interval is relevant for many applications in finance, e.g. Lobato & Velasco (2000), Andersen et al. (2001), and Andersen et al. (2003). In particular, it is the empirically relevant region for the volatility processes that we study.

To estimate the parameters of the model (8) we use the maximum likelihood procedure of Sowell (1992). The maximum likelihood estimator is \sqrt{T} -consistent and asymptotically normal. For details on the asymptotic normal distribution and the associated regularity conditions, we refer the reader to Sowell (1992). The estimator is asymptotically efficient in the classical sense when the model is correctly specified and Gaussian.

The well documented long memory property of volatility suggests replacing the Poterba & Summers (1986) short memory volatility process (4) by

$$\phi(L) \Delta^d (\sigma_t^2 - \bar{\sigma}^2) = \theta(L) \mu_t, \quad (10)$$

where $\bar{\sigma}^2$ denotes the unconditional mean of σ_t^2 , but retaining their other assumptions. In particular, we maintain relations (2) and (3). This allows isolating the effect of long memory in volatility on the level of stock prices. We have the following.

Proposition 1 *If stock return volatility is governed by the general ARFIMA(p, d, q) process (10), then the elasticity of the stock market level with respect to volatility, at the mean values of the risk premium and the dividend yield, is given by*

$$\frac{d \log P_t}{d \log \sigma_t^2} = \sum_{j=0}^{\infty} \psi_j \frac{-\bar{\alpha}(1+g)^j}{(1+r_f+\bar{\alpha})^{j+1}}, \quad (11)$$

where ψ_j is the j 'th impulse-response of σ_t^2 .

This proposition shows that the simple form (5) of the elasticity is special and a consequence of imposing the short memory AR(1) process (4) on volatility. The general form of the elasticity from Proposition 1 is an infinite weighted sum of terms similar to the elasticity (5), with weights given by the impulse-responses ψ_j , which are simply the coefficients in the infinite-order moving average representation of volatility,

$$\sigma_t^2 - \bar{\sigma}^2 = \sum_{j=0}^{\infty} \psi_j \mu_{t-j}. \quad (12)$$

Square summability of the impulse-responses follows from stationarity of the AR polynomial $\phi(\cdot)$ and since $d < 1/2$.

Of course, higher order short memory ARMA(p, q) models for volatility, instead of the simple AR(1), do not in general reproduce the summable geometric series and hence closed form solution for the elasticity (5). Nonetheless, we include long memory from the outset, due to the recent empirical evidence favoring this property.

Within the long memory class there are simplifications in relevant special cases. In particular, when $p = q = 0$, the impulse-responses are given by $\psi_j = \Gamma(j+d)/(\Gamma(d)\Gamma(j+1))$, and we have the following.

Corollary 2 *In the situation from Proposition 1, the special case of an ARFIMA(0, d , 0) or fractionally integrated noise process for volatility yields the elasticity*

$$\frac{d \log P_t}{d \log \sigma_t^2} = \sum_{j=0}^{\infty} \frac{-\bar{\alpha}(1+g)^j \Gamma(j+d)}{(1+r_f+\bar{\alpha})^{j+1} \Gamma(d) \Gamma(j+1)}.$$

It is worth investigating the quantitative economic consequences of these changes in asset price elasticities. Poterba & Summers (1986) evaluate the short memory elasticity (5) for the

representative (monthly) parameter values $r_f = .035$, $\bar{\alpha} = .006$, and $g = .00087$. They find that the elasticity varies from $-.006$ to $-.387$ as ρ_1 varies from 0 to .99. In comparison, for the ARFIMA(0, d , 0) volatility specification from Corollary 2 and using the same values for r_f , $\bar{\alpha}$, and g , we get the long memory elasticities shown in Table 1. Clearly, these are less variable than the short memory elasticities. Even as the memory parameter d approaches 1/2, the boundary for stationarity, the elasticity is only about $-.03$, an order of magnitude less than the short memory elasticities near the boundary for stationarity of the AR(1) given by $\rho_1 = 1$. Given the mounting empirical evidence favoring long memory in volatility, this suggests that the high elasticities reported earlier should be interpreted with considerable caution. More generally, the results indicate that the way in which volatility is entered in the model is crucial and should be considered carefully.

Table 1 about here

Inspection of relation (3) shows that in fact the memory properties of the volatility process carry over to the stock return process through the risk premium link. This is cause of concern since it implies that if volatility is stationary fractionally integrated, as empirical literature demonstrates, then so are stock returns, and indeed, the two are stationary fractionally cointegrated (see Christensen & Nielsen (2005)). Such strong serial dependence in stock returns seems empirically highly implausible, thus calling specification (3) into question. Of course, as finance theory suggests, a link between risk and return should be allowed. The problem with the tight link (3) is that the entire movement in the deviation from the mean $\sigma_t^2 - \bar{\sigma}^2$ in volatility spills over fully into risk premium deviations, $\alpha_t - \bar{\alpha} = \gamma(\sigma_t^2 - \bar{\sigma}^2)$, including long memory components, which is not empirically warranted. Hence, an alternative form of the general link (7) is needed, and we consider the specification

$$\alpha_t - \bar{\alpha} = \gamma \Delta^d (\sigma_t^2 - \bar{\sigma}^2), \tag{13}$$

where the fractional filter $\Delta^d = (1 - L)^d$ removes spill-over of the long memory components into the risk premium, while still permitting a risk-return trade-off. This specification is in the spirit of Ang et al. (2005), i.e. only the short memory component of volatility impacts expected returns. If increases in this component primarily occur in bad states and returns co-vary with volatility, then potentially the required premium is negative in bad states, and

this is accommodated through the possibility of a negative value of γ . Lacking a model for the conditional mean of volatility, Ang et al. (2005) regress changes in VIX on stock returns to construct a factor mimicking innovations in market volatility, and similarly insert the innovations in a linear risk-return specification. Our long memory model (10) in effect provides the model for the conditional mean of volatility and hence allows directly identifying the short memory component of volatility, for use in (13).

With the new specification (13), the elasticity of the stock market level with respect to volatility is again given by (11), but now with the terms ψ_j given by the impulse-responses of $\Delta^d(\sigma_t^2 - \bar{\sigma}^2)$, i.e. an ARMA(p, q) process. For example, if volatility exhibits long memory of the ARFIMA(1, d , 0) type, then the resulting elasticity is again of the original form (5). Thus, what is important in assessing the relevance of the latter form of the stock price-volatility elasticity is not only the memory properties of volatility, but the combination of these with the risk premium link.

The specification of the risk premium link (13) still imposes ARMA properties on α_t if volatility is governed by the ARFIMA process (10) with p or q positive. The same problem arises in the original setup, with specifications (3) and (4), since α_t is restricted to follow an AR(1) process in that case, and with the same memory parameter ρ_1 as volatility. Even ARMA memory may be too much to force upon returns, and a martingale difference scheme for returns would seem a competitive alternative, based on casual empiricism. Hence, in the following we also consider the specification

$$\alpha_t - \bar{\alpha} = \gamma\mu_t, \tag{14}$$

that is, it is exactly the innovation to volatility from (10) that impacts the current risk premium and generates the risk-return relation, c.f. Ang et al. (2005). By the fundamental structure of the stock return equation (2), the risk premium in (14) guarantees that the stock return (dividend included) is a martingale difference sequence. Whether this specification is preferred now becomes an empirical question. Under (14), the elasticity of the stock market level with respect to volatility is simply the constant $-\bar{\alpha}$, i.e. minus the average risk premium. Using the same parameter value as in Table 1, this is $-.006$, again much smaller in magnitude than the elasticities from the short memory model for positive ρ_1 .

Of course, testable cross-restrictions are implied by the risk premium links (3), (13), and (14), since volatility and risk premia share the same innovation sequence. This is explored

below.

3 Bivariate Modelling Strategy

In this section, we consider explicitly the bivariate time series model for returns and volatility. Rewriting (2) explicitly in terms of the expectation error $\varepsilon_t = \sigma_t e_t$, we have

$$r_t = r_f + \alpha_t + \varepsilon_t. \quad (15)$$

We write ω_ε^2 for $var(\varepsilon_t)$ and $\omega_{\varepsilon\mu}$ for $cov(\varepsilon_t, \mu_t)$, with μ_t the innovation in the volatility equation (4). The model based on (3) and (4) then in fact implies the bivariate structure

$$\begin{bmatrix} r_t - \bar{r} \\ \sigma_t^2 - \bar{\sigma}^2 \end{bmatrix} = \begin{bmatrix} 0 & \gamma\rho_1 \\ 0 & \rho_1 \end{bmatrix} \begin{bmatrix} r_{t-1} - \bar{r} \\ \sigma_{t-1}^2 - \bar{\sigma}^2 \end{bmatrix} + u_t, \quad (16)$$

where $\bar{r} = r_f + \gamma\bar{\sigma}^2$ and $\bar{\sigma}^2 = \rho_0/(1 - \rho_1)$ are the unconditional means of the return and volatility processes, respectively, and the vector $u_t = (u_{1t}, u_{2t})'$ of system errors is given by

$$u_t = \begin{bmatrix} 1 & \gamma \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \mu_t \end{bmatrix}, \quad var(u_t) = \Omega_u = \begin{bmatrix} \omega_\varepsilon^2 + 2\gamma\omega_{\varepsilon\mu} + \gamma^2\omega_\mu^2 & \gamma\omega_\mu^2 + \omega_{\varepsilon\mu} \\ \gamma\omega_\mu^2 + \omega_{\varepsilon\mu} & \omega_\mu^2 \end{bmatrix}. \quad (17)$$

Note that (16) is a first order vector autoregressive (VAR(1)) model with two testable zero restrictions. Furthermore, from (17) the errors u_{1t} and u_{2t} are correlated (through the risk premium parameter γ) even when ε_t and μ_t are not. Indeed, imposing $\omega_{\varepsilon\mu} = 0$ generates a testable cross-equation restriction linking mean and variance parameters (the ratio of the non-zero elements in the autoregressive coefficient matrix from (16) equals the ratio of $cov(u_{1t}, u_{2t})$ to $var(u_{2t})$ from (17)). Whether or not $\omega_{\varepsilon\mu} = 0$, efficient estimation of this model would require simultaneous estimation of both equations, since the system errors in u_t are contemporaneously correlated in any case.

It is important to note that even though the unconditional variance of returns $var(r_t) = (1 - \gamma^2\rho_1^2)^{-1}(\omega_\varepsilon^2 + 2\gamma\omega_{\varepsilon\mu} + \gamma^2\omega_\mu^2)$ is constant in the model, it is the time-varying $\sigma_t^2 = var(r_t|\mathcal{F}_{t-1})$ which is modelled in the second equation of the system (16). In our empirical work below, we use either the VIX model free implied volatility for σ_t^2 , or realized volatility estimated by a sum of squared intra-period returns (see Section 4). Our estimation procedure uses the constant unconditional variance Ω_u from (17) for the VAR system errors, which entails

a potential efficiency loss relative to explicitly allowing for time-varying conditional variance of u_{1t} in the estimation, but retains consistency and asymptotic normality of the parameter estimates and is much simpler to implement and interpret, e.g. in terms of the above parameter restrictions.

The VAR(1) model (16) is not empirically realistic, since it imposes short memory on volatility, contradicting findings of long memory in volatility from the literature. Long memory in volatility may be entered through the ARFIMA(p, d, q) process (10). This yields the alternative triangular fractionally cointegrated system

$$r_t - \bar{r} = \gamma(\sigma_t^2 - \bar{\sigma}^2) + \varepsilon_t \quad (18)$$

$$\Delta^d(\sigma_t^2 - \bar{\sigma}^2) = \phi^{-1}(L)\theta(L)\mu_t \quad (19)$$

where \bar{r} and $\bar{\sigma}^2$ continue to denote the unconditional means, although these may now be different from before. The problem with this model is that it imposes long memory in the return equation (18), which is unwarranted empirically.

As a consequence of the empirical problems arising in the models (16) and (18)-(19) (short memory in volatility resp. long memory in returns), we turn to alternative specifications, where (13) or (14) is substituted for (3) to produce a better empirically founded return equation. An additional important and well documented empirical feature in financial time series is the financial leverage effect, see Black (1976), Engle & Ng (1993), and Yu (2005). The standard argument from Black (1976) is that bad news decreases the stock price and hence increases the debt-to-equity ratio (i.e. financial leverage), making the stock riskier and increasing future expected volatility. This generates a negative relationship between volatility and the lagged return. Thus, we consider either (10) for the volatility process or the specification

$$\phi(L)\Delta^d(\sigma_t^2 - \bar{\sigma}^2) = \lambda r_{t-1} + \theta(L)\mu_t, \quad (20)$$

where λr_{t-1} accommodates the financial leverage effect. The implications for the resulting bivariate systems are explored in the following propositions.

Proposition 3 *Let $X_t = (r_t - \bar{r}, \Delta^d(\sigma_t^2 - \bar{\sigma}^2))'$. Then the bivariate system for returns and volatilities is given by the vector ARMA (VARMA) model*

$$A(L)X_t = B(L)u_t, \quad (21)$$

where u_t is given by (17), and the risk premium is given by (13).

A) *No leverage effect:* Assume stock return volatility is governed by the ARFIMA(p, d, q) process (10). Then the lag polynomials $A(L) = I_2 - \sum_{i=1}^p A_i L^i$ and $B(L) = I_2 + \sum_{i=1}^q B_i L^i$ satisfy the restrictions

$$A_i = \begin{bmatrix} 0 & \gamma\phi_i \\ 0 & \phi_i \end{bmatrix}, \quad i = 1, \dots, p,$$

$$B_i = \begin{bmatrix} 0 & \gamma\theta_i \\ 0 & \theta_i \end{bmatrix}, \quad i = 1, \dots, q,$$

where ϕ_i and θ_i are defined as in (20). In particular, there are $2(p+q)$ zero restrictions and $\max(p+q-1, 0)$ cross-equation restrictions. If, in addition, we assume $\omega_{\varepsilon\mu} = 0$, then there is one additional cross-equation restriction when $p+q \geq 1$. Finally, γ is not identified if $\omega_{\varepsilon\mu} \neq 0$ and $p = q = 0$.

B) *Leverage effect included:* Assume stock return volatility is governed by the levered ARFIMA(p, d, q) process (20). Then the lag polynomials $A(L) = I_2 - \sum_{i=1}^{\max\{1,p\}} A_i L^i$ and $B(L) = I_2 + \sum_{i=1}^q B_i L^i$ satisfy the restrictions

$$A_i = \begin{bmatrix} \gamma\lambda 1(i=1) & \gamma\phi_i 1(p \neq 0) \\ \lambda 1(i=1) & \phi_i 1(p \neq 0) \end{bmatrix}, \quad i = 1, \dots, \max\{1, p\},$$

$$B_i = \begin{bmatrix} 0 & \gamma\theta_i \\ 0 & \theta_i \end{bmatrix}, \quad i = 1, \dots, q,$$

with $1(\cdot)$ the usual indicator function. There are $2(1(p=0) + \max\{0, p-1\} + q)$ zero restrictions and $p+q$ cross-equation restrictions. If we assume $\omega_{\varepsilon\mu} = 0$, then there is one additional cross-equation restriction.

The simultaneous system approach for asset returns and volatility behind this proposition is novel and the main contribution of this paper. Far more information can be obtained from the system than from analysis of each equation in isolation. Efficient estimation and inference requires a system approach such as full information maximum likelihood. Note that by construction X_t has zero mean. The number of empirically testable restrictions from asset pricing theory is considerable. In part A with no leverage, the total number of parameters in the conditional mean equations of the corresponding unrestricted VARMA(p, q) model is

$4(p + q)$. When leverage is included in part B, the total number of parameters in the unrestricted VARMA($\max\{1, p\}, q$) model is $4(\max\{1, p\} + q)$. To illustrate the restrictions in this case, consider the special case where stock return volatility is governed by an ARFIMA(0, d , 1) and the financial leverage effect is included. The bivariate process for returns and volatility is given by the VARMA(1,1) model

$$\begin{aligned} r_t - \bar{r} &= \gamma\lambda(r_{t-1} - \bar{r}) + \varepsilon_t + \gamma\mu_t + \gamma\theta\mu_{t-1}, \\ \Delta^d(\sigma_t^2 - \bar{\sigma}^2) &= \lambda(r_{t-1} - \bar{r}) + \mu_t + \theta\mu_{t-1}. \end{aligned}$$

In the associated unrestricted VARMA(1,1) model there are two 2×2 matrices in the mean equations, for a total of eight parameters. Only the first order moving average coefficient θ (corresponding to θ_1 in (20)) and the parameters γ and λ appear in the restricted system above, i.e., there are five constraints. Thus, this count matches the general number of restrictions from Proposition 3B for the case $(p, q) = (0, 1)$.

Although the autoregressive polynomial in the ARFIMA model for volatility is of order $p = 0$ in this example, the autoregressive order of the resulting simultaneous VARMA system is one. This corresponds to the more general result in Proposition 3B, that an autoregression of order p for volatility translates into an autoregressive component in the full system of order $\max\{1, p\}$. All these restrictions are of course empirically testable. Finally, in the specific example considered here, if the additional assumption $\omega_{\varepsilon\mu} = 0$ is invoked, i.e. the expectation revisions in returns are uncorrelated with the innovations to volatility, then the risk premium parameter, γ , may be recovered from the estimated covariance matrix for the (still correlated) system errors u_t defined in (17) as $\gamma = \text{cov}(u_{1t}, u_{2t})/\text{var}(u_{2t})$, and this generates one additional restriction, for a total of six constraints.

Both the A_i and B_i coefficient matrices ($i \geq 1$) exhibit non-zero terms in the first row (the return equation), even in the restricted form of the model in Proposition 3. This degree of serial dependence in asset returns is potentially in conflict with empirical evidence, as well as efficient markets theory. Hence, for the empirical analysis, it is worthwhile to compare with a version based on the alternative risk premium specification (14).

Proposition 4 *Let the risk premium be given by (14). Then the bivariate system for returns and volatilities is given by the VARMA model (21).*

A) *No leverage effect: Assume stock return volatility is governed by the ARFIMA(p, d, q) process (10). Then the lag polynomials $A(L) = I_2 - \sum_{i=1}^p A_i L^i$ and $B(L) = I_2 + \sum_{i=1}^q B_i L^i$ satisfy the restrictions*

$$A_i = \begin{bmatrix} 0 & 0 \\ 0 & \phi_i \end{bmatrix}, \quad i = 1, \dots, p,$$

$$B_i = \begin{bmatrix} 0 & 0 \\ 0 & \theta_i \end{bmatrix}, \quad i = 1, \dots, q,$$

with ϕ_i, θ_i defined as in (10). In particular, there are $3(p+q)$ zero restrictions.

B) *Leverage effect included: Assume stock return volatility is governed by the levered ARFIMA(p, d, q) process (20). Then the lag polynomials $A(L) = I_2 - \sum_{i=1}^{\max\{1,p\}} A_i L^i$ and $B(L) = I_2 + \sum_{i=1}^q B_i L^i$ satisfy the restrictions*

$$A_i = \begin{bmatrix} 0 & 0 \\ \lambda 1(i=1) & \phi_i 1(p \neq 0) \end{bmatrix}, \quad i = 1, \dots, \max\{1, p\},$$

$$B_i = \begin{bmatrix} 0 & 0 \\ 0 & \theta_i \end{bmatrix}, \quad i = 1, \dots, q,$$

with $1(\cdot)$ the usual indicator function and ϕ_i, θ_i defined as in (20). Thus, there are $3(\max\{1, p\} + q) + 1(p=0) - 1$ zero restrictions.

Finally, in both parts A and B, if $\omega_{\varepsilon\mu} \neq 0$ the parameter γ is not identified, and if $\omega_{\varepsilon\mu} = 0$ then $\gamma = \text{cov}(u_{1t}, u_{2t})/\text{var}(u_{2t})$.

With this specification, there are $4(p+q)$ VARMA parameters in the unrestricted model in part A and $4(\max\{1, p\} + q)$ in part B. The restricted coefficient matrices A_i and B_i have zeros in the first row (the return equation) for $i \geq 1$. Hence, the process for asset returns is a martingale difference sequence in the restricted model. This potentially accords better with data than the model from Proposition 3, a possibility explored in the empirical work below.

Before turning to the empirical analysis, we briefly return to the simple example from above, where volatility is governed by an ARFIMA(0, d ,1) model and financial leverage is included. In this case, with $q = 1$, the restricted simultaneous VARMA(1,1) system from part B of

Proposition 4 is given by

$$\begin{aligned} r_t - \bar{r} &= \varepsilon_t + \gamma\mu_t, \\ \Delta^d(\sigma_t^2 - \bar{\sigma}^2) &= \lambda(r_{t-1} - \bar{r}) + \mu_t + \theta\mu_{t-1}. \end{aligned}$$

Even though returns form a martingale difference sequence, the system errors remain contemporaneously correlated, and simultaneous estimation is called for. This is so even if $\omega_{\varepsilon\mu} = 0$, i.e. if expectation revisions in returns are uncorrelated with innovations to volatility. In this case γ may be recovered as stated in Proposition 4. Note that there are six testable zero restrictions on the VARMA parameters in this example.

The various models above are explored empirically in the following section.

4 Empirical Analysis

We use 5-minute returns on the S&P 500 stock market index using linear interpolation following Müller, Dacorogna, Olsen, Pictet, Schwarz & Morgenegg (1990), Dacorogna, Müller, Nagler, Olsen & Pictet (1993), and Barucci & Reno (2002), among others, to form monthly realized volatilities (annualized sample sum-of-squares), resulting in roughly 2,000 intra-monthly return observations for each realized volatility (97 per day and approximately 20 trading days per month). The returns data cover the period January 1, 1988, to December 31, 2002. For more details on the construction of the realized volatility data, see Andersen, Bollerslev & Diebold (2005) or Andersen, Bollerslev, Diebold & Vega (2004). The use of 5-minute returns can be justified on the basis of bias considerations when dealing with market microstructure noise induced effects which are expected to be small at the 5-minute frequency, see e.g. the simulations by Nielsen & Frederiksen (2004). Indeed, much recent work has been devoted to integrated variance estimation in the presence of noise, e.g. Bandi & Russell (2005*a*, 2005*b*), Hansen & Lunde (2004), Oomen (2005*a*, 2005*b*), Zhang, Mykland & Ait-Sahalia (2003) – Bandi & Russell (2005*c*) and Barndorff-Nielsen & Shephard (2005) review this literature. There is some evidence that the optimal (in a mean squared error sense) sampling frequency is higher than 5 minutes when estimating daily realized volatilities. However, we consider monthly observations on realized volatility, hence employing roughly 2,000 return observations for each realized volatility. Consequently, we expect our monthly realized volatility measures to be very precise and not contaminated by market microstructure induced bias.

In addition, we use the monthly nominal dividend series for the index to form monthly growth rates g_t^{nom} . For the riskfree rate (r_{ft}^{nom}), we use monthly data on the 3-month Treasury bill rate (secondary market, middle rate), and we use the growth rate in the consumer price index (nondurables) for the inflation series (π_t). These data and the one-month nominal returns (r_t^{nom}) are obtained from Datastream. As Poterba & Summers (1986), we also compare with results where the monthly volatility series is based on implied volatilities from option prices. We use the model free (new) VIX implied volatility series from the Chicago Board of Options Exchange (CBOE). The new VIX is based on Britten-Jones & Neuberger (2000) and measures expected volatility over a 30-day period by averaging the weighted prices of out-of-the money puts and calls. Unlike the old VIX (now VXO), the new VIX is based on the S&P 500 index and is available from January 1990.³ Our sample covers the period January 1990 to February 2005, and all variables are stated on a monthly basis.

Table 2 about here

Summary statistics are in Table 2. Panel A shows statistics for the data set covering the period 1988:1-2002:12 (180 observations), for which we have high-frequency realized volatility data for σ_t^2 . Panel B shows statistics for the data set covering the period 1990:1-2005:2 (182 observations), for which we use VIX implied volatilities for σ_t^2 . As the standard deviations of returns (first column) are about equal in the two panels, the higher average volatility (last column) in Panel B is due to the use of implied instead of realized volatility, perhaps reflecting a negative volatility risk premium in option prices. Furthermore, the standard deviation of volatility is considerably lower in Panel B (.17 versus .25), consistent with the notion that implied volatility measures expected future volatility which should be less volatile than actual realized volatility. In both data sets, average returns are considerably higher than interest rates, which in turn exceed inflation. As expected, the standard deviations of returns far exceed those of the riskfree rates.

Table 3 about here

Before proceeding to the empirical analysis of the bivariate systems of Section 3, we examine the univariate time series properties of the two individual volatility series (realized and implied),

³ Ang et al. (2005) use the old VIX which is based on the S&P 100 index, but our high-frequency realized volatilities are for the S&P 500 index, so we restrict attention to the new VIX.

using the ARFIMA(p, d, q) model (10). Table 3 shows the results of the univariate analysis. The results for the realized variance data set, from 1988-2002, are in Panel A, and results for the VIX implied variance data set, from 1990-2005, are in Panel B. The reported estimates of the long memory parameter d and the autoregressive and moving average parameters ϕ and θ are found by application of the Sowell (1992) maximum likelihood method to deviations from the sample mean, using the arfima package v1.01 for PcGive 10.1, see Doornik (2001) and Doornik & Ooms (2001). Asymptotic standard errors are in parentheses. Also reported are log-lik, the value of the maximized log-likelihood function, AIC and BIC, the Akaike and Schwarz information criteria, and Q(12), the Ljung-Box statistic for 12 lags which is asymptotically χ^2 distributed with $12 - 1 - p - q$ degrees of freedom. For the Ljung-Box test, one and two asterisks denote rejection at 5% and 1% significance level, respectively. Across both panels, the long memory parameter d is very strongly significant and estimated to values between .3 and .5, except a single case in Panel A where a negative d is accompanied by a near unit root in the autoregressive polynomial.

In Panel A, the realized variance data set, the $(0, d, 3)$ model gets the best (lowest) AIC criterion and the highest likelihood value, the third-order MA parameter θ_3 gets a t -statistic of 2.78, and the Ljung-Box statistic is by far the lowest in the panel. In principle, the likelihood ratio test does not reject reduction in the model, and the residuals actually appear serially uncorrelated in some of the other models estimated, but we hesitate to discard a statistically significant parameter (θ_3). This is even more so because estimators in oversimplified models are inconsistent, whereas the only cost of overparametrization is inefficiency, but not inconsistency. We thus prefer the ARFIMA($0, d, 3$) model for realized volatilities. We also estimated higher order models (results not reported), but they were clearly inferior to the chosen model.

In Panel B, the implied variance data, model selection is less obvious. The $(1, d, 0)$ model gets the best (lowest) SIC criterion and the second best AIC in the panel. The estimated first-order AR coefficient ϕ_1 is estimated to .45, with a t -ratio of 3.27, and the Ljung-Box statistic shows no sign of misspecification. From the table, adding parameters such as ϕ_2 or θ_1 yields t -tests below one. Based on these observations we select the ARFIMA($1, d, 0$) model for VIX implied volatilities.

In the chosen $(0, d, 3)$ and $(1, d, 0)$ models, the long memory parameter d takes the values .30 and .31 and is strongly significant. Hence, the results in Table 3 confirm the long memory

property of volatility found in the literature, and our estimates of the long memory parameter are in line with those in e.g. Andersen et al. (2001) and Andersen et al. (2003).⁴ In particular, they are well within the stationary long memory region $0 < d < 1/2$.

Table 4 about here

The long memory property of volatility changes the implications for the level of the stock market. We have seen the first indication of this in Table 1, for the simple case of a $(0,d,0)$ model for volatility. We now examine this issue in more detail for the empirically relevant volatility models from Table 3. Table 4 shows the parameters needed for the comparison. The first row of the table gives the Poterba & Summers (1986) values for the real risk premium, riskfree rate and dividend growth rate. The subsequent rows show the corresponding values for our two data sets, estimated from the averages in Table 2.

Figure 1 about here

The parameters from Table 4 may now be used with Proposition 1 to assess the stock market elasticity with respect to volatility changes. To this end, we also need the impulse-responses ψ_j for the volatility process. These are exhibited in Figure 1, for each of our two data sets. Initially, we use the impulse-response functions for the estimated ARFIMA models for σ_t^2 (using realized resp. implied volatility and shown as solid lines), but the figure also shows the impulse-responses for the corresponding ARMA models, i.e. for $\Delta^d \sigma_t^2$ (shown as dotted lines), which we will need below.

Table 5 about here

Table 5 shows the resulting elasticities based on Proposition 1. Panel A shows elasticities using the parameter values from the first row of Table 4, and Panel B shows elasticities using the parameter estimates from our data (last two rows of Table 4). In each panel of Table 5,

⁴We also estimated the same models for realized and implied standard deviations, as opposed to variances, following Andersen et al. (2001) and Andersen et al. (2003), and got slightly higher d -values, much in line with results in these studies. Applying the alternative semiparametric d estimator of Geweke & Porter-Hudak (1983) using bandwidth $m = 20$ to the variance measurements yields $\hat{d} = 0.40$ for VIX and $\hat{d} = 0.29$ for realized, both with a standard error of .18.

elasticities are shown for the preferred model for each of our two data sets. The first column of elasticities is based on the ARFIMA impulse-responses from Figure 1 and the risk premium link (3). All four estimated elasticities are of the same order of magnitude as those reported for short memory autoregressive models by Poterba & Summers (1986, Table 2) in their empirical section ranging from $-.014$ to $-.048$, whereas their reported elasticities for integrated moving average models are an order of magnitude larger, at $-.175$ to $-.225$. The largest of our elasticity estimates, at $-.047$, is obtained using implied volatility.

As discussed in Section 2, use of the risk premium link (3) imposes long memory on returns, which is not empirically warranted. Thus, the next column of elasticities in Table 5 is based on the alternative risk premium link (13). Calculation of elasticities in this case amounts to substituting the impulse-responses for the corresponding ARMA models (i.e. the models for $\Delta^d \sigma_t^2$) into (11). These impulse-response functions are also exhibited in Figure 1 (dotted lines). In all cases, the resulting elasticities are even lower than in the previous column, although the parameters from Table 4 are unchanged within each row of Table 5. A glance at Figure 1 shows that, as expected, the impulse-responses for each of the ARMA models are much lower than for the associated ARFIMA model for the same volatility series. It is these lower values for ψ_j that feed into (11) and yield lower elasticities.

If we turn to risk premium link (14), thus allowing for martingale difference returns, the elasticity is given simply as $-\bar{\alpha}$, shown in the last column of Table 5. This produces estimates that are slightly lower but of the same order of magnitude as in the second to last column. Thus, both risk premium links that do not force long memory upon returns yield elasticity estimates of $-.01$ or less, i.e., smaller than in the first column of elasticities and smaller than the estimates in Panel A. All in all, our results suggest that when allowing for long memory in volatility, the stock market elasticity with respect to volatility is small, probably no larger than about $-.01$, and smaller than the estimates reported in Poterba & Summers (1986), which in light of the evidence on long memory appear to have been exaggerated.

The elasticities considered so far reveal something, but not everything about the relation between volatility and the level of the stock market. To gain further insight into the dynamics of this relation, we next consider the explicit simultaneous system of changes in both, i.e., the multivariate model for stock market returns and volatility changes. This development is unique to the present paper. As volatility is governed by a fractionally integrated process, the volatility

changes considered in the multivariate system are fractional differences. Thus, we implement in turn the various system specifications of Section 3 for the bivariate process X_t . Here, the components of X_t are the deviations from the unconditional means of the returns r_t and the volatility changes $\Delta^d \sigma_t^2$. The unconditional means are .0090 and .0006, respectively, for the realized volatility data set, and .0085 resp. .0005 for the data set with implied volatilities.

Table 6 about here

We first apply the general unrestricted VARMA model (1) to the data set with realized volatilities. Based on the univariate analysis, the orders of the vector AR and MA polynomials are set at 1 and 3, respectively, i.e., we implement an unrestricted VARMA(1,3) model for $X_t = (r_t - \bar{r}, \Delta^d(\sigma_t^2 - \bar{\sigma}^2))'$. The autoregressive term accommodates the leverage effect from Part B of Propositions 3 and 4. In Part A, the presence of this autoregressive term generates an additional 4 zero restrictions, since $p = 0$ in the preferred univariate model for realized volatility. Estimation results appear in Table 6, Panel A. From Proposition 3, Part A, asset pricing theory implies 8 restrictions on the 16 coefficients in the AR and MA polynomials. In particular, with $(p, q) = (0, 3)$, there are $2(0+3) = 6$ zero restrictions and $\max\{0+3-1, 0\} = 2$ cross-equation restrictions. Adding the 4 zero restrictions on the vector AR matrix yields a total of 12 restrictions. For Part B of Proposition 3, there is one less restriction, due to the inclusion of the additional leverage effect. We also estimate the model as restricted under Part A and Part B, separately, and the results are shown in Panels B and C. In each panel, the first two columns of estimates correspond to the first order AR matrix A_1 , with coefficients for the return equation in the first row, and coefficients for the volatility equation in the second row. The following columns correspond to the MA matrices. The last two columns show the 2×2 variance-covariance matrix (times 1,000) of the VARMA residuals u_t . The estimation method is Gaussian maximum likelihood, and standard errors based on the sum of outer products of the score contributions are in parentheses. Strictly, Gaussianity is not needed for consistency and asymptotic normality of the parameter estimates in stationary VARMA models (Lütkepohl (1991)), and we get similar inferences from using robust (sandwich-type) standard errors.

The unrestricted estimates in Panel A are all insignificant, suggesting that the model is overfitted, but three of the six estimates in the restricted model in Panel B (corresponding to Proposition 3A) are significant. Note that in Panel B, 6 estimates are reported in the 8 columns

of VARMA coefficients, but these correspond to only 4 distinct parameters, due to the 2 cross-equation restrictions. Proposition 3A shows how the restricted parameter estimates relate to the parameters from the univariate analysis in Table 3. For example, the lower right corner of B_i corresponds to θ_i . In particular, the parameter θ_3 that was significant at the 5% level in the univariate analysis remains significant in the multivariate analysis, although the point estimate is somewhat lower than the univariate estimate. In Panel C (restricted model based on Proposition 3B), the lower left corner of A_1 estimates the additional leverage parameter λ , which is negative as expected, at $-.013$, and strongly significant, with a t -statistic of about -4 . This clearly shows the empirical relevance of the leverage effect, inducing an inverse relation between returns and future volatility.

The theory restrictions from Proposition 3 are based on the risk premium link (13). Working within the VARMA framework, we have already abandoned the link (3), which has the empirically unrealistic triangular fractional cointegration implication, see (18)-(19). The possibility remains that returns should be modelled as martingale differences. This is allowed under the alternative risk premium link (14). From Proposition 4A, there are 9 asset pricing theory restrictions on the VARMA coefficients in this case, in addition to the 4 zero restrictions on A_1 . The estimated model thus restricted appears in Panel D of Table 6. Here, 2 of the 3 restricted coefficient estimates are significant. Panel E shows estimates for the restricted model corresponding to Proposition 4B, including the leverage effect, so a total of 12 restrictions are imposed. The estimated leverage effect λ is very similar to that in Panel C, at $-.015$, with a t -statistic in excess of 5 in magnitude. Finally, the estimated residual covariance matrices are nearly identical in all five panels, indicating that not much information is lost by imposing the restrictions.

Table 7 about here

We also carry out the similar analysis using the data set with VIX implied volatilities. Table 7 shows the results. Based on the univariate results, we estimate a VARMA(1,0) model. Unlike in the realized volatility data set, 3 of the 4 coefficients in the unrestricted model (Panel A) are now significant, and the residual covariance matrices for the restricted models without leverage (Panels B and D) are quite different from those of the unrestricted model (Panel A) and the restricted models including the leverage effect (Panels C and E). Proposition 3A imposes 2 zero

restrictions on the model, and Proposition 3B imposes a single cross-equation restriction, so the 4 coefficient estimates in Panel C correspond to only 3 parameters. Proposition 4A imposes 3 zero restrictions, and Proposition 4B imposes 2. Throughout the table, all VARMA coefficients in the restricted models are significant. The point estimates of the leverage coefficient λ are similar to those in Table 6, and the t -statistics now exceed 10 in magnitude.

Table 8 about here

Table 8 shows the results of formal hypothesis testing within the VARMA framework. Panel A shows results for the data set with realized volatilities, and Panel B for the data set with implied volatilities. The first column shows the maximized log likelihood values. The second column shows the likelihood ratio tests of the restrictions from Propositions 3 and 4, respectively, against the appropriate unrestricted model. The degrees of freedom and p -values based on the asymptotic χ^2 -distribution are given in the next two columns. The following three columns show the estimated standard deviations of the expectations revisions and volatility innovations and their correlation. The final two columns show the estimates of the risk-return trade-off γ and the leverage effect λ (asymptotic standard errors in parentheses). Note that γ and the structural parameters related to the expectations revisions and volatility innovations are only reported for the model corresponding to the Proposition 3 restrictions, which is where they are identified without imposing $\omega_{\varepsilon\mu} = 0$ (the estimated correlation between ε_t and μ_t ranges between $-.24$ and $-.82$ in the table). Thus, using Proposition 3, the reported γ -estimate is calculated off the entries in the leading MA coefficient matrix, B_1 , as $\gamma = B_{1,12}/B_{1,22}$ in the data set with realized volatilities. In the data set with implied volatilities, γ is calculated off the entries in the leading AR coefficient matrix, A_1 , as $\gamma = A_{1,12}/A_{1,22}$. From these expressions, asymptotic standard errors are calculated by the delta method.

In the data set with realized volatilities (Panel A), we reject the restrictions from Propositions 3A and 4A at the 5% level, but not at the 1% level. Strikingly, there is no evidence against the models including the leverage effect (Propositions 3B and 4B), which get p -values of 55% and 56%. Consistent with the results in Table 6 (parameter significance and analysis of error covariance), we prefer the VARMA(1,3) model restricted according to Proposition 4B for this data set. The pattern is similar but even more striking in the data set with implied volatilities (Panel B), where the restrictions from Propositions 3A and 4A are overwhelmingly

rejected, whereas the restrictions from Propositions 3B and 4B are not rejected at the 10% level. Thus, the results from both data sets confirm the empirical importance of the leverage effect. We prefer the model corresponding to Proposition 3B for the data set with implied volatilities, since the two VARMA parameters dropped when moving from Panel C to Panel E in Table 7 are significant at conventional levels.

The risk-return trade-off parameter γ from (13) is estimated to values between 4.7 and 21.2 in the two data sets, and gets t -values of 2.70 and 1.94 in the data set with implied volatilities. Based on this evidence, γ appears significantly positive, reflecting a positive risk-return relation. Volatility increases are compensated through increased risk premia. The leverage parameter λ is strongly significant, with t -statistics above 4 in magnitude throughout, showing that the leverage effect is indispensable for the model. This is further reinforced by noting that the point estimates vary very little across specifications, from $-.013$ to $-.015$. Thus, negative return innovations spur volatility increases. In such bad states, stocks that co-vary with volatility pay off, and Ang et al. (2005) find evidence of correspondingly lower required premia for these stocks in the cross-section. Our aggregate time series evidence is that the overall market risk premium is positive, consistent with e.g. Brandt & Kang (2004) and Ghysels et al. (2005).

We end this section by examining further the preferred model for the data set with VIX implied volatilities, subject to the restrictions of Proposition 3B including both a risk-return trade-off and the leverage effect. The VARMA form of this model is expressed in terms of the system errors u_t . Recasting the model in terms of the original return expectation revisions ε_t and volatility innovations μ_t , using (17), we obtain

$$r_t - \bar{r} = \underset{(3.30)}{6.41} \Delta^d(\sigma_t^2 - \bar{\sigma}^2) + \varepsilon_t, \quad (22)$$

$$\Delta^d(\sigma_t^2 - \bar{\sigma}^2) = \underset{(0.02)}{0.48} \Delta^d(\sigma_{t-1}^2 - \bar{\sigma}^2) - \underset{(0.001)}{0.01} (r_{t-1} - \bar{r}) + \mu_t, \quad (23)$$

with structural error covariance matrix

$$\begin{bmatrix} \omega_\varepsilon^2 & \omega_{\varepsilon\mu} \\ \omega_{\varepsilon\mu} & \omega_\mu^2 \end{bmatrix} = \begin{bmatrix} 1668.7 & -4.33 \\ -4.33 & 0.79 \end{bmatrix} \times 10^{-3}. \quad (24)$$

In the estimated version of the restricted model, returns depend positively on volatility changes which in turn depend positively on lagged volatility changes and negatively on lagged returns. The variance of the expectation revision ε_t is more than 1,000 times that of the

volatility shock μ_t . The estimated covariance between the two, $\omega_{\varepsilon\mu}$, exceeds the variance of the volatility shock ω_μ^2 in magnitude. Probably the true covariance is non-zero, implying that it is not appropriate to impose this zero restriction in order to calculate the risk premium parameter γ in the unrestricted and Proposition 3A and 4A models in Table 8. Fortunately, this restriction is not needed for estimation of γ in the preferred Proposition 3B model for the data set with implied volatility, where the estimate is 6.41 and significant.

The risk-return trade-off is thus positive (p -value of 5.19%), and the leverage hypothesis is confirmed through the negative values of both λ and the covariance $\omega_{\varepsilon\mu}$ between innovations to returns and volatility in (24). Obviously, the model implies that returns are not martingale differences. In particular, substituting (23) in (22) shows the dependence of returns on lagged returns and volatility changes.

Figure 2 about here

With the estimated multivariate models in hand, it is possible to give a much more detailed dynamic picture of the relation between volatility and stock prices than that entailed in the simple elasticities considered earlier. In particular, we may solve (22) for the impulse-response function for future returns with respect to a unit shock in current volatility. This is not an ordinary univariate impulse-response function, but a response function for the variable in the first equation of the system with respect to a shock in the variable from the second equation, so our multivariate framework is crucial for this approach. The resulting impulse-response function is shown as the dotted line in Figure 2. The dynamics inherent in the estimated system is evident from the shape of the impulse-response function. This is in contrast to the simpler dynamics in the short memory multivariate system (16) which results from the model (3)-(4). The j 'th impulse-response in the simple system is $\gamma\rho_1^j$, i.e., geometrically declining, whereas our impulse-response function in Figure 2 is clearly non-monotonic, a result of the more complicated multivariate model structure.

To better complement the information contained in the elasticities considered earlier, we may in addition examine the impulse-responses of log-stock prices with respect to volatility shocks. Expressing log-prices as infinite sums of future returns, $\log P_{t+j} = -\sum_{k=0}^{\infty} r_{t+j+k}$, we calculate the impulse-response function of log-prices from the corresponding sums of impulse-responses of returns. The result is shown as the solid line in Figure 2. Note that after the

initial price drop in reaction to the volatility shock, the market reverts back to the original level, to finally stabilize at the original level after around six months. The initial drop is slightly less than the value of γ . These dynamic response patterns reinforce the impression that the true impact of volatility changes on the stock market level, appropriately accounting for long memory in volatility, is quite modest and short-lived.

5 Concluding Remarks

Recent empirical literature documents the presence of a long memory component in volatility. In addition, at least part of the recent literature finds a significantly positive risk-return trade-off. In conjunction, the two would seem to make for a strong and long-lasting effect on asset values of shocks to volatility. We find that on the contrary the instantaneous as well as dynamic impacts of volatility changes on stock prices are modest and short-lived, with an elasticity of stock price with respect to volatility changes of no more than 1%, and the impulse-response function dying out after about six months in our empirically supported model. Thus, the effect is not as first anticipated. We introduce a new VARMA approach which allows for general joint dynamics of returns and volatility, and there is evidence in favor of the theoretical restrictions in our preferred model. The results are consistent with earlier findings of both a positive risk-return trade-off, long memory in volatility, and a strong financial leverage effect, showing in addition that these interact to yield a perhaps surprisingly low impact of volatility shocks on asset values. Our work complements the literatures on realized and implied volatility, the financial leverage effect, and the risk-return relation, and contributes towards an understanding of the roles they play in determining asset values.

Appendix: Proofs

Proof of Proposition 1. By the chain rule for differentiation

$$\frac{dP_t}{d\sigma_t^2} = E_t \left(\sum_{j=0}^{\infty} \frac{dP_t}{d\alpha_{t+j}} \frac{d\alpha_{t+j}}{d\alpha_t} \frac{d\alpha_t}{d\sigma_t^2} \right),$$

where the first term on the right-hand side is the derivative in an infinite order Taylor expansion of P_t around $\bar{\alpha}$, given by equation (6) of Poterba & Summers (1986) as

$$\frac{dP_t}{d\alpha_{t+j}} = \frac{-D_t(1+g)^j}{(1+r_f+\bar{\alpha})^{j+1}(r_f+\bar{\alpha}-g)},$$

the second term is given by the impulse-response (12), i.e.

$$\frac{d\alpha_{t+j}}{d\alpha_t} = \psi_j,$$

and the last term is

$$\frac{d\alpha_t}{d\sigma_t^2} = \gamma$$

by (3). Collecting the terms we find that

$$\frac{d \log P_t}{d \log \sigma_t^2} = \gamma \frac{\sigma_t^2}{P_t} \sum_{j=0}^{\infty} \psi_j \frac{-D_t(1+g)^j}{(1+r_f+\bar{\alpha})^{j+1}(r_f+\bar{\alpha}-g)}.$$

Applying (3) and setting the risk premium α_t and the dividend yield D_t/P_t equal to their mean values $\bar{\alpha}$ and $r_f + \bar{\alpha} - g$, we get the desired result. ■

Proof of Corollary 2. Follows immediately from Proposition 1 and the formula for the impulse-responses, $\psi_j = \Gamma(j+d)/(\Gamma(d)\Gamma(j+1))$. ■

Proof of Proposition 3. Rewrite (10) as

$$X_{2t} = \Delta^d(\sigma_t^2 - \bar{\sigma}^2) = \lambda X_{1,t-1} + \sum_{i=1}^p \phi_i X_{2,t-i} + \theta(L) \mu_t. \quad (25)$$

Inserting into the definition of the returns (15) and using the risk premium link (13) we find that $X_{1t} = r_t - \bar{r} = \gamma X_{2t} + \varepsilon_t$, which implies that

$$X_{1t} = \gamma \lambda X_{1,t-1} + \gamma \sum_{i=1}^p \phi_i X_{2,t-i} + \gamma \theta(L) \mu_t + \varepsilon_t. \quad (26)$$

Combining (25) and (26) and defining u_t as in (17) we get the system (21) with the stated restrictions for part B, and part A follows by setting $\lambda = 0$.

In part A, there are $2p$ zero restrictions and p cross-equation restrictions in $A(L)$ and there are $2q$ zero restrictions and q cross-equation restrictions in $B(L)$. However, one of those is used to identify γ since when $\omega_{\varepsilon\mu} \neq 0$, γ cannot be determined from the elements of Ω_u in (17) but

is determined in either $A(L)$ or $B(L)$ provided $p + q \geq 1$, which gives the $\max(p + q - 1, 0)$ cross-equation restrictions. If, in addition, we assume $\omega_{\varepsilon\mu} = 0$ we find that

$$\Omega_u = \begin{bmatrix} \omega_\varepsilon^2 + \gamma^2\omega_\mu^2 & \gamma\omega_\mu^2 \\ \gamma\omega_\mu^2 & \omega_\mu^2 \end{bmatrix}$$

and γ is determined in Ω_u as $\gamma = \text{cov}(u_{1t}, u_{2t})/\text{var}(u_{2t})$. Hence, in this case there is one additional cross-equation restriction, provided $p + q \geq 1$.

In part B, there are $2(1(p = 0) + \max\{0, p - 1\})$ zero restrictions and $p + 1$ cross-equation restrictions in $A(L)$ and there are $2q$ zero restrictions and q cross-equation restrictions in $B(L)$. Again, one of those is used to identify γ since when $\omega_{\varepsilon\mu} \neq 0$, γ cannot be determined from the elements of Ω_u in (17) but is determined in either $A(L)$ or $B(L)$, which gives the $p + q$ cross-equation restrictions. If, in addition, we assume $\omega_{\varepsilon\mu} = 0$, γ is determined in Ω_u as before, and there is one additional cross-equation restriction. ■

Proof of Proposition 4. As in the proof of Proposition 3, the system follows by inserting X_{2t} from (25) into the definition of the returns (15) but now using the risk premium link (14). In part A, there are $3p$ and $3q$ zero restrictions in $A(L)$ and $B(L)$, respectively. In part B, there are $3\max\{1, p\} - 1$ and $3q$ zero restrictions in $A(L)$ and $B(L)$, respectively. As before, γ cannot be determined from the elements of Ω_u when $\omega_{\varepsilon\mu} \neq 0$, but if $\omega_{\varepsilon\mu} = 0$ then γ is identified as in the proof of Proposition 3. ■

References

- Andersen, T., Bollerslev, T., Diebold, F. X. & Vega, C. (2004), ‘Real-time price discovery in stock, bond and foreign exchange markets’, *Working Paper, University of Pennsylvania*.
- Andersen, T. G. & Bollerslev, T. (1998), ‘Answering the skeptics: Yes, standard volatility models do provide accurate forecasts’, *International Economic Review* **39**, 885–905.
- Andersen, T. G., Bollerslev, T. & Diebold, F. X. (2005), ‘Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility’, *Working Paper, University of Pennsylvania*.
- Andersen, T. G., Bollerslev, T., Diebold, F. X. & Ebens, H. (2001), ‘The distribution of realized stock return volatility’, *Journal of Financial Economics* **61**, 43–76.

- Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2003), ‘Modelling and forecasting realized volatility’, *Econometrica* **71**, 579–625.
- Ang, A., Hodrick, R. J., Xing, Y. & Zhang, X. (2005), ‘The cross-section of volatility and expected returns’, *Forthcoming in Journal of Finance* .
- Backus, D. K. & Gregory, A. W. (1993), ‘Theoretical relations between risk premiums and conditional variances’, *Journal of Business and Economic Statistics* **11**, 177–185.
- Baillie, R. T. (1996), ‘Long memory processes and fractional integration in econometrics’, *Journal of Econometrics* **73**, 5–59.
- Baillie, R. T., Bollerslev, T. & Mikkelsen, H. O. (1996), ‘Fractionally integrated generalized autoregressive conditional heteroskedasticity’, *Journal of Econometrics* **74**, 3–30.
- Bandi, F. M. & Russell, J. R. (2005*a*), ‘Microstructure noise, realized variance, and optimal sampling’, *Working paper, University of Chicago Graduate School of Business* .
- Bandi, F. M. & Russell, J. R. (2005*b*), ‘Separating market microstructure noise from volatility’, *Forthcoming in Journal of Financial Economics* .
- Bandi, F. M. & Russell, J. R. (2005*c*), Volatility, *in* J. R. Birge & V. Linetsky, eds, ‘Handbook of Financial Engineering (Forthcoming)’, North-Holland, Amsterdam.
- Barndorff-Nielsen, O. E. & Shephard, N. (2002), ‘Econometric analysis of realized volatility and its use in estimating stochastic volatility models’, *Journal of the Royal Statistical Society Series B* **64**, 253–280.
- Barndorff-Nielsen, O. E. & Shephard, N. (2005), ‘Variation, jumps, market frictions, and high frequency data in financial econometrics’, *World Congress of the Econometric Society, Invited Lecture* .
- Barucci, E. & Reno, R. (2002), ‘On measuring volatility and the GARCH forecasting performance’, *Journal of International Financial Markets, Institutions and Money* **12**, 183–200.
- Beran, J. (1994), *Statistics for Long-Memory Processes*, Chapman-Hall, New York.

- Black, F. (1976), ‘Studies of stock market volatility changes’, *Proceedings of the American Statistical Association, Business and Economic Statistics Section* pp. 177–181.
- Bollerslev, T., Engle, R. F. & Wooldridge, J. M. (1988), ‘A capital asset pricing model with time-varying covariances’, *Journal of Political Economy* **96**, 116–131.
- Brandt, M. W. & Kang, Q. (2004), ‘On the relationship between the conditional mean and volatility of stock returns: A latent VAR approach’, *Journal of Financial Economics* **72**, 217–257.
- Breidt, F. J., Crato, N. & de Lima, P. (1998), ‘The detection and estimation of long-memory in stochastic volatility’, *Journal of Econometrics* **83**, 325–348.
- Britten-Jones, M. & Neuberger, A. (2000), ‘Option prices, implied price processes, and stochastic volatility’, *Journal of Finance* **55**, 839–866.
- Campbell, J. Y. & Hentschel, L. (1992), ‘No news is good news: An asymmetric model of changing volatility in stock returns’, *Journal of Financial Economics* **31**, 281–318.
- Chou, R., Engle, R. F. & Kane, A. (1992), ‘Measuring risk aversion from excess returns on a stock index’, *Journal of Econometrics* **52**, 201–224.
- Chou, R. Y. (1988), ‘Volatility persistence and stock valuations: Some empirical evidence using GARCH’, *Journal of Applied Econometrics* **3**, 279–294.
- Christensen, B. J. & Nielsen, M. Ø. (2005), ‘Asymptotic normality of narrow-band least squares in the stationary fractional cointegration model and volatility forecasting’, *Forthcoming in Journal of Econometrics* .
- Comte, F. & Renault, E. (1998), ‘Long-memory in continuous-time stochastic volatility models’, *Mathematical Finance* **8**, 291–323.
- Crato, N. & de Lima, P. (1994), ‘Long-range dependence in the conditional variance of stock returns’, *Economics Letters* **45**, 281–285.
- Dacorogna, M. M., Müller, U. A., Nagler, R. J., Olsen, R. B. & Pictet, O. V. (1993), ‘A geographical model for the daily and weekly seasonal volatility in the foreign exchange market’, *Journal of International Money and Finance* **12**, 413–438.

- Diebold, F. X. & Inoue, A. (2001), ‘Long memory and regime switching’, *Journal of Econometrics* **105**, 131–159.
- Ding, Z. & Granger, C. W. J. (1996), ‘Modelling volatility persistence of speculative returns: A new approach’, *Journal of Econometrics* **73**, 185–215.
- Doornik, J. A. (2001), *Ox: An Object-Oriented Matrix Language*, 4th edn, Timberlake Consultants Press, London.
- Doornik, J. A. & Ooms, M. (2001), ‘A package for estimating, forecasting and simulating arfima models: Arfima package 1.01 for Ox’, *Working Paper, Nuffield College, Oxford* .
- Engle, R. F., Lilien, D. M. & Robins, R. P. (1987), ‘Estimating time-varying risk premia in the term structure: The ARCH-M model’, *Econometrica* **55**, 391–407.
- Engle, R. F. & Ng, V. (1993), ‘Measuring and testing the impact of news in volatility’, *Journal of Finance* **43**, 1749–1778.
- French, K. R., Schwert, G. W. & Stambaugh, R. F. (1987), ‘Expected stock returns and volatility’, *Journal of Financial Economics* **19**, 3–30.
- Geweke, J. & Porter-Hudak, S. (1983), ‘The estimation and application of long memory time series models’, *Journal of Time Series Analysis* **4**, 221–238.
- Ghysels, E., Santa-Clara, P. & Valkanov, R. (2005), ‘There is a risk-return tradeoff after all’, *Journal of Financial Economics* **76**, 509–548.
- Granger, C. W. J. & Ding, Z. (1996), ‘Varieties of long memory models’, *Journal of Econometrics* **73**, 61–77.
- Granger, C. W. J. & Joyeux, R. (1980), ‘An introduction to long memory time series models and fractional differencing’, *Journal of Time Series Analysis* **1**, 15–29.
- Hansen, P. R. & Lunde, A. (2004), ‘Realized variance and market microstructure noise’, *Forthcoming in Journal of Business and Economic Statistics* .
- Harrison, P. & Zhang, H. H. (1999), ‘An investigation of the risk and return relation at long horizons’, *Review of Economics and Statistics* **81**, 399–408.

- Hosking, J. R. M. (1981), ‘Fractional differencing’, *Biometrika* **68**, 165–176.
- Lettau, M. & Ludvigson, S. C. (2004), Measuring and modeling variation in the risk-return tradeoff, *in* Y. Ait-Sahalia & L. P. Hansen, eds, ‘Handbook of Financial Econometrics (Forthcoming)’, North-Holland, Amsterdam.
- Lobato, I. N. & Velasco, C. (2000), ‘Long memory in stock-market trading volume’, *Journal of Business and Economic Statistics* **18**, 410–427.
- Lütkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.
- Merton, R. C. (1973), ‘An intertemporal capital asset pricing model’, *Econometrica* **41**, 867–887.
- Merton, R. C. (1980), ‘On estimating the expected return on the market’, *Journal of Financial Economics* **8**, 323–361.
- Müller, U. A., Dacorogna, M. M., Olsen, R. B., Pictet, O. V., Schwarz, M. & Morgeneegg, C. (1990), ‘Statistical study of foreign exchange rates, empirical evidence of a price scaling law, and intraday analysis’, *Journal of Banking and Finance* **14**, 1189–1208.
- Nielsen, M. Ø. & Frederiksen, P. H. (2004), ‘Finite sample accuracy of integrated volatility estimators’, *Working paper, Cornell University* .
- Oomen, R. C. A. (2005a), ‘Properties of bias-corrected realized variance under alternative sampling schemes’, *Journal of Financial Econometrics* **3**, 555–577.
- Oomen, R. C. A. (2005b), ‘Properties of realized variance under alternative sampling schemes’, *Working Paper, Warwick Business School* .
- Poterba, J. M. & Summers, L. H. (1986), ‘The persistence of volatility and stock market fluctuations’, *American Economic Review* **76**, 1142–1151.
- Robinson, P. M. (1991), ‘Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regressions’, *Journal of Econometrics* **47**, 67–84.
- Robinson, P. M. (1994), Time series with strong dependence, *in* C. A. Sims, ed., ‘Advances in Econometrics’, Cambridge University Press, Cambridge, pp. 47–95.

- Robinson, P. M. (2001), ‘The memory of stochastic volatility models’, *Journal of Econometrics* **101**, 195–218.
- Robinson, P. M. (2003), Long-memory time series, *in* P. M. Robinson, ed., ‘Time Series With Long Memory’, Oxford University Press, Oxford, pp. 4–32.
- Schwert, G. W. (1989), ‘Why does stock market volatility change over time?’, *Journal of Finance* **44**, 1115–1153.
- Scruggs, J. T. (1998), ‘Resolving the puzzling intertemporal relation between the market risk premium and conditional variance: A two-factor approach’, *Journal of Finance* **53**, 575–603.
- Sowell, F. B. (1992), ‘Maximum likelihood estimation of stationary univariate fractionally integrated time series models’, *Journal of Econometrics* **53**, 165–188.
- Yu, J. (2005), ‘On leverage in a stochastic volatility model’, *Journal of Econometrics* **127**, 165–178.
- Zhang, L., Mykland, P. A. & Ait-Sahalia, Y. (2003), ‘A tale of two time scales: Determining integrated volatility with noisy high-frequency data’, *Forthcoming in Journal of the American Statistical Association* .

Table 1: Elasticity Estimates

Memory Parameter (d)	$\frac{d \log P_t}{d \log \sigma_t^2}$
0.0	-0.006
0.1	-0.007982
0.2	-0.011053
0.3	-0.015307
0.4	-0.021198
0.49	-0.028415

Table 2: Summary Statistics

Data Series:	r_t^{nom}	r_{ft}^{nom}	g_t^{nom}	π_t	variance $\times 100$
Panel A: Realized variance 1988:1-2002:12					
Mean	0.9006%	0.4229%	0.3108%	0.2183%	0.2260
Std. dev.	4.3389%	0.1521%	2.4481%	0.5904%	0.2457
Panel B: VIX implied variance 1990:1-2005:2					
Mean	0.8507%	0.3493%	0.3363%	0.2057%	0.2521
Std. dev.	4.2619%	0.1559%	2.4535%	0.6466%	0.1665

Table 3: Univariate Analysis

Model	d	ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2	θ_3	log-lik	AIC	SIC	Q(12)
Panel A: S&P 500 realized variance 1988:1-2002:12											
(0, d ,0)	0.3978 (0.0517)	—	—	—	—	—	—	868.70	-1733.40	-1730.21	18.17
(0, d ,1)	0.3757 (0.0810)	—	—	—	0.0446 (0.1252)	—	—	868.76	-1731.53	-1725.14	18.33
(0, d ,2)	0.4163 (0.0817)	—	—	—	0.0314 (0.1138)	-0.1087 (0.0998)	—	869.34	-1730.69	-1721.11	18.37*
(0, d ,3)	0.3035 (0.0936)	—	—	—	0.1376 (0.1138)	0.0092 (0.0951)	0.2046 (0.0736)	872.43	-1734.87	-1722.10	13.76
(1, d ,0)	0.3803 (0.0783)	0.0347 (0.1127)	—	—	—	—	—	868.75	-1731.50	-1725.11	18.28
(1, d ,1)	0.4116 (0.1156)	0.7447 (0.4688)	—	—	-0.7594 (0.4501)	—	—	868.71	-1729.42	-1719.84	18.14*
(1, d ,2)	-0.4863 (0.1673)	0.9934 (0.0129)	—	—	-0.0598 (0.1797)	-0.1517 (0.1247)	—	870.11	-1730.22	-1717.45	18.50*
(2, d ,0)	0.4323 (0.0773)	-0.0209 (0.1115)	-0.0856 (0.0923)	—	—	—	—	869.14	-1730.28	-1720.70	18.25*
(2, d ,1)	0.3842 (0.1296)	0.4688 (1.361)	-0.0178 (0.0790)	—	-0.4381 (1.3130)	—	—	868.75	-1727.50	-1714.73	18.30*
(3, d ,0)	0.3454 (0.1271)	0.0759 (0.1489)	-0.0482 (0.0939)	0.2034 (0.0850)	—	—	—	872.03	-1734.07	-1721.30	15.02
Panel B: VIX implied variance 1990:1-2005:2											
(0, d ,0)	0.4877 (0.0167)	—	—	—	—	—	—	982.92	-1961.83	-1958.64	35.64**
(0, d ,1)	0.4263 (0.0593)	—	—	—	0.3256 (0.0749)	—	—	991.61	-1977.22	-1970.83	16.53
(0, d ,2)	0.3198 (0.0782)	—	—	—	0.4808 (0.1020)	0.2194 (0.0986)	—	993.81	-1979.62	-1970.04	15.73
(0, d ,3)	0.3469 (0.1111)	—	—	—	0.4444 (0.1457)	0.1849 (0.1385)	-0.0370 (0.1055)	993.87	-1977.75	-1964.98	15.56*
(1, d ,0)	0.3146 (0.1172)	0.4509 (0.1379)	—	—	—	—	—	992.48	-1978.96	-1972.57	18.06
(1, d ,1)	0.3253 (0.1130)	0.3474 (0.2039)	—	—	0.1172 (0.1416)	—	—	992.83	-1977.65	-1968.07	16.30
(1, d ,2)	0.3312 (0.0888)	-0.0790 (0.3196)	—	—	0.5441 (0.2720)	0.2416 (0.1269)	—	993.84	-1977.68	-1964.91	15.68*
(2, d ,0)	0.3421 (0.1130)	0.4559 (0.1372)	-0.0760 (0.0763)	—	—	—	—	992.97	-1977.95	-1968.37	15.60
(2, d ,1)	0.3857 (0.1188)	0.7713 (0.3766)	-0.2303 (0.1563)	—	-0.3579 (0.4184)	—	—	993.23	-1976.46	-1963.69	14.79
(3, d ,0)	0.4171 (0.0899)	0.3712 (0.1179)	-0.0434 (0.0799)	-0.1119 (0.0811)	—	—	—	993.86	-1977.71	-1964.94	14.02

Note: Log-lik is the value of the maximized log-likelihood function, AIC and BIC are the Akaike and Schwarz information criteria, and Q(12) is the Ljung-Box statistic for 12 lags which is asymptotically χ^2 distributed with $12 - 1 - p - q$ degrees of freedom. For the Ljung-Box test, one and two asterisks denote rejection at 5% and 1% significance level, respectively. Asymptotic standard errors are in parentheses.

Table 4: Parameter Estimates

Volatility Series	$\bar{\alpha}$	r_f	g
Estimates from Poterba & Summers (1986)	0.6%	0.035%	0.087%
Realized variance 1988:1-2002:12	0.4777% (4.1867%)	0.2046% (0.5770%)	0.0926% (2.5169%)
VIX implied variance 1990:1-2005:2	0.5011% (4.1060%)	0.1436% (0.6524%)	0.1306% (2.5389%)

Note: Standard errors in parentheses.

Table 5: Elasticities Based on Estimated Parameter Values

Volatility Series	Specification	Elasticity (11)	Elasticity using (13)	Elasticity using (14)
Panel A: Using parameter estimates from Poterba & Summers (1986)				
Realized variance 1988:1-2002:12	(0,d,3)	-0.020477	-0.007623	-0.006
VIX implied variance 1990:1-2005:2	(1,d,0)	-0.028339	-0.010175	-0.006
Panel B: Using parameter estimates from our datasets				
Realized variance 1988:1-2002:12	(0,d,3)	-0.030403	-0.006391	-0.004777
VIX implied variance 1990:1-2005:2	(1,d,0)	-0.047471	-0.009031	-0.005011

Table 6: Bivariate Estimates Using Realized Volatility 1969:1-2003:5

Equation	A_1		B_1		B_2		B_3		$Cov(u_t) \times 1,000$	
Panel A: Unrestricted estimates										
$X_{1t} = r_t - \bar{r}$	-0.0153 (0.9755)	6.6099 (14.1792)	-0.0000 (0.9779)	-4.5319 (13.8687)	0.0054 (0.1761)	-2.5300 (2.4581)	0.1060 (0.1160)	1.5193 (3.6222)	1.7580	-0.0293
$X_{2t} = \Delta^d(\sigma_t^2 - \bar{\sigma}^2)$	0.0098 (0.0622)	-0.1097 (0.6861)	-0.0233 (0.0623)	0.0935 (0.6814)	-0.0057 (0.0108)	-0.0272 (0.2134)	0.0026 (0.0074)	0.2423 (0.2028)	-0.0293	0.0035
Panel B: Restricted (Prop. 3) estimates										
$X_{1t} = r_t - \bar{r}$	0	0	0	2.5306 (1.2982)	0	-1.1701 (0.8574)	0	2.2715 (1.5017)	1.7424	-0.0287
$X_{2t} = \Delta^d(\sigma_t^2 - \bar{\sigma}^2)$	0	0	0	0.1194 (0.0445)	0	-0.0552 (0.0508)	0	0.1072 (0.0314)	-0.0287	0.0037
Panel C: Restricted (Prop. 3 with leverage) estimates										
$X_{1t} = r_t - \bar{r}$	-0.0619 (0.0742)	0	0	0.0646 (0.2521)	0	0.0136 (0.3255)	0	0.8744 (0.9198)	1.7767	-0.0283
$X_{2t} = \Delta^d(\sigma_t^2 - \bar{\sigma}^2)$	-0.0133 (0.0031)	0	0	0.0139 (0.0469)	0	0.0029 (0.0708)	0	0.1878 (0.0612)	-0.0283	0.0034
Panel D: Restricted (Prop. 4) estimates										
$X_{1t} = r_t - \bar{r}$	0	0	0	0	0	0	0	0	1.7710	-0.0283
$X_{2t} = \Delta^d(\sigma_t^2 - \bar{\sigma}^2)$	0	0	0	0.1539 (0.0348)	0	-0.0277 (0.0788)	0	0.1833 (0.0464)	-0.0283	0.0036
Panel E: Restricted (Prop. 4 with leverage) estimates										
$X_{1t} = r_t - \bar{r}$	0	0	0	0	0	0	0	0	1.7760	-0.0282
$X_{2t} = \Delta^d(\sigma_t^2 - \bar{\sigma}^2)$	-0.0149 (0.0027)	0	0	-0.0022 (0.0512)	0	0.0175 (0.0869)	0	0.2052 (0.0418)	-0.0282	0.0034

Note: Asymptotic standard errors in parentheses.

Table 7: Bivariate Estimates Using Implied Volatility 1986:4-2003:4

Equation	A_1		$Cov(u_t) \times 1,000$	
Panel A: Unrestricted estimates				
$X_{1t} = r_t - \bar{r}$	-0.0156 (0.0722)	5.2332 (1.9321)	1.6595	-4.27×10^{-3}
$X_{2t} = \Delta^d(\sigma_t^2 - \bar{\sigma}^2)$	-0.0131 (0.0011)	0.4653 (0.0210)	-4.27×10^{-3}	7.95×10^{-4}
Panel B: Restricted (Prop. 3) estimates				
$X_{1t} = r_t - \bar{r}$	0	5.2376 (1.9023)	1.6507	-3.90×10^{-3}
$X_{2t} = \Delta^d(\sigma_t^2 - \bar{\sigma}^2)$	0	0.4691 (0.0256)	-3.90×10^{-3}	1.08×10^{-3}
Panel C: Restricted (Prop. 3 with leverage) estimates				
$X_{1t} = r_t - \bar{r}$	-0.0819 (0.0426)	3.0607 (1.5761)	1.6687	-4.33×10^{-3}
$X_{2t} = \Delta^d(\sigma_t^2 - \bar{\sigma}^2)$	-0.0128 (0.0011)	0.4774 (0.0211)	-4.33×10^{-3}	7.93×10^{-4}
Panel D: Restricted (Prop. 4) estimates				
$X_{1t} = r_t - \bar{r}$	0	0	1.6792	-3.98×10^{-3}
$X_{2t} = \Delta^d(\sigma_t^2 - \bar{\sigma}^2)$	0	0.4815 (0.0263)	-3.98×10^{-3}	1.08×10^{-3}
Panel E: Restricted (Prop. 4 with leverage) estimates				
$X_{1t} = r_t - \bar{r}$	0	0	1.6792	-4.35×10^{-3}
$X_{2t} = \Delta^d(\sigma_t^2 - \bar{\sigma}^2)$	-0.0132 (0.0011)	0.4788 (0.0217)	-4.35×10^{-3}	7.95×10^{-4}

Note: Asymptotic standard errors in parentheses.

Table 8: Bivariate Analysis

Model	log-lik	LR	d.f.	p -value	ω_ε	ω_μ	$\omega_{\varepsilon\mu}/(\omega_\varepsilon\omega_\mu)$	$\hat{\gamma}$	$\hat{\lambda}$
Panel A: Realized variance 1988:1-2002:12									
Unrestricted	1213.68	—	—	—	—	—	—	—	—
Restricted (Prop. 3)	1201.77	23.82	12	0.0215	0.0677	0.0019	-0.8195	21.1876 (16.9451)	—
Restricted (Prop. 3 w. leverage)	1208.81	9.74	11	0.5539	0.0457	0.0018	-0.5222	4.6559 (6.0366)	-0.0133 (0.0031)
Restricted (Prop. 4)	1199.87	27.63	13	0.0102	—	—	—	—	—
Restricted (Prop. 4 w. leverage)	1208.37	10.64	12	0.5600	—	—	—	—	-0.0149 (0.0027)
Panel B: VIX implied variance 1990:1-2005:2									
Unrestricted	1347.42	—	—	—	—	—	—	—	—
Restricted (Prop. 3)	1318.36	58.13	2	0.0000	0.0426	0.0010	-0.3466	11.1657 (4.1309)	—
Restricted (Prop. 3 w. leverage)	1346.65	1.54	1	0.2141	0.0413	0.0009	-0.2393	6.4113 (3.2975)	-0.0128 (0.0011)
Restricted (Prop. 4)	1316.30	62.25	3	0.0000	—	—	—	—	—
Restricted (Prop. 4 w. leverage)	1345.34	4.17	2	0.1245	—	—	—	—	-0.0132 (0.0011)

Note: Asymptotic standard errors in parentheses.

Figure 1: Impulse-response functions for the preferred models for realized and implied variance.

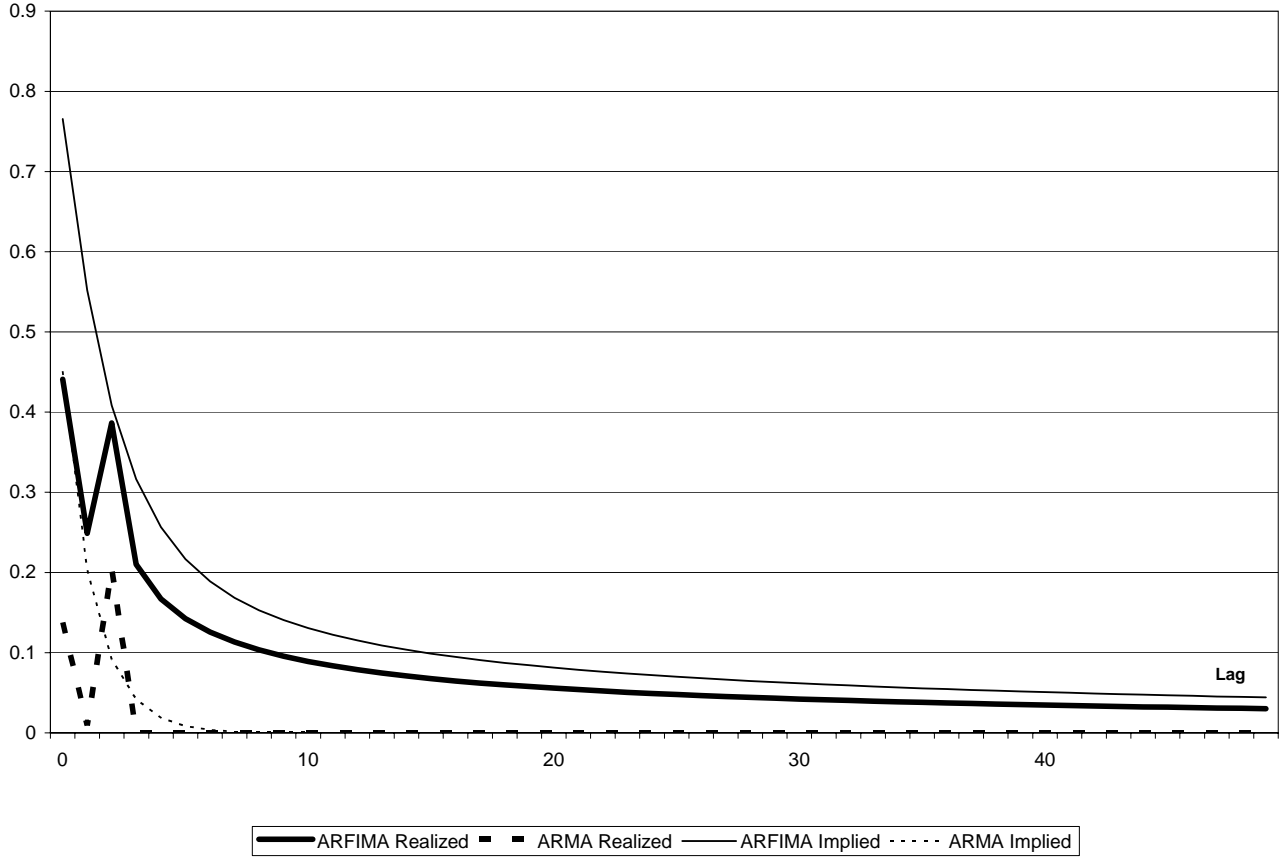
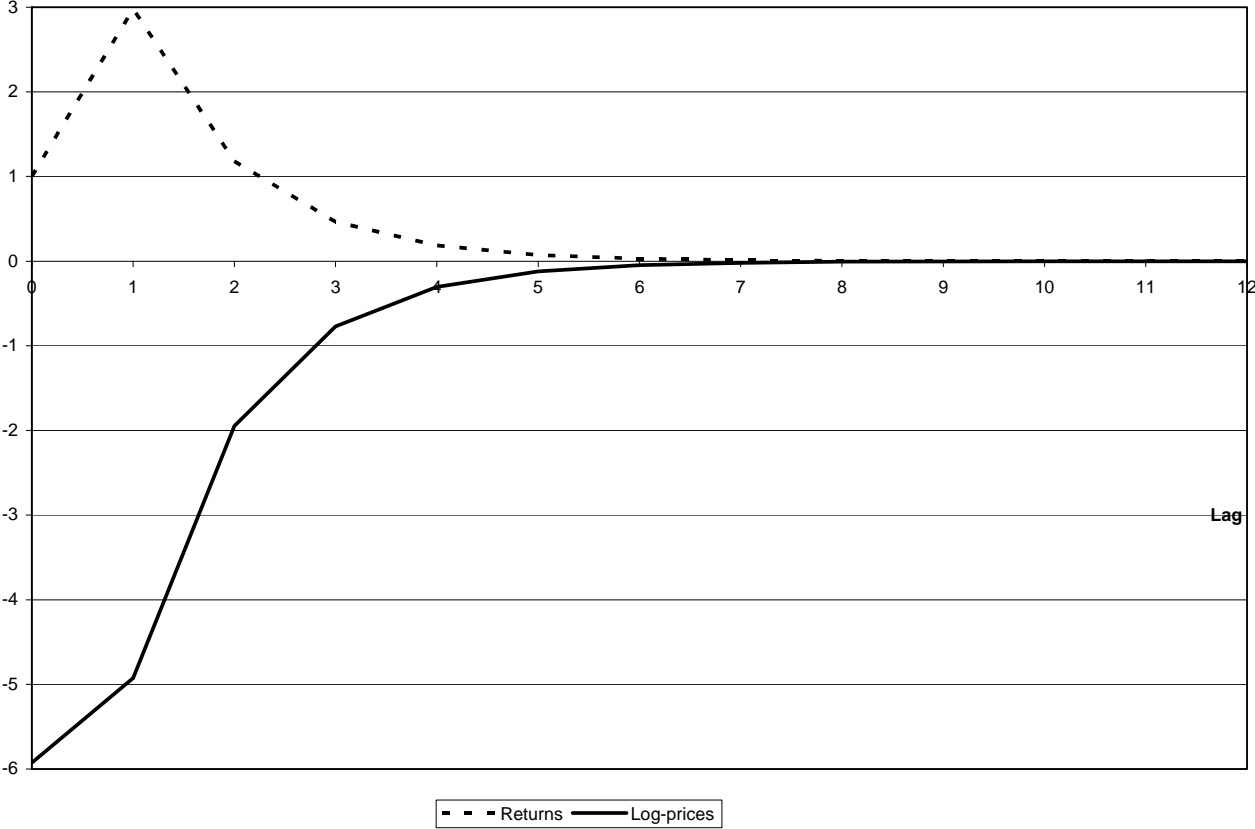


Figure 2: Impulse-response functions of returns and log-prices to unit shock in volatility innovations.



Working Papers published by the Department of Management

- 2005-1: Esben Koling Lastrup and Johannes Raaballe: Udbytteannonceringseffekten i Danmark.
- 2005-2: Carina Sponholtz: Separating the Stock Market's Reaction to Simultaneous Dividend and Earnings Announcements.
- 2005-3: Kyle Bagwell and Per Baltzer Overgaard: Look How Little I'm Advertising!
- 2005-4: Jerker Nilsson and Ole Øhlenschläger Madsen: Cross-border Mergers Between Agricultural Co-operatives - a Governance Perspective.
- 2005-5: Bent Jesper Christensen and Morten Ørregaard Nielsen: The Effect of Long Memory in Volatility on Stock Market Fluctuations.