MONASH University

Australia

Department of Econometrics and Business Statistics

<http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/>

**Common non-linearities in multiple series of stock
market volatility**

Heather M. Anderson and Farshid Vahid

January 2013

Working Paper 01/13

Common non-linearities in multiple series of stock market volatility*

Heather M. Anderson and Farshid Vahid

Department of Econometrics and Business Statistics

Monash University

Australia [†]

January 2013

Abstract

Decreases in stock market returns often lead to higher increases in volatility than increases in returns of the same magnitude, and it is common to incorporate these so-called leverage effects in GARCH and stochastic volatility models. Recent research has also found it useful to account for leverage in models of realized volatility, as well as in models of the continuous and jump components of realized volatility. This paper explores the use of smooth transition autoregressive (STAR) models for capturing leverage effects in multiple series of the continuous components of realized volatility. We find that the leverage effect is well captured by a common nonlinear factor driven by returns, even though the persistence in each country's volatility is country specific. A three country model that incorporates both country specific persistence and a common leverage effect offers slight forecast improvements for mid-range horizons, relative to other models that do not allow for the common nonlinearity.

Keywords: Realized Volatility, Bipower Variation, Common Factors, Forecasting, Leverage, Smooth Transition Models.

JEL classification: C13, C32, C52, C53, G17, G32

*We thank Bertrand Candelon, Christophe Hurlin, Timo Teräsvirta and the participants of the Nonlinear Time Series Econometrics Conference in Denmark 2012, for helpful comments. We also acknowledge financial assistance from the Australian Research Council Discovery Grants DP0449995 and DP0665710.

[†]E-mail: heather.anderson@monash.edu, farshid.vahid@monash.edu

1 Introduction

The modelling of volatility in asset markets has always been an important pursuit for financial practitioners, who often require volatility forecasts for purposes such as asset pricing and risk management. The literature on this topic is voluminous, with important developments including the family of autoregressive conditional heteroskedastic (ARCH) models first put forward by Engle (1982), stochastic volatility models as surveyed by Ghysels et al. (1996) and the use of realized volatility and its components in simple time series models as popularized in a series of papers by Andersen, Bollerslev, Diebold and various coauthors (see, for instance, Andersen et al., 2001, 2003, 2007). Much of this literature has focussed on univariate modelling and forecasting, although multivariate models are potentially more important given that financial phenomena and decisions typically involve many assets.

Multivariate modelling of volatility is more challenging than univariate modelling, although there is a sizable literature on this topic as well. Bauwens et al. (2006) provide a useful survey. Prominent in this literature is the use of factor models of volatility, firstly because theoretical asset-pricing models are often based on factor representations, and secondly because factor models embody parsimony, and this facilitates modelling and often leads to improved forecast performance. Important aspects of volatility models that matter for forecasting include strong persistence in volatility and an asymmetry that is often called "the leverage effect". Each of these aspects have been incorporated into GARCH factor models, with common persistence featuring in an early paper by Bollerslev and Engle (1993) and asymmetries featuring in Kroner and Ng (1998). Early factor models in the stochastic and realized volatility literatures include work by Harvey et al. (1994) and Anderson and Vahid (2007) respectively, and these models both incorporate persistence in the factors.

There are many models of the leverage effect in the univariate GARCH and stochastic volatility literatures, and the literature on realized volatility has also started to develop models of leverage. Examples include work by Bollerslev et al. (2006), Bollerslev et al. (2009), McAleer and Medeiros (2008), Scharth and Medeiros (2009) and Martens et al. (2009). Although all of this latter work has been done in a univariate setting, McAleer and Medeiros (2008) model the volatility of sixteen Dow Jones Industrial Average (DJIA) stocks and find that very

similar specifications (and similar parameter estimates) provide a good fit to the data. Scharth and Medeiros (2009) extend this work and find that closely related nonlinear specifications for each of these sixteen stocks lead to improved forecasts of realized volatility. The success of these nonlinear models together with the closeness of their specifications suggest the possibility of a common nonlinear factor in these DJIA series of realized volatility. Given that the presence of common factors offers potential modelling and forecasting gains, we believe that an empirical exploration for such factors in multiple series of realized volatility is justified.

We study the use of a smooth transition autoregressive (STAR) factor model for capturing leverage effects in three European series of realized volatility, choosing the STAR specification because it can capture many types of nonlinear behavior, as well as strong persistence in data. A further consideration is that the strategy developed by Teräsvirta (1994) for specifying, estimating and evaluating these models has proven to be very useful, and it is relatively simple to adapt and apply this modelling strategy to our multivariate context. Our factor representation is closely related to that containing "common nonlinear components" as developed by Anderson and Vahid (1998), although the idiosyncratic contributions to each country's volatility series in the current context depend on country specific rather than common explanatory variables, and we have to modify our testing and model building strategies to incorporate this difference.

We focus on whether the data supports a STAR factor model, and on whether such a model can outperform other univariate and multivariate models with respect to capturing the important characteristics of the data and forecasting multiple periods ahead. We find that although each of our three volatility series exhibit strong idiosyncratic persistence that is linear and does not lend itself to a factor representation, the three series also exhibit nonlinear dependence on past returns and this is well represented by a common nonlinear factor.

The rest of this paper is organised as follows. Section 2 provides a preliminary description of our data, and then we build some preliminary single equation linear models of volatility in Section 3. These equations capture the persistence in volatility, and we include returns in some of these equations to further account for the possible effects of past price shocks on volatility. We then discuss smooth transition specifications in Section 4, test for STAR behaviour in each series and also test for a common STAR factor. We use the results of our tests to specify the

nonlinear factor model of volatility. We turn to the construction of forecasts in Section 5, and since the volatility equations incorporate returns and we need to be able to incorporate forecasts of returns in our volatility forecasts, we firstly build a model of returns. We then construct and present our forecast results, which show that the STAR factor model can slightly outperform linear models of volatility at five-step and ten-step ahead forecast horizons. Section 6 then presents some brief conclusions.

2 Data

Our empirical analysis is based on data drawn from version 0.2 of the library constructed by Heber et al. (2009) at the Oxford-Man Institute of Quantitative Finance. This library contains financial data for many stock market indices, and we focus on the FTSE100, the DAX and the CAC40, as representative indices for the three largest European economies, i.e. the United Kingdom, Germany and France. The data base contains a wide variety of daily volatility series calculated from high frequency observations, and we choose to work with bi-power variation series, since bi-power variation provides a consistent measure of the integrated volatility associated with standard theoretical continuous time diffusion models of the logarithm of an asset's price, and it is free from the effects of price jumps that might unduly influence the analysis. We work with bi-power variation series based on five-minute returns that have been subsampled and averaged over equidistant subgrids that are one minute apart. The subsampling delivers very efficient estimates of daily volatility that are largely free from the effects of microstructure noise. We scale the raw data by 10^4 , and then take logarithms to obtain series that we call v_{ut} , v_{gt} and v_{ft} that measure return volatilities in the United Kingdom, Germany and France respectively. Figure 1 plots these three series and illustrates clear co-movement that is also obvious from the associated correlation matrix.

Our daily returns series (denoted by r_{ut} , r_{gt} and r_{ft}) are also drawn from the Heber et al. (2009) data base. These are calculated as the differences between the logarithms of the closing and opening prices for each day and then converted into percentage measures. These series are also strongly cross-correlated, as shown in Figure 2, but this cross-correlation is weaker than that for the volatility series.

[Insert Figures 1 and 2 about here]

Our data set spans from January 4 2000 to March 9 2012, and includes five days of trade a week over most of this period. We remove a few days for which at least one of the three markets was closed, which leaves a sample of 3027 observations. We used 2728 of these observations (from January 4 2000 to December 31 2010) for modelling purposes and retained the remaining 299 observations for out of sample forecast analysis. Thus our model building sample includes the period of global financial crisis, and although stock markets had partially recovered by the end of 2010, the sample retained for out-of-sample forecast analysis nevertheless relates to a "time of financial uncertainty" in Europe. Table 1 provides some summary statistics for the initial sample. Although the log volatilities have quite different means, their standard deviations, skewness and kurtosis are very similar. The augmented Dickey and Fuller (1979) tests in Table 1 indicate that none of the three volatility series are $I(1)$, while the estimated memory coefficients (d) indicate that these series have very similar and strongly persistent autocorrelation properties. The latter were estimated via a Geweke and Porter-Hudak (1983) regression. Table 2 shows that although average returns are negative over the initial sample and they are serially correlated for several rather than one or two days, their other properties seem to be quite typical of daily returns. We find strong evidence of ARCH in each series, and upon estimating AR(5)-TGARCH(1,1) versions of the threshold switching GARCH model proposed by Glosten et al. (1993), we find that the coefficients that measure the effects of negative return shocks on conditional volatility are strongly statistically significant in each case. This last test provides preliminary evidence of leverage effects in our data set.

[Insert Tables 1 and 2 about here]

3 Linear models of volatility

We start the modelling process by developing (single equation) linear models of volatility that can capture the persistence in each series. This is in line with the strategy outlined in Teräsvirta (1994), that advocates the use of tests of linearity against nonlinear alternatives to justify the estimation of nonlinear models, but firstly requires that the dynamics of these linear models to be well specified for this testing purpose. We first work with univariate specifications of volatility, and then consider the inclusion of lagged volatilities from the other countries and lags

of the three returns series.

ARMA models provide a standard way of modelling univariate time series, and we experiment with various forms of these models for our volatility series. Our experiments include pure autoregressive structures and versions of the heterogeneous autoregressive (HAR) model proposed by Corsi (2009). The latter model typically works with a lag structure that incorporates past week and past month moving averages of (unlogged) volatility V_t as predictors and has a standard specification of

$$\ln(V_{t+1}) = \beta_0 + \beta_D \ln(V_{t-1,t}) + \beta_W \ln(V_{t-5,t}) + \beta_M \ln(V_{t-22,t}) + \varepsilon_{t+1}, \quad (1)$$

with $V_{t,t+h} = h^{-1}[V_{t+1} + V_{t+2} + \dots + V_{t+h}]$ for $h = 5$ and $h = 22$. Here, we use a linear version of (1) given by

$$\ln(V_{t+1}) = \beta_0 + \beta_D \overline{\ln(V_{t-1,t})} + \beta_W \overline{\ln(V_{t-5,t})} + \beta_M \overline{\ln(V_{t-22,t})} + \varepsilon_{t+1}, \quad (2)$$

with $\overline{\ln(V_{t,t+h})} = h^{-1}[\ln V_{t+1} + \ln V_{t+2} + \dots + \ln V_{t+h}]$ for $h = 5$ and $h = 22$. Equation (2) is simply a restricted version of a linear AR(22) in $\ln V_t$, and it fits each of our three $\ln V_t$ series (i.e. v_{ut} , v_{gt} and v_{ft}) slightly better than (1) and leads to less serial correlation in the residuals as well.

We present information criteria (BIC) and residual diagnostics (p-values for Ljung-Box Q(15) serial correlation tests) for various specifications in Tables 3a and 3b. These show that although the parsimonious ARMA and HAR specifications fit the data quite well, they perform poorly with respect to capturing the dynamics of $\ln(V_{t+1})$. This leads us to base our subsequent univariate analysis on relatively long AR specifications, and we find that sparse versions of AR(10) specifications can model these medium run dynamics quite well. The inclusion of past lags of volatility from other countries is generally unhelpful, although the inclusion of one lag of either v_{ut} or $\bar{v}_t = (v_{ut} + v_{gt} + v_{ft})/3$ leads to marginal improvements in the fit of the v_{gt} and v_{ft} equations. The inclusion of four lags of returns improves the fit of all three volatility equations, and in each case the use of an average return $\bar{r}_t = (r_{ut} + r_{gt} + r_{ft})/3$ improves fit, relative to the use of (four lags of) own country returns (r_{ut} , r_{gt} , or r_{ft}) in the modelling of volatility.

[Insert Tables 3a and 3b about here]

The three linear specifications that we choose on the basis of this modelling

exercise are given by

$$v_{it} = \varphi_{i0} + \Psi_i(L)v_{it} + \phi_i v_{ut-1} + \Phi_i(L)\bar{r}_t + \varepsilon_{it} \quad (3)$$

where $i = u, g, f$ for UK, Germany and France, respectively. We restrict $\phi_u = 0$ to avoid multicollinearity and report estimates of the remaining coefficients in Table 4. Interesting characteristics of the estimated versions of these models are the very strong similarities in own lag coefficients for all three series (excepting the first lag), as well as strong similarities in the lagged return coefficients. A (HAC corrected) test that the system can be treated as a panel (with country specific intercepts) is soundly rejected, but for the lagged return coefficients, a test that the four Φ_i coefficients for the lags on $\bar{r}_t$ in the UK equations are the same as the corresponding Φ_i coefficients for the German equation, and the same as the corresponding Φ_i coefficients for the French equation (eight cross equation restrictions) gives rise to a (HAC corrected) p-value of 0.5697. Thus all three series evolve in a very similar way, although only some of these dynamics (i.e. those related to past returns) can be attributed to a common source ($\bar{r}_t$). Since returns do not in themselves exhibit strong persistence, it appears that the persistence in each volatility series can be mainly attributed to its own lags. The equations fit the data quite well and the serial correlation diagnostics provided at the bottom of Table 4 indicate that they provide suitable linear baselines for conducting nonlinearity tests. Not surprisingly, each equation exhibits evidence of non-normality and heteroskedasticity, but the linearity tests can be corrected for heteroskedasticity and the absence of normality does not present problems when the sample size is large.

[Insert Table 4 about here]

4 Smooth transition models of volatility

The models in the previous section account for persistence in volatility and provide initial evidence that the dynamic behaviour of the three volatility series is quite similar. However, these linear specifications do not yet incorporate any mechanisms for modelling asymmetries in volatility, so we now expand our modelling framework to examine whether this sort of nonlinearity is present in the data. Like McAleer and Medeiros (2008) and Scharth and Medeiros (2009) we use a

smooth transition autoregressive (STAR) model as a starting point for our analysis, but we depart from their set-up by adapting a multivariate version of the STAR model put forward by Anderson and Vahid (1998) that explicitly allows for common nonlinear factors.

4.1 Theoretical considerations

A generic version of a multivariate STAR model is given by

$$y_t = \pi_{A0} + \pi_{A1}(L)y_t + \pi_{A2}(L)x_t + F(z_t) \cdot [\pi_{B0} + \pi_{B1}(L)y_t + \pi_{B2}(L)x_t] + \varepsilon_t \quad (4)$$

where y_t is the $n \times 1$ vector of time series under analysis, π_{A0} and π_{B0} are $n \times 1$ vector of constants, $\pi_{A1}(L)$ and $\pi_{B1}(L)$ are p -th order matrix polynomials in the lag operator with $\pi_{A1}(0) = \pi_{B1}(0) = 0$, and ε_t is a $n \times 1$ i.i.d. $(0, \Sigma)$ sequence. The variable x_t is a $k \times 1$ vector of k weakly exogenous explanatory variables for y_t and $\pi_{A2}(L)$ and $\pi_{B2}(L)$ are q -th order matrix polynomials in the lag operator with $\pi_{A2}(0) = \pi_{B2}(0) = 0$. The function $F(z_t)$ is a $(n \times n)$ diagonal transition matrix with typical diagonal entry $F_i(z_{it})$, and z_{it} is one of the np lagged regressors in $ylags_t = (y'_{t-1}, y'_{t-2}, \dots, y'_{t-p})'$ or one of the kq lagged regressors in $xlags_t = (x'_{t-1}, x'_{t-2}, \dots, x'_{t-q})'$. The specification that each transition function $F_i(z_{it})$ is a logistic function given by $F_i(z_{it}) = (1 + \exp[-\gamma_i(z_{it} - c_i)])^{-1}$, (where $\gamma_i > 0$), leads to a multivariate LSTAR model when there are no weakly exogenous variables (i.e. $xlags_t = \emptyset$), and a multivariate LSTR model when there are variables other than $ylags_t$ in the model.¹

Luukkonen et al. (1988) and Teräsvirta (1994) study univariate LSTAR models in the case where $z_{it} = y_{it-d}$, and show that the hypothesis of linearity of the specification against an LSTAR alternative can be tested by means of a usual test of

$$H_0 : \beta_{2j} = 0 \quad (j = 1, \dots, p)$$

in the following artificial regression that is obtained from a first order expansion of the transition function around $z_{it} = c_i$

$$y_{it} = \beta_0 + \beta'_1(ylags_{it}) + \beta'_2(ylags_{it} \cdot z_{it}) + \zeta_{1it} \quad (5)$$

¹Although we could consider other specifications of the $F_i(z_{it})$, we use the logistic function in this setting because it is simple, but allows for asymmetries in y_{it} that depend on the values taken by z_{it} .

where $ylags_{it} = (y_{it-1}, y_{it-2}, \dots, y_{it-p})'$. This test of p restrictions collectively tests the restriction that $\gamma_i = 0$, and rejection of the null supports a nonlinear alternative in which changes in z_{it} drive regime changes in y_{it} . The assumption that the v_{it} are normally distributed gives rise to this test as a standard Lagrange multiplier test, but the test statistic has an asymptotic χ_p^2 distribution if one simply interprets it as a moment test. However, it is clear that this test will not have any power against the alternative when the only parameter that interacts with $F(z_{it})$ is the constant term. For this reason, Luukkonen et al. (1988) consider a third order approximation to $F(z_{it})$ and suggest a third order linearity test of

$$H_0 : \beta_{2j} = \beta_{3j} = \beta_{4j} = 0 \quad (j = 1, \dots, p)$$

in the artificial regression

$$y_{it} = \beta_0 + \beta'_1 (ylags_{it}) + \beta'_2 (ylags_{it} \cdot z_{it}) + \beta'_3 (ylags_{it} \cdot z_{it}^2) + \beta'_4 (ylags_{it} \cdot z_{it}^3) + \zeta_{3it} \quad (6)$$

This test of $3p$ restrictions collectively tests the restriction that $\gamma_i = 0$, and has good power against LSTAR alternatives including those against which the first order test lacks power. However, this additional power comes at the expense of estimating three times as many parameters as in the first order test, which can be unattractive, in particular in our multivariate extension. To remedy this, Luukkonen et al. (1988) suggest an “augmented first order” test, which is a test of

$$H_0 : \beta_{2j} = \beta_4^* = 0 \quad (j = 1, \dots, p)$$

in the artificial regression

$$y_{it} = \beta_0 + \beta'_1 (ylags_{it}) + \beta'_2 (ylags_{it} \cdot z_{it}) + \beta_4^* z_{it}^3 + \zeta_{1Ait} \quad (7)$$

The addition of z_{it}^3 term to the list of regressors of the first order test gives the test power against the LSTAR alternative where the only term in the “second regime” (i.e. the part that multiplies $F(z_{it})$) is a constant. The associated test statistic for the augmented first order test follows a χ_{p+1}^2 distribution asymptotically when the null hypothesis is true. Simulations in Luukkonen et al. (1988) show that the augmented first order tests can often offer a good compromise between power and degrees of freedom concerns.

The adaptation of these tests to the multivariate LSTR model (4) discussed above involves replacing $ylags_{it}$ in (6) and (7) with $alllags_t$ where $alllags_t$ now contains p lags of *each* of the n endogenous variables and q lags of *each* of the k weakly

exogenous variables (i.e. $alllags_t = (y'_{t-1}, y'_{t-2}, \dots, y'_{t-p}, x'_{t-1}, x'_{t-2}, \dots, x'_{t-q})'$). We choose z_{it} so that $z_{it} \in \{alllags_t\}$. When the null hypothesis is true (and the model is linear), the first order, the third order and the augmented first order test statistics have $\chi^2_{(np+kq)}$, $\chi^2_{3(np+kq)}$ and $\chi^2_{np+kq+1}$ distributions asymptotically. Each of these tests can be conducted separately on each of the equations in the system.

The possibility that some of the y_i variables might have a common LSTR nonlinearity arises if, when we set $z_{it} = z_t$ in the single equation LSTR tests on y_i , the tests reject the null of linearity for more than one i . The evidence then suggests that these y_i have LSTR nonlinearities with the same transition variable, leading to the possibility that the same nonlinearity is common to at least two of the n variables. If the LSTR nonlinearity is common, then there will be at least one α making $\alpha' y_t$ linear in mean, in which case

$$E(\alpha' y_t^\dagger \otimes w_t^\dagger) = 0, \quad (8)$$

(where in this context w_t is the set of test regressors $\{(alllags_t \cdot z_t), (alllags_t \cdot z_t^2)$ and $(alllags_t \cdot z_t^3)\}$ for the third order test or $\{(alllags_t \cdot z_t)$ and $(z_t^3)\}$ for the first order augmented test). The dagger superscripts on y_t and w_t indicate that the linear effects of the intercept and $alllags_t$ have been removed. From Lemma 2.2 in Anderson and Vahid (1998), the test that there are s such linear combinations of the y_i which are linear in mean (and therefore $n - s$ common nonlinearities if all n of the y_i are nonlinear) is given by the test statistic

$$J = T \sum_{i=1}^{i=s} \lambda_i^2$$

where the $\lambda_1^2 < \lambda_2^2 < \dots < \lambda_n^2$ are the estimated squared canonical correlations between $y_t^\dagger$ and $w_t^\dagger$. The test statistic will have a $\chi^2_{(3np+3kq-n)s+s^2}$ distribution for the third order version of the common nonlinearities test, and a $\chi^2_{(np+kq+1-n)s+s^2}$ distribution for the augmented first order version of the test. Rejection of the null hypothesis will provide evidence that there are more than $n - s$ nonlinearities in the system. The special case of this test in which $s = n$ is a “system linearity test” of the null hypothesis that the system contains no LSTR nonlinearities.

The above theory assumes that p , x_t , q and z_t are known, but typically economic theory is used to choose x_t and then p and q are chosen using a model selection criterion such as AIC or BIC. The choice of z_t is sometimes guided

by context or economic considerations, but when this is not the case, Teräsvirta (1994) suggests running a series of linearity tests that use the chosen lag structure, and then consecutively set z_t equal to one of the choice variables in $alllags_t$ and conduct a (single equation) linearity test that uses z_t as a transition variable. One can then choose the z_t that minimizes the p-value of the set of linearity tests.

4.2 Common nonlinearities in volatility

We take the three volatility variables as our series of interest (i.e. $y_t = (v_{ut}, v_{gt}, v_{ft})'$) and use the results of the preliminary modelling exercise in Section 3 to include $\bar{r}_t$ (rather than r_{ut}, r_{gt} and r_{ft}) in the weakly exogenous regressor set x_t . We maintain the previously chosen lag structures for these variables (lags 1 to 5 and lag 10 for each variable in y_t and lags 1 to 4 for $\bar{r}_t$), and then consider which lags of y_t and x_t seem to be suitable candidates for z_t . The lags of $\bar{r}_t$ are a-priori candidates because leverage effects in volatility are typically associated with movements in returns. The lags of v_{ut}, v_{gt} and v_{ft} are also candidates on the basis that the effects of price shocks on volatility can depend on the current volatility state, as discussed in Fornari and Mele (1997) and Anderson et al. (1999). However, since v_{ut-1} is the only variable from this latter set that appears in all three preliminary models and we plan to test for common nonlinearities, we exclude all other lags of v_{ut}, v_{gt} and v_{ft} from further consideration as a transition variable. Thus we consider $z_t \in \{v_{ut-1}, \bar{r}_{t-1}, \bar{r}_{t-2}, \bar{r}_{t-3}, \bar{r}_{t-4}\}$ as potential transition variables in the work that follows.

Implementation in our setting is based on a slightly broader linear model than that given in Table 4, in that our version of (4) includes lags of all *three* volatility variables in the separate country equations, rather than just the own country lags. This is because our common nonlinearity tests require that we take the same linear effects out of each equation prior to testing for common nonlinearity, so we need the same set of explanators in each equation. Satisfying this requirement amounts to including some statistically insignificant variables in the linear portion of (4) while conducting our linearity tests. This has potential to slightly reduce the power of our single equation tests of linearity, but it is unlikely to affect the test outcomes.

The top and third panel of Table 5 contain p-values that respectively relate to the third order and augmented first-order single equation linearity tests. The

third order tests provide very strong evidence of nonlinearity in the volatility of each country, regardless of the proposed transition variable. Some of the p-values are extremely small, and although this is not clear from the reported results, we note that the transition variable that leads to the smallest p-value for each country is $\bar{r}_{t-1}$. The p-values from the augmented first order tests in the third panel of Table 5 are in line with those in top panel, although they are now all somewhat higher. We still see statistically significant evidence of nonlinearity, and also see that the transition variable that led to the smallest p-value for each country was $\bar{r}_{t-1}$. This is consistent with a leverage effect story, in that the corresponding STR model for each country predicts that current volatility behaviour depends on past shocks to prices (and hence to returns).

[Insert Table 5 about here]

The evidence from the above tests supports the hypothesis that each volatility series has LSTR nonlinearities with the same transition variable ($\bar{r}_{t-1}$), so we now explore the possibility that the same nonlinearity (with transition variable $\bar{r}_{t-1}$) is common to at least two of the three volatilities. The second and fourth panels of Table 5 contain the relevant test statistics. In each case, the tests fail to reject the null hypothesis that there is an α making $\alpha'y_t$ linear in mean, which is consistent with the system having two separate, rather than three LSTR type nonlinearities. The p-values for these tests are 0.6837 and 0.8711. Further, the p-values of 0.3778 in the second panel and 0.8076 in the fourth panel indicate that the data also supports the null hypothesis that there are two independent α that make $\alpha'y_t$ linear in mean, consistent with the system having one nonlinear component, rather than two. The final test in each of the second and fourth panels provides strong evidence against the null hypothesis that none of the equations have LSTR type nonlinearities. We therefore conclude that the system of three volatility equations contains a single nonlinear component that is common to each volatility. This component is consistent with an LSTR nonlinearity with a transition variable $\bar{r}_{t-1}$ that drives smooth transitional changes between volatility regimes associated with high and low returns.

We estimate the implied model by minimizing the (log of the) determinant of the 3×3 residual matrix (i.e. maximum likelihood if the errors are normally distributed) using a general to specific methodology. Our initial model included four lags of the composite return $\bar{r}_t$ and lags of all three volatility variables in all

three linear portions of the system as well as in the common nonlinear portion, and then we removed statistically insignificant variables one by one, until all variables in the remaining model were statistically significant. The final estimated model is reported in Table 6. The accompanying diagnostics indicate that this specification fits the data better than the linear model according to BIC which accounts for the number of parameters. Further, it has captured most of the nonlinearity in the data.

Table 6 shows quite strong similarities across the three equations with respect to their estimated country specific coefficients, but overall the specification is not consistent with a Gonzalez et al. (2005) type PSTR (panel smooth transition regression) specification. Interesting features of this model are that the lag structure of our basic linear specification in Table 4 essentially reappears as the linear component in our nonlinear common factor model, and the nonlinear component enters each volatility equation with approximately the same coefficients (the factor loadings are estimated to be 1.106 for UK, 1.213 for France and normalized to 1.000 for Germany). The nonlinear part of the model does not contain any lagged volatility terms at all, so it can be attributed entirely to movements in returns.

[Insert Table 6 about here]

Empirical estimates of the transition function are presented in Figure 3. The top figure shows how the transition function varies over time and it is visually somewhat similar to the volatility illustrations provided in Figure 1. This suggests that the broad swings that we see in volatility are related nonlinearly to returns. The bottom figure shows how the transition function varies with returns, and we see that this operates in a range where most of the returns are, although only a relatively small proportion of returns (about 12%) fall into the region where $\bar{r}_{t-1} > 1.290$ and the transition function takes a value that is greater than one half.

[Insert Figure 3 about here]

The interesting interpretative aspect of the model is that negative returns increase return volatility in all three countries for several days. This is evident from the linear component of the model. Small positive returns decrease volatility, but for larger positive returns most of the decrease implied by the linear component

of the model is cancelled out by the nonlinear terms, so that large positive returns have very little effect on volatility at all. Figure 4 plots the contribution to each country’s volatility made by the previous day’s (average) return. This is mostly in line with usual leverage effect arguments, but it is also in line with the reversion effects discussed in Fornari and Mele (1997). A final interesting interpretive aspect of our model is that apart from the presence of v_{ut-1} in the French volatility equation, the volatility in each country is driven by own country lagged volatility. Given the recent literature that documents interdependence and contagion across stock markets (see, for instance, Forbes and Rigobon, 2002), this finding is initially quite surprising. However, as discussed in Aït-Sahalia et al. (2010), one can attribute most of the interdependence in volatility to jumps. The significant effect of *average* returns on each of the three volatility series can be interpreted as evidence of the effects of common price jumps on volatility.

5 Forecasts

We use the period from January 4 2011 until March 9 2012 (299 observations) for 1-step ahead, 5-step ahead and 10 step ahead forecast analysis, and compare the forecast performance of our non-linear common component model with the linear models considered in Section 3. In all forecasting exercises that follow, we keep the specifications developed in Sections 3 and 4 fixed, but we re-estimate the model parameters over expanding windows as we expand the sample forward through the forecasting period. We evaluate our forecasts in terms of mean square forecast errors (MSFE), and conduct (HAC corrected) tests due to Diebold and Mariano (1995) to assess whether forecast performances from different models are similar. We also calculate model confidence sets (MCS) due to Hansen et al. (2011) in order to classify the models into two groups that contain the best forecasting model(s) and poorer forecasting models².

The calculation of one step ahead forecasts are straightforward for all model specifications, but for the nonlinear model we need to model the average returns in order to obtain more than one-step-ahead volatility forecasts. Using the average returns, as opposed to three individual returns, simplifies the task, but we do not

²Our MCSs are pseudo-MCSs, given that we are working with expanding window forecasts. Hansen et al (2011) comment that differences between MCS and pseudo-MCS results are small in the empirical applications that they studied.

have data for the volatility of average returns (this of course could be constructed from the raw high frequency returns data, but we do not have access to that). Hence, we build a GARCH model for the average returns, and allow this GARCH specification to be non-linear. In particular we develop an ANST-GARCH model as in Anderson et al. (1999). The estimated ANST-GARCH model for average returns is

$$\begin{aligned}
\bar{r}_t &= \frac{-0.026}{(0.015)} - \frac{0.047}{(0.021)}\bar{r}_{t-1} - \frac{0.035}{(0.019)}\bar{r}_{t-5} + \varepsilon_t \\
\varepsilon_t &= \sigma_t \eta_t \quad \eta_t \sim \text{Student's t with df} = \frac{17.023}{(5.153)} \\
\sigma_t^2 &= \frac{0.023}{(0.005)} + \frac{0.083}{(0.014)}\varepsilon_{t-1}^2 + \frac{1.023}{(0.033)}\sigma_{t-1}^2 \\
&\quad + \frac{1}{1 + \exp(-3.862 \varepsilon_{t-1})} \times \left(\frac{-0.023}{(0.005)} - \frac{0.060}{(0.020)}\varepsilon_{t-1}^2 - \frac{0.188}{(0.062)}\sigma_{t-1}^2 \right).
\end{aligned} \tag{9}$$

where the numbers in brackets under the coefficient estimates are HAC corrected standard errors. We do not supply the standard error for the γ coefficient in the transition function because the non-normal distribution of $\hat{\gamma}$ renders the interpretation of a t-statistic for this estimate inappropriate. The appearance of the $1.023\sigma_{t-1}^2$ in the first part of equation may raise the question of stationarity of this specification and existence of a conditional variance. However, as Anderson et al. (1999) show, the condition for stationarity of ANST-GARCH is $\alpha_1 + \beta_1 + 0.5(\alpha_2 + \beta_2) < 1$, which is satisfied for the above estimated model. The presence of the negative coefficients in the second part of the ANST-GARCH equation is not an issue for the same reason. The news impact curve implied by the ANST-GARCH model is illustrated in Figure 5. It has a similar shape to the response curves in Figure 4, providing reassurance that this model of the volatility of average returns is consistent with our joint model of the volatilities associated with the separate index returns.

[Insert Figure 5 about here]

Since lags of $\bar{r}_t$ appear in the common non-linear component of the log-volatility equations, the forecast of this component is not simply the forecast of $\bar{r}_t$ plugged in the non-linear function. To forecast the non-linear component

$$f(\bar{r}_{t-1}, \bar{r}_{t-2}) = \frac{1}{1 + \exp\{-3.361(\bar{r}_{t-1} - 1.290)\}} \times \left(\frac{0.110}{(0.019)}\bar{r}_{t-1} + \frac{0.050}{(0.016)}\bar{r}_{t-2} \right)$$

1 to 10 steps ahead, we simulate 10,000 sets of 10 random draws from the t -distribution, use $\hat{\sigma}_T^2$ and $\hat{\varepsilon}_T$ from the estimated sample to initiate the conditional variance equation to generate ε_{T+1} to ε_{T+10} recursively, and then use these to generate 10,000 sample paths for $\{\bar{r}_{T+1}, \dots, \bar{r}_{T+10}\}$ and hence 10,000 sets of $\{f(\bar{r}_T, \bar{r}_{T-1}), \dots, f(\bar{r}_{T+9}, \bar{r}_{T+8})\}$. We use the average of these latter sets to calculate our 1-step to 10-step ahead forecasts of the common non-linear component. We add these to the forecast of the linear parts to form the forecast of log-volatilities implied by our non-linear model. For the returns that appear in the linear part of the model, we only need to plug in the forecast of the returns from the mean equation in the returns model.

Table 7 provides some forecast comparisons for one step ahead, five step ahead and ten step ahead forecast horizons. We report the mean squared forecast error (MSFE) for all models and discuss these first, before making some more formal comparisons. Initial inspection of each MSFE column suggests that the single forecasting equations all have very similar forecasting ability, while the two systems (that also include lagged returns) seem to improve forecasts of volatility, with there being little difference in forecasting ability between the linear and nonlinear systems. This pattern is apparent for all forecasting horizons, so that the models seem to fall into two groups - single equation models that seem to lead to inferior forecasts, and then the two systems with returns, that seem to lead to similar, but better forecasts.

[Insert Table 7 about here]

For one-step ahead forecasts, the linear system delivers slightly better forecasts for UK and German volatility, while the nonlinear system delivers slightly better forecasts for French volatility. Diebold and Mariano (1995) tests applied to single country one step ahead volatility forecasts do not provide any evidence that the forecast performance of either system is better than the other. However, the trace of the MSFE matrix for the nonlinear system (0.6190) is lower than that for the linear system (0.6200), suggesting that the overall forecast performance of the nonlinear system is slightly better when forecasting just one period ahead. The performance of the nonlinear system relative to the linear system seems to improve once we consider five-step ahead forecasts, and now the common non-linear component model delivers better volatility forecasts for all three countries. Nevertheless the differences between forecasts delivered by the two systems are

still quite small and Diebold and Mariano (1995) tests are unable to detect any significant differences. More pronounced differences become apparent once the forecast horizon is increased to 10 steps ahead, and the nonlinear model continues to lead to smaller (although not significantly smaller) MSFE for each country.³

Tests of the null hypothesis that the forecasts from one model are no better than those from another are complicated by the fact that we have eight models to consider for each country and horizon, and it is hard to control the size associated with more than one pairwise test. We deal with this problem by presenting p-values associated with model confidence sets (MCS) defined by Hansen et al. (2011), and as explained by these authors these p-values allow statements about statistical significance that are valid in the usual sense. Lower p-values indicate poorer forecasting ability. Given a set of models and a loss function, the $MCS_{1-\alpha}$ consists of the subset of those models that are "best" with a confidence level of $1 - \alpha$, and model i belongs in the $MCS_{1-\alpha}$ if and only if its MCS p-value $\hat{p}_i \geq \alpha$. An MCS will always contain at least one model, but it can contain many or all models if the data is not informative.

MCS p-values were calculated using the MULCOM package (Hansen and Lunde, 2010) and are displayed on the right of the reported MSFE for each model. These p-values show that there is not enough discriminating information in the data to identify superior forecasting models with very high levels of confidence. For example, the 95% confidence set includes all models for all countries and all horizons. However, there is sufficient evidence to classify models at the 80% level, and our following discussion therefore focusses on finding sets of models that include the "best" model with a probability of 80%. Models with p-values greater or equal to 0.2 fall into the $MCS_{0.8}$. We see that for the UK volatility models, the $MCS_{0.8}$ contains only the two systems, regardless of the forecast horizon. The same is true for the German volatility models used for five-step and ten step ahead forecasts, but for the one step ahead German volatility forecasts the $MCS_{0.8}$ consists of all models. For France, we see that the $MCS_{0.8}$ for five step ahead forecasting consists of just the nonlinear system, whereas it also contains the linear system and the VAR(10) equation for one step ahead forecasts, and all but the sparse VAR(10)

³Patton and Sheppard (2009) show that the Diebold-Mariano test has higher power when the loss function is QLIKE rather than MSFE. We also compared the forecasts from the linear and nonlinear systems using the QLIKE loss function, but did not find statistically significant differences.

equation for ten step ahead forecasting. Thus, although we have some ability to discriminate between the UK models, and to a lesser extent between the German and French models, we cannot conclude that the common nonlinear factor model is best. We can, however, observe that the two systems are often contained in the set of "best models", which points to improvements in forecasting when models incorporate effects driven by lagged returns.

6 Conclusion

The empirical literature on stock market volatility emphasises persistence and leverage effects, since these characteristics are found in almost all individual stock volatility series, as well as in the volatility associated with stock market indices. This paper studies the joint modelling of these characteristics across three European stock indices, using a carefully constructed lag structure to capture persistence and a smooth transition specification to address the leverage effects. Our analysis finds that although the persistence in each volatility series is essentially country specific, the leverage effect is well described by a common nonlinear smooth transition factor. The factor is driven by the lagged average return across countries, and it implies that negative index returns raise volatility by more than positive returns of the same size, and that large positive returns are not associated with increases in volatility.

References

- Aït-Sahalia, Y., J. Cacho-Diaz, and R. J. A. Laeven (2010). Modeling financial contagion using mutually exciting jump processes. Technical report, N.B.E.R.
- Andersen, T., T. Bollerslev, and F. X. Diebold (2007). Roughing it up: including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics* 89 (4), 701 – 720.
- Andersen, T., T. Bollerslev, F. X. Diebold, and H. Ebens (2001). The distribution of realized stock volatility. *Journal of Financial Economics* 61, 43 – 76.
- Andersen, T., T. Bollerslev, and P. Diebold, F. X. and Labys (2003). Modeling and forecasting realized volatility. *Econometrica* 71, 579 – 625.
- Anderson, H. M., K. Nam, and F. Vahid (1999). Asymmetric nonlinear smooth-transition GARCH models. In P. Rothman (Ed.), *Dynamic Modeling and Econometrics in Economics and Finance*. Kluwer.
- Anderson, H. M. and F. Vahid (1998). Testing multiple equation systems for common nonlinear components. *Journal of Econometrics* 84, 1–36.
- Anderson, H. M. and F. Vahid (2007). Forecasting the volatility of Australian stock returns: Do common factors help? *Journal of Business and Economic Statistics* 25, 76–90.
- Bauwens, L., S. Laurent, and J. V. K. Rombouts (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics* 21, 79 – 109.
- Bollerslev, T. and R. F. Engle (1993). Common persistence in conditional variances. *Econometrica* 61, 167 – 186.
- Bollerslev, T., U. Kretschmer, C. Pigorsch, and G. Tauchen (2009). A discrete-time model for daily s&p500 returns and realized variations: jumps and leverage effects. *Journal of Econometrics* 150, 151 – 166.
- Bollerslev, T., J. Litvinova, and G. Tauchen (2006). Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics* 4, 353 – 384.

- Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* 7, 174–196.
- Dickey, D. and W. A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427 – 431.
- Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13, 253 – 263.
- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of variance of UK inflation. *Econometrica* 50, 987–1008.
- Forbes, K. J. and R. Rigobon (2002). No contagion, only interdependence. Measuring stock market comovements. *Journal of Finance* 57, 2223 – 2261.
- Fornari, F. and A. Mele (1997). Sign and volatility switching ARCH models: theory and applications to international stock markets. *Journal of Applied Econometrics* 12, 49–66.
- Geweke, J. and S. Porter-Hudak (1983). The estimation and application of long memory time series. *Journal of Time Series Analysis* 4, 221 – 238.
- Ghysels, E., A. C. Harvey, and E. Renault (1996). Stochastic volatility. In G. S. Maddala and C. R. Rao (Eds.), *Handbook of Statistics*, Volume 14. Elsevier Science.
- Glosten, L., R. Jagannathan, and R. Runkle (1993). On the relationship between the expected value and the volatility of the nominal return on stocks. *Journal of Finance* 48, 1779 – 1801.
- Gonzalez, A., T. Teräsvirta, and D. van Dijk (2005). Panel smooth transition models. Technical report, University of Technology, Sydney. Wp 165.
- Hansen, P. R. and A. Lunde (2010). *MULCOM 2.00 Econometric Toolkit for Multiple Comparisons* http://mit.econ.au.dk/vip_hm/alunde/MULCOM/MULCOM.HTM.
- Hansen, P. R., A. Lunde, and J. M. Nason (2011). Model confidence sets. *Econometrica* 79, 453 – 497.

- Harvey, A. C., E. Ruiz, and N. Shephard (1994). Multivariate stochastic volatility models: *Review of Economic Studies* 61, 247 – 264.
- Heber, G., A. Lunde, N. Shephard, and K. Sheppard (2009). Oxford-Man Institute’s realized library (version 0.2). Oxford-Man Institute, University of Oxford.
- Kroner, K. F. and V. K. Ng (1998). Modelling asymmetric comovements of asset returns. *The Review of Financial Studies* 11, 4, 817 – 844.
- Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988). Testing linearity against smooth transition autoregressive models. *Biometrika* 75, 491 – 499.
- Martens, M., D. van Dijk, and M. de Pooter (2009). Forecasting S&P500 volatility: long memory, level shifts, leverage effects, day-of-the-week seasonality, and macroeconomic announcements. *International Journal of Forecasting* 25, 282 – 303.
- McAleer, M. and M. Medeiros (2008). A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries. *Journal of Econometrics* 147, 104 – 119.
- Patton, A. J. and K. Sheppard (2009). Evaluating volatility and correlation forecasts. In T. G. Andersen, R. A. Davis, J.-P. Kreiss, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pp. 801–838. Springer.
- Scharth, M. and M. Medeiros (2009). Asymmetric effects and long memory in the volatility of Dow Jones stocks. *International Journal of Forecasting* 25, 304 – 327.
- Teräsvirta, T. (1994). Specification, estimation and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association* 89, 208 – 218.

Table 1: Descriptive statistics of (log) daily volatility

Statistic/stock	v_{ut}	v_{gt}	v_{ft}
Mean	-0.6886	-0.0116	-0.1411
Std dev	1.0400	1.0438	1.0177
Max	3.4413	4.0246	3.7734
Min	-3.1674	-3.0799	-2.8914
Skewness	0.3498	0.3307	0.2450
Kurtosis	2.9776	3.0021	2.8902
ADF _c p-value	0.0030	0.0023	0.0016
Memory (d)	0.1694	0.1673	0.1837
Days sampled	2728	2728	2728

Notes: Volatility has been scaled by 10^4 . The ADF_c test incorporates a lag length chosen by BIC. The HAC standard errors for the fractional differencing parameter (d) are approximately 0.025 in each case.

Table 2: Descriptive statistics of daily returns

Statistic/stock	r_{ut}	r_{gt}	r_{ft}
Mean	-0.0520	-0.0495	-0.0621
Std dev	1.0559	1.4417	1.3013
Max	7.0441	9.9934	7.2825
Min	-5.7603	-9.4122	-8.1239
Skewness	-0.0897	-0.0520	-0.1300
Kurtosis	6.5979	7.5311	6.9675
ADF _c p-value	0.0001	0.0001	0.0001
$\rho(1)$	-0.026	0.014	-0.052
$\rho(5)$	-0.026	-0.057	-0.056
ARCH(1) p-value	0.0000	0.0000	0.0000
TGARCH p-value	0.0000	0.0000	0.0000
Days sampled	2728	2728	2728

Notes: Returns are expressed in % terms. The ADF_c test incorporates a lag length chosen by BIC. The TGARCH p-value relates to the statistical significance of the "negative news" coefficient in a threshold AR(5)-GARCH(1,1) model.

Table 3a: Model selection (BIC for single eqns)

Model\dependent variable	v_{ut}	v_{gt}	v_{ft}
HAR	1.260	1.260	1.226
ARMA(2,1)	1.260	1.258	1.225
AR(10)	1.272	1.268	1.233
AR (10) sparse	1.262	1.260	1.224
VAR(10)	1.319	1.315	1.272
VAR(10) sparse	1.289	1.285	1.242
AR(10) sparse + $\bar{v}_{t-1}$	1.264	1.260	1.217
AR(10) sparse + v_{ut-1}	1.262	1.259	1.217
AR(10) sparse + $v_{ut-1} + r_{it-1..it-4}$	1.231	1.224	1.178
AR(10) sparse + $v_{ut-1} + \bar{r}_{t-1..t-4}$	1.224	1.219	1.172

Table 3b: Model selection (residual p-values for Q(15) stats for single eqns)

Model\dependent variable	v_{ut}	v_{gt}	v_{ft}
HAR	0.013	0.000	0.000
ARMA(2,1)	0.001	0.000	0.000
AR(10)	0.977	0.895	0.617
AR (10) sparse	0.852	0.383	0.227
VAR(10)	0.965	0.904	0.736
VAR(10) sparse	0.855	0.446	0.388
AR(10) sparse + $\bar{v}_{t-1}$	0.860	0.392	0.343
AR(10) sparse + v_{ut-1}	0.852	0.428	0.309
AR(10) sparse + $v_{ut-1} + r_{t-1..t-4}$	0.622	0.318	0.344
AR(10) sparse + $v_{ut-1} + \bar{r}_{t-1..t-4}$	0.641	0.306	0.276

Table 4: Baseline linear models

Variable	v_{ut}	v_{gt}	v_{ft}
Constant	-0.044 (0.010)	0.021 (0.015)	0.044 (0.016)
Own lag 1	0.361 (0.025)	0.357 (0.029)	0.285 (0.031)
Own lag 2	0.134 (0.020)	0.152 (0.023)	0.167 (0.020)
Own lag 3	0.109 (0.021)	0.057 (0.024)	0.061 (0.022)
Own lag 4	0.111 (0.020)	0.109 (0.023)	0.091 (0.022)
Own lag 5	0.112 (0.020)	0.101 (0.020)	0.117 (0.021)
Own lag 10	0.125 (0.018)	0.131 (0.017)	0.120 (0.018)
$\bar{r}_{t-1}$	-0.078 (0.010)	-0.080 (0.010)	-0.084 (0.011)
$\bar{r}_{t-2}$	-0.038 (0.008)	-0.036 (0.008)	-0.034 (0.008)
$\bar{r}_{t-3}$	-0.024 (0.009)	-0.028 (0.009)	-0.020 (0.008)
$\bar{r}_{t-4}$	-0.018 (0.007)	-0.023 (0.007)	-0.022 (0.007)
v_{ut-1}		0.050 (0.017)	0.112 (0.024)
$\bar{R}^2$	0.822	0.824	0.824
BIC	1.224	1.219	1.172
p-val s.c. Q(15)	0.641	0.306	0.276
p-val ARCH(10)	0.000	0.015	0.000
p-val Jacque-Bera	0.000	0.000	0.000
$ln \widehat{\Omega} $	-7.7186		
BIC	0.8973		

Note: Standard errors and p-values are HAC corrected. The diagnostic statistics include tests for (the absence of) serial correlation, ARCH and non-normality.

Table 5: P-values for linearity tests

Transition variable\dependent variable	v_{ut}	v_{gt}	v_{ft}
Third order linearity tests			
$\bar{v}_{t-1}$	0.0001	0.0001	0.0055
$\bar{r}_{t-1}$	0.0000	0.0000	0.0000
$\bar{r}_{t-2}$	0.0000	0.0000	0.0000
$\bar{r}_{t-3}$	0.0000	0.0000	0.0000
$\bar{r}_{t-4}$	0.0257	0.0008	0.0206
Third order common nonlinear factor tests			
Two nonlinear factors (vs three)	0.6837		
One nonlinear factor (vs two)	0.3778		
No nonlinearity	0.0000		
Augmented first order linearity tests			
$\bar{v}_{t-1}$	0.1766	0.0077	0.0549
$\bar{r}_{t-1}$	0.0000	0.0000	0.0000
$\bar{r}_{t-2}$	0.1090	0.0535	0.0768
$\bar{r}_{t-3}$	0.0059	0.2658	0.0026
$\bar{r}_{t-4}$	0.6921	0.1205	0.6960
First order augmented common nonlinear factor test			
Two nonlinear factors (vs three)	0.8711		
One nonlinear factor (vs two)	0.8076		
No nonlinearity	0.0000		

Note : The tests are HAC corrected. See the text for details regarding how the tests are constructed.

Table 6: Estimated nonlinear common factor model

$$\begin{pmatrix} \hat{v}_{ut} \\ \hat{v}_{gt} \\ \hat{v}_{ft} \end{pmatrix} = \begin{pmatrix} 1.106 \\ (0.169) \\ 1 \\ 1.213 \\ (0.117) \end{pmatrix} \times \frac{1}{1 + \exp \{-3.361 (\bar{r}_{t-1} - 1.290)\}} \times \begin{pmatrix} 0.110\bar{r}_{t-1} + 0.050\bar{r}_{t-2} \\ (0.019) \quad (0.016) \end{pmatrix} + \begin{pmatrix} \hat{l}_{ut} \\ \hat{l}_{gt} \\ \hat{l}_{ft} \end{pmatrix}$$

with equations for $(\hat{l}_{ut}, \hat{l}_{gt}, \hat{l}_{ft})'$ and model diagnostics as follows:

Variable	$\hat{l}_{ut}$	$\hat{l}_{gt}$	$\hat{l}_{ft}$
Constant	-0.059 (0.010)	0.023 (0.006)	
Own lag 1	0.323 (0.012)	0.308 (0.011)	0.224 (0.031)
Own lag 2	0.148 (0.013)	0.160 (0.012)	0.171 (0.020)
Own lag 3	0.101 (0.014)	0.084 (0.013)	0.082 (0.022)
Own lag 4	0.126 (0.013)	0.117 (0.013)	0.099 (0.022)
Own lag 5	0.134 (0.013)	0.109 (0.013)	0.131 (0.021)
Own lag 10	0.110 (0.011)	0.111 (0.011)	0.109 (0.018)
$\bar{r}_{t-1}$	-0.119 (0.009)	-0.118 (0.009)	-0.129 (0.009)
$\bar{r}_{t-2}$	-0.046 (0.008)	-0.045 (0.008)	-0.046 (0.008)
$\bar{r}_{t-3}$	-0.024 (0.007)	-0.025 (0.007)	-0.019 (0.007)
$\bar{r}_{t-4}$		-0.020 (0.005)	-0.020 (0.005)
v_{ut-1}			0.125 (0.009)
Model diagnostics			
P-val s.c. Q(15)	0.489	0.043	0.090
$\ln \hat{\Omega} $	-7.742		
BIC	0.882		
P-val nonlinearity due to v_{ut}	0.042		
P-val nonlinearity due to $\bar{r}_{t-1}$	0.185		
P-val nonlinearity due to $\bar{r}_{t-2}$	0.203		
P-val nonlinearity due to $\bar{r}_{t-3}$	0.235		
P-val nonlinearity due to $\bar{r}_{t-4}$	0.149		

Note: The nonlinearity tests are system tests of the null that there is no additional nonlinearity. See Appendix D in Anderson and Vahid (1998). Coefficient standard errors and p-values for test statistics are HAC corrected.

Table 7: Mean-squared error of one step ahead forecasts and associated MCS p-values

Model	UK		Germany		France	
	<i>MSFE</i>	p-val	<i>MSFE</i>	p-val	<i>MSFE</i>	p-val
One step ahead forecasts						
HAR	0.2197	0.11	0.2276	0.47	0.2078	0.17
ARMA(2,1)	0.2203	0.11	0.2273	0.47	0.2102	0.15
AR(10)	0.2181	0.14	0.2249	0.48	0.2094	0.15
AR(10) sparse	0.2195	0.11	0.2271	0.47	0.2085	0.16
VAR(10)	0.2201	0.11	0.2237	0.53	0.2069	0.21
VAR(10) sparse	0.2195	0.11	0.2258	0.47	0.2070	0.17
Linear system (with $v_{ut-1} + \bar{r}_{t-1..t-4}$)	0.2061	1.00	0.2174	1.00	0.1964	0.61
Nonlinear factor model	0.2069	0.85	0.2176	0.95	0.1945	1.00
Five step ahead forecasts						
HAR	0.3685	0.17	0.4234	0.07	0.3659	0.12
ARMA(2,1)	0.3681	0.17	0.4225	0.07	0.3719	0.12
AR(10)	0.3666	0.17	0.4221	0.07	0.3720	0.12
AR(10) sparse	0.3665	0.17	0.4198	0.07	0.3701	0.12
VAR(10)	0.3702	0.17	0.4284	0.07	0.3827	0.12
VAR(10) sparse	0.3671	0.17	0.4248	0.07	0.3812	0.12
Linear system (with $v_{ut-1} + \bar{r}_{t-1..t-4}$)	0.3460	0.36	0.3833	0.56	0.3476	0.18
Nonlinear system	0.3416	1.00	0.3807	1.00	0.3415	1.00
Ten step ahead forecasts						
HAR	0.4725	0.07	0.5470	0.08	0.4650	0.40
ARMA(2,1)	0.4718	0.07	0.5389	0.08	0.4681	0.30
AR(10)	0.4724	0.06	0.5395	0.08	0.4705	0.30
AR(10) sparse	0.4710	0.07	0.5384	0.08	0.4702	0.30
VAR(10)	0.4710	0.07	0.5469	0.07	0.4852	0.21
VAR(10) sparse	0.4648	0.08	0.5459	0.08	0.4880	0.13
Linear system (with $v_{ut-1} + \bar{r}_{t-1..t-4}$)	0.4549	0.20	0.5200	0.35	0.4642	0.30
Nonlinear system	0.4474	1.00	0.5144	1.00	0.4538	1.00

Figure 1: Log daily volatility ($\times 10^4$) of the three indices

The correlation between the three series is $\begin{pmatrix} 1 & 0.888 & 0.945 \\ 0.888 & 1 & 0.950 \\ 0.945 & 0.950 & 1 \end{pmatrix}$.

Figure 2: Daily returns for the three indices ($\times 100$)

The correlation between the three series is $\begin{pmatrix} 1 & 0.777 & 0.764 \\ 0.777 & 1 & 0.846 \\ 0.764 & 0.846 & 1 \end{pmatrix}$.

Figure 3: Plots of the empirical transition function

Figure 4: The effect of last period's return on current log volatility

Figure 5: News impact curves implied by the ANST-GARCH model

Note: We set $\sigma_{t-1}^2 = 1$ in the model in equation (9) to draw the low volatility curve and $\sigma_{t-1}^2 = 2$ to draw the high volatility curve. The graph shows predicted log volatilities.