The Return, Volatility and Interaction of Stock and Bond Markets around Macroeconomic Announcements

EVERT B. VRUGT*

This version: April 2009

ABSTRACT

Using a novel dataset with macroeconomic announcements and expectations, I estimate a variety of GARCH models to analyze the behavior of stock and bond markets around macroeconomic news releases. Macroeconomic announcements strongly affect conditional means, variances and covariances of stocks and especially bonds. Conditional variances and covariances increase on announcement days, but this is followed by drops of as much as 30% on days following announcements. This largely offsets the initial increase on announcement days and the aggregate impact of news on the covariance structure of stocks and bonds on and around announcement days is transitory, rather than permanent. The impact of macroeconomic news releases is not only statistically significant, but also economically. Estimated Value-at-Risk provisions are more precise when macroeconomic information is taken into account.

JEL classifications: C32, G12

Keywords: Macroeconomic news, Time-varying covariances, Multivariate GARCH, Value-at-Risk

* *E-mail address:* evrugt@xs4all.nl (E.B. Vrugt). I thank Ivo Arnold, Peter de Goeij, Roman Kräussl and Pieter Jelle van der Sluis for useful comments and Charles Bos and Michael Creel for making parts of the Ox code available.

1. Introduction

The covariance structure of assets plays a central role in financial economics. It is crucial in areas as diverse as asset allocation, risk management, and the pricing of derivative contracts on the values of multiple assets. Yet, the fundamental determinants of changes in the covariance structure of assets remain poorly understood.

This paper analyzes the impact of macroeconomic news releases on the conditional returns, variances, and covariance of U.S. stocks and bonds. More specifically, I try to answer three main empirical questions.

First, does the release of macroeconomic information affect returns, variances and covariances of stocks and bonds? How asset prices incorporate new information is among the most fundamental questions in finance. Unfortunately, starting with the work of Cutler, Poterba and Summers (1989) and Schwert (1989), it has been difficult to establish an empirical link between asset prices and news on fundamentals.

Second, is the impact of macroeconomic news on the covariance structure of stocks and bonds confined to announcement days, or are the days immediately before and after the actual announcement day also special? And closely related to this question, does the aggregate impact of macroeconomic announcements over all these days have a permanent or transitory effect on the covariance structure of stocks and bonds? The impact of scheduled macroeconomic news on volatility is a priori unclear. If news releases resolve uncertainty about future fundamentals among market participants, as in Pasquariello (2007), volatility is likely to decrease after announcements. If, on the other hand, macroeconomic news increases disagreement among market participants, perhaps as a result of differences in the way economic agents process information, volatility may increase. The question as to how financial markets process new information is therefore ultimately an empirical one.

Third and finally, is the release of macroeconomic news also economically important? Although news announcements may affect returns, variances and covariances statistically, it is unclear whether these effects are significant enough to matter in practice. I illustrate the relevance of macroeconomic news in the context of risk management, by estimating Value-at-Risk provisions with and without taking information about fundamentals into account.

The goal of this paper is thus to analyze whether macroeconomic information can explain variation in the covariance structure of asset returns, not only statistically but also economically.

Although the effect of macroeconomic announcements on the first and second moments of asset returns has received considerable attention in the literature, previous studies typically analyze a single asset class and often a limited number of announcements. Studies set within a univariate framework include Jones, Lamont and Lumsdaine (1998), Li and Engle (1998), Fleming and Remolona (1999), Balduzzi, Elton and Green (2001), Flannery and Protopapadakis (2002), Andersen, Bollerslev, Diebold and Vega (2003), Bomfim (2003), and Boyd, Hu and Jagannathan (2005). Macroeconomic news, however, affects both stock and bond markets simultaneously although not necessarily in the same way. Inflation surprises, for example, are thought to be negative for bonds and positive for stocks insofar as stocks provide a hedge against inflation. Announcements thus influence stocks and bonds differently, thereby changing their covariance. A multivariate perspective is therefore required to analyze these effects. I study the response of stocks and bonds jointly for a broad set of 13 types of macroeconomic announcements.

In addition, I explicitly consider the behavior of conditional variances and covariances on days surrounding macroeconomic announcements. These pre- and post-announcement days have received comparably little attention in the extant literature, but are empirically important, as this paper shows.

Finally, no study to-date has explicitly considered the practical relevance of the link between asset price dynamics and news on fundamentals. The final contribution of this paper is to check how the inclusion of macroeconomic information affects the accuracy of daily Value-at-Risk provisions for stocks, bonds and a portfolio consisting of stocks and bonds. Perhaps closest to the out-of-sample economic evaluation is the recent paper of De Goeij and Marquering (2008). However, their economic evaluation is set within the context of portfolio construction rather than risk management and their models do not include macroeconomic information.

Apart from these contributions, I employ a novel source for macroeconomic expectations data from Bloomberg. The Bloomberg survey is an important benchmark in the investment management industry. Market participants refer to the median of the Bloomberg survey as the “market consensus”. All prior studies on the link between macroeconomic news and asset returns and volatilities that I am aware of use data from MMS. But MMS has stopped their survey services and the dataset ends in 2002 (see Brenner, Pasquariello and Subrahmanyam, 2008). Although the starting date of the Bloomberg survey is later, the resulting dataset is without structural break and on a consistent basis. These considerations make the Bloomberg survey an attractive dataset for studying the link between asset price dynamics and news on fundamentals.

In establishing these new stylized facts, I employ various univariate and multivariate GARCH models with multiplicative news effects. Given the large number of parameters, multivariate GARCH models with news effects are computationally challenging to estimate with standard gradient search routines. In addition, constraints must be imposed to ensure that the covariance matrix is positive definite, which adds to the complexity. Since conditional variances and covariances are time-varying, these constraints must hold for each of the many observations in the sample period. The few prior studies that estimate multivariate GARCH models with news effects have dealt with these issues by imposing constant correlations (Christiansen, 2000) or sacrificing efficiency by estimating the model in two stages (De Goeij and Marquering, 2006, 2008; Andersen, Bollerslev, Diebold and Vega, 2007; and Brenner, Pasquariello and Subrahmanyam, 2008). I opt for a different solution by estimating the models with a more robust optimization algorithm: simulated annealing developed by Corana, Marchesi, Martini and Ridella (1987). Since very few assumptions are made on the function to be evaluated, simulated annealing is quite robust and hence very useful in the present challenging context.

The main findings of this paper can be summarized as follows.

First, I find strong evidence for the importance of macroeconomic information in modeling conditional means, variances, and covariances of stocks and bonds. Tests for exclusions of macroeconomic variables from the model are rejected stocks and bonds and for both univariate and multivariate specifications. The response of the bond market is stronger than that of the equity market. This holds for both the impact on the conditional mean return and the conditional variance. A one-standard deviation improvement in economic conditions leads, on average, to a significant eight-basis-point decrease in bond returns and an insignificant increase in equity returns of 3 basis points. The conditional variance of bond returns increases approximately 30% on announcement days. For equities, the increase is around 10% which is comparable to the increase in the stock-bond covariance.

Second, there is interesting variation in the reaction of conditional variances and covariances surrounding macroeconomic announcements. Conditional variances as well as the covariance between stocks and bonds are sharply lower on days immediately following news announcements: around 12% for stocks, 30% for bonds and 25% for the stock-bond conditional covariance. I label this phenomenon "post-macroeconomic-announcement relief". This finding is very robust: it is present for both conditional stock and bond variances as well as the stock-bond covariance and shows up in all univariate and multivariate specifications. When I separate the four most influential announcements from the rest of the announcements, post-macroeconomic-announcement relief remains significant for both selections.

Conditional variances and covariances increase on average on announcement days, but decrease significantly on days immediately following announcement days. The aggregate impact of news during the two days surrounding the announcement and the announcement days itself is insignificantly different from zero. This suggests that changes in the covariance structure of stocks and bonds on and around macroeconomic announcement days are transitory, rather than permanent.

Third, apart from statistical significance, macroeconomic information is economically important. I use the various GARCH specifications with and without macroeconomic news to form daily Value-at-Risk (VaR) provisions. The backtests show that the predicted VaR provisions are in general of good quality, since most forecasts fall in the "green zone", as defined by the Basle Committee on Banking Supervision (1996). Adding macroeconomic information enhances accuracy, especially during periods when the model performance falls in the "yellow zone". Incorporating macroeconomic information therefore helps to reduce these violations. This finding is of practical relevance to financial risk managers.

The rest of this paper is organized as follows. Section 2 describes the datasets for the financial futures and the macroeconomic expectations and realizations. Section 3 provides the methodological details and Section 4 summarizes the estimation results of the univariate and multivariate GARCH models. In Section 5, I analyze the economic significance of the findings by calculating Value-at-Risk provisions with and without macroeconomic information. Finally, Section 6 concludes.

2. Data

2.1. Financial futures data

As discussed above, I examine the impact of regularly scheduled U.S. macroeconomic announcements on stocks and bonds. I use the S&P 500 future and the 10-year Treasury bond future to measure the performance of stocks and bonds. Both contracts are traded in Chicago: the former at the Chicago Mercantile Exchange and the latter at the Chicago Board of Trade. Contracts are rolled over from the first nearest-to-expiry contract to the second if trading volume and open interest exceed that of the first (see Ederington and Lee, 1995 for a similar approach). Based on the series of prices, I create a sample of daily logarithmic price changes over the period 1996 – 2007. Table 1 shows summary statistics for the 3076 equity and 2977 bond future observations.

[Insert Table 1 here]

The mean logarithmic change in prices is slightly higher for the S&P 500 future than for 10-year Treasury bond future over the sample period. The standard deviation of the equity future is substantially higher than that of the bond future. Both series exhibit negative skewness and excess kurtosis, leading to significant departures from normality as the Jarque-Bera tests show. The unconditional correlation between stocks and bonds is slightly negative (-0.108). Finally, the Ljung-Box statistics show that there is a small, but significant degree of autocorrelation in both stock and bond returns. Squared returns, on the other hand, are highly and significantly autocorrelated. These are well-known properties of daily data which have motivated the use of GARCH models for this type of data.

2.2. Macroeconomic announcements and expectations data

The Bloomberg market survey starts in 1996 and contains a range of widely-followed macroeconomic indicators. The Bloomberg survey serves as benchmark for market participants. International brokerage firms with economics teams refer to the Bloomberg survey median as “the consensus” and contrast in-house forecasts and subsequent realizations with the consensus. Previous papers on the empirical link between macroeconomic surprises and asset prices exclusively analyze data from Money Market Services International (MMS). Examples are Jones, Lamont and Lumsdaine (1998), Li and Engle (1998), Balduzzi, Elton and Green (2001), Andersen, Bollerslev, Diebold and Vega (2003), Flannery and Protopapadakis (2003), Ehrmann and Fratzscher (2005), De Goeij and Marquering (2006), and Brenner, Pasquariello and Subrahmanyam (2008). The MMS survey has been the longest running expectations dataset available, but unfortunately, MMS has stopped with their services in 2002. It is not clear whether Action Economics, which currently runs the survey, uses the same methodology as MMS (see Brenner, Pasquariello and Subrahmanyam, 2008). Recent studies therefore either have to work with a dataset that contains a structural break, or use a sample that ends in 2002, as Brenner, Pasquariello and Subrahmanyam (2008) do. Bloomberg, on the other hand, has been running a consistent survey since inception. Although the start of the sample is later, the dataset also includes recent observations and is on a consistent basis. Table 2 provides an overview of the Bloomberg survey data.

[Insert Table 2 here]

I collect the Bloomberg median for 13 macroeconomic news releases. The indicators are from various categories: economic activity (nonfarm payrolls, retail sales less autos, industrial production, housing starts, durable goods orders, and GDP), international trade (trade balance), prices (PPI and CPI), surveys (consumer confidence, Chicago purchasing managers, and ISM manufacturing), and monetary policy (the Fed funds target rate). The dataset contains many important U.S. announcements and should paint a fairly complete picture of the impact of news on stocks, bonds and the way they interrelate around times with major macroeconomic news releases.

Most indicators are released between 7:30 and 9:00 in the morning. The Fed funds target rate announcement is an exception, as it is usually released at 13:15 Chicago time. The median number of contributing forecasters varies between 47 (Chicago purchasing managers) and 87 (the Fed funds target rate). Compared to the roughly 40 forecasters surveyed by MMS, this compares favorably, especially given that I also include non-key indicators which are probably less widely followed. The number of observations ranges from a high of 134 monthly ISM manufacturing announcements to a low of 43 for the quarterly GDP announcements.

Following the convention in the literature, I construct standardized macroeconomic surprises for each variable *i* as follows (see Balduzzi, Elton and Green (2001) for example):

$$S_{i,t} = \frac{A_{i,t} - E_{i,t} \mid \Omega_{t-1}}{\sigma_i}, \tag{1}$$

where $A_{i,t}$ is the actual, initial, unrevised announcement of variable *i* at time *t*, $E_{i,t} \mid \Omega_{t-1}$ is the Bloomberg survey median for the release of variable *i* at time *t* based on information available prior to the releases of the indicator (Ω_{t-1}), and σ_i is the sample standard deviation of the time-series with surprises for indicator *i*. Scaling by the sample standard deviation enhances comparability across indicators, but does not affect significance levels or the log likelihood of the models in the remainder of this paper. The monetary policy target rate deserves special attention. Since it moves in discrete steps of 25 basis points, the survey median is less informative than the mean for this variable and I take the latter for expected target rate changes. This is consistent with Ehrmann and Fratzscher (2005). I combine the standardized surprises of the 13 variables into a single surprise variable, S_t[1]. If there is more than one announcement on a particular day, I take the average of all contemporaneously announced surprises. In the empirical section I also separate variables into key- (nonfarm payrolls, CPI, PPI, and the Fed funds target rate) and non-key announcements (consumer confidence, Chicago purchasing managers, ISM manufacturing, the trade balance, retail sales, industrial production, housing starts, durable goods orders, and GDP) to assess the robustness of the main conclusions. In addition to surprise variable S_t, I construct a dummy variable D_t that takes a value of one on announcement days and zero on non-announcement days.

2.3. Accuracy of Bloomberg forecasts

I evaluate the accuracy of the Bloomberg survey by examining unbiasedness and efficiency of the survey median using standard regression-based tests. In the interest of brevity, the full results are not reported here, but available upon request. For the unbiasedness tests, I estimate the following equation for each macroeconomic announcement *i*:

$$A_{i,t} = c_i + b_i E_{i,t} \mid \Omega_{t-1} + u_{i,t}, \tag{2}$$

where $A_{i,t}$ is the initial unrevised announcement *i* at time *t*, and $E_{i,t} \mid \Omega_{t-1}$ is the Bloomberg survey median based on information available up to and including time *t-1*. If forecasts are on average unbiasedness, $c_i = 0$ and $b_i = 1$. I use a Wald-test for each of the 13 macroeconomic variables to test this hypothesis. At the 1%-level of significance, unbiasedness is rejected for the price series (PPI and CPI), retail sales and durable goods orders. Visual inspection of the residuals shows that accuracy deteriorates somewhat during the last years of the sample when uncertainty about economic prospects increases. Furthermore, a single outlier observation is sometimes responsible for rejection of the unbiasedness hypothesis. Overall, these results are comparable to recent tests for MMS (see Ehrmann and Fratzscher, 2005).

[1] Previous research shows that the equity market reacts negatively to increases in producer and consumer prices and the Federal funds target rate. This is confirmed in my dataset. In constructing the pooled announcement variable, I therefore reverse the signs of surprises in these variables to bring them in line with the interpretation of the other announcements.

Efficiency of the forecasts is tested by the hypothesis $H_0 : b_{i,1} = b_{i,2} = ... = b_{i,K} = 0$ in the regression:

$$A_{i,t} - E_{i,t} \mid \Omega_{t-1} = c_i + \sum_{k=1}^{K} b_{i,k} A_{i,t-k} + u_{i,t} , \qquad (3)$$

where all variables are as previously defined. I implement the efficiency test with three lags, but results are qualitatively the same for other lag lengths. At the 1%-level of significance, the efficiency hypothesis is only rejected for CPI and durable goods orders. In conclusion, the unbiasedness and efficiency tests show that the Bloomberg survey data are in general of good quality.

The accuracy of the Federal funds target rate predictions deserve special attention. Previous authors have used a multitude of data sources for the market expectations of the Federal funds target rate. Ehrmann and Fratzscher (2005) use the Reuters poll, Andersen, Bollerslev, Diebold and Vega (2003) the MMS survey and Kuttner (2001) extracts expectations from the Fed funds future. Research shows that the expectations extracted from futures prices perform better than survey data in terms of accuracy, unbiasdness and efficiency. Since I am the first to use expectations data from Bloomberg, it is interesting to consider the performance of Bloomberg monetary policy expectations data versus the futures market, which seems a hard-to-beat benchmark. I extract expectations for the Fed funds rate from the futures market with the methodology introduced by Kuttner (2001) and subsequently used by Bomfim (2003), among others[2].

There are 72 policy meetings since 1998 for which Bloomberg provides mean expectations. Table 3 shows the errors in the expectations for the futures market and the Bloomberg survey.

[Insert Table 3 about here]

Table 3 shows that both the Bloomberg survey and the Federal funds futures market are very accurate in predicting target rate changes. The mean absolute difference between the actual target rate and the predicted rate is less than 2.5 bp. The error of 2.45 for the Bloomberg survey is virtually identical to the 2.44 bp of the Federal funds futures market. The standard deviation of absolute differences and the maximum are both slightly higher for the Bloomberg survey. The middle and last panel separate the 72 FOMC meetings into 35 meetings with and 37 without changes in the target rate. Interestingly, the Fed funds futures market is somewhat more accurate than the survey when there is a change in the

[2] In the absence of changes in the forward premium, the surprise can be computed as the scaled one-day change in the first-nearby Fed funds future contract. The futures market expectation is defined as the announced target rate minus the surprise in the futures contract.

target rate (4.32 bp versus 3.47 bp). The Bloomberg survey, on the other hand, is more accurate when the target rate is unchanged (0.67 bp versus 1.46 bp). Prior studies conclude that the Fed funds futures rate outperforms survey data and represents the most efficient prediction of the target rate. Table 3 shows that the accuracy of the Bloomberg survey is virtually identical to that of the futures market. In the light of this evidence and in the interest of consistency with the other variables, I use the Bloomberg survey for Fed funds target rate predictions in remainder of this paper.

3. Methodology

3.1. Univariate GARCH models

I use the standard AR(1)-GARCH(1,1) model of Bollerslev (1986) as the baseline model for the univariate analyses without news announcements. I extend the model to incorporate macroeconomic news in a framework comparable to Jones, Lamont and Lumsdaine (1998):

$$r_t = \mu + \gamma r_{t-1} + \beta S_t + \varepsilon_t \tag{4a}$$

$$\varepsilon_t \mid \Omega_{t-1} \sim N(0, h_t) \tag{4b}$$

$$h_t = s_t\left(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1}\right) \tag{4c}$$

$$s_t = 1 + \lambda_{Ann} D_t^{Ann} + \lambda_{Ann-} D_{t-1}^{Ann} + \lambda_{Ann+} D_{t+1}^{Ann} \tag{4d}$$

The return on stocks or bonds (r_t) depends on a constant, the lagged return (r_{t-1}), and the pooled macroeconomic surprise indicator S_t. The error term has a conditional mean of zero and time-varying variance h_t. The function s_t scales the conditional variance up multiplicatively as a function of an announcement day dummy (D_t^{Ann}) and pre- and post-announcement day dummies (D_{t+1}^{Ann} and D_{t-1}^{Ann}, respectively). The unconditional variance is used as the starting value for the conditional variance series. This set-up allows conditional variance (h_t) to differ on announcement and non-announcement days by a factor $1 + \lambda_{Ann}$. The same holds for pre- and post-announcement days.

Compared to Jones, Lamont and Lumsdaine (1998), the model is extended in two ways. First, it includes actual surprises, rather than announcement day dummies in the mean equation. Second, the model also explicitly incorporates pre- and post-announcement day dummies in the conditional variance specification. Both modifications turn out to be empirically important.

3.2. Multivariate GARCH models

Multivariate GARCH models are only gradually becomeing more popular in the empirical financial economics literature. Bauwens, Laurent and Rombouts (2006) and Silvennoinen and Teräsvirta (2007) provide recent overviews. The multivariate model specifications that I use are extensions of the univariate models. The univariate model without macroeconomic information generalizes to the Vech-model of Bollerslev, Engle and Wooldridge (1988)[3]:

$$\mathbf{r}_t = \boldsymbol{\mu} + \boldsymbol{\gamma}\mathbf{r}_{t-1} + \boldsymbol{\varepsilon}_t \tag{5a}$$

$$\boldsymbol{\varepsilon}_t \mid \Omega_{t-1} \sim N(\mathbf{0}, \mathbf{H}_t) \tag{5b}$$

$$\mathbf{H}_t = \{h_{ij,t}\} \tag{5c}$$

$$h_{ij,t} = \alpha_{ij,0} + \alpha_{ij,1}\varepsilon_{i,t-1}\varepsilon_{j,t-1} + \alpha_{ij,2}h_{ij,t-1} \tag{5d}$$

The returns on stocks and bonds ($\mathbf{r}_t$) depend on constant terms ($\boldsymbol{\mu}$) and their own lagged returns ($\mathbf{r}_{t-1}$), i.e. a vector autoregressive model of order 1 where $\boldsymbol{\gamma}$ is diagonal. The error terms are conditional normally distributed with zero mean and time-varying covariance matrix $\mathbf{H}_t$. Each element $h_{ij,t}$ of the covariance matrix depends on a constant ($\alpha_{ij,0}$), an ARCH-term ($\varepsilon_{i,t-1}\varepsilon_{j,t-1}$), and a GARCH-term ($h_{ij,t-1}$). Incorporating macroeconomic news effects follows analogously:

$$\mathbf{r}_t = \boldsymbol{\mu} + \boldsymbol{\gamma}\mathbf{r}_{t-1} + \boldsymbol{\beta}\mathbf{S}_t + \boldsymbol{\varepsilon}_t, \tag{6a}$$

$$\boldsymbol{\varepsilon}_t \mid \Omega_{t-1} \sim N(\mathbf{0}, \mathbf{H}_t) \tag{6b}$$

$$\mathbf{H}_t = \{h_{ij,t}\} \tag{6c}$$

$$h_{ij,t} = s_{ij,t}\left(\alpha_{ij,0} + \alpha_{ij,1}\varepsilon_{i,t-1}\varepsilon_{j,t-1} + \alpha_{ij,2}h_{ij,t-1}\right) \tag{6d}$$

$$s_{ij,t} = 1 + \lambda_{ij,Ann}D_t^{Ann} + \lambda_{ij,Ann-}D_{t-1}^{Ann} + \lambda_{ij,Ann+}D_{t+1}^{Ann} \tag{6e}$$

The parameters and variables are as discussed above. The model can be estimated by (quasi) maximum likelihood and the log-likelihood function of the multivariate model is:

$$\ell(\boldsymbol{\theta}) = -\frac{1}{2}nT\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\ln|\mathbf{H}_t| - \frac{1}{2}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\mathbf{H}_t^{-1}\boldsymbol{\varepsilon}_t, \tag{7}$$

where $\boldsymbol{\theta}$ is the vector of parameters, n is the number of asset, T is the number of observations, and $|\mathbf{H}_t|$ is the determinant of the time-varying covariance matrix at time *t*.

[3] The Vech operator stacks the lower triangle of a matrix into columns. Since the covariance matrix is symmetric, only the lower triangle of the matrix has to be modeled.

The Vech-model is probably the most general multivariate GARCH model available, as parameters are not restricted. The drawback of this specification, however, is that the conditional covariance matrix is not guaranteed to be positive semi-definite for each time t. An additional issue is the large number of parameters that has to be estimated. Finally, stationarity conditions have to be imposed on the parameter estimates. These factors make estimation of the multivariate models far from straightforward, even in a setting with only two assets. Standard gradient search optimization routines have great difficulty in optimizing the likelihood and the final convergence is highly dependent on the chosen starting parameters. The literature has traditionally dealt with this problem in two ways: (1) by using parametrizations that assure that the covariance matrix is positive semi-definite or (2) by directly imposing constraints on the parameters. A popular example of the former is the BEKK model of Engle and Kroner (1995), where parameter matrices are squared. A full discussion of the various parametrizations of multivariate GARCH models that achieve positive semi-definiteness is offered in Ding and Engle (2001). The drawback of many of these parametrizations, however, is that coefficient estimates are no longer easy to interpret. To overcome this problem, various researchers constrain the parameters directly. But this presents an additional hurdle to the optimization algorithm which is extremely hard to overcome by traditional optimization routines. Researchers have therefore chosen not to impose the constraints directly, but merely check whether the constraints are satisfied in the non-constrained optimum. It is unclear what can be done if the constraints are not satisfied.

The few prior empirical studies that estimate multivariate GARCH models with news effects impose constant correlations (Christiansen, 2000) or sacrifice efficiency by estimating the model in two stages (De Goeij and Marquering, 2006, 2008; Andersen, Bollerslev, Diebold and Vega, 2007; and Brenner, Pasquariello and Subrahmanyam, 2008). I opt for a different solution. First, I impose the stationarity conditions directly on the parameters (i.e. $\alpha_{ij,1} + \alpha_{ij,2} \leq 1$ in equations 5(d) and 6(d)). Second, to guarantee positive semi-definiteness of the covariance matrix for each observation, I impose a penalty if it is not. This is implemented by returning an arbitrarily large negative likelihood value if not all eigen values of the covariance matrix are larger than zero. This follows the approach of De Goeij and Marquering (2006, 2008), among others. I find this more attractive than imposing constraints on the parameters directly that (potentially) violate the nature of the data. However, this implementation calls for a robust optimization routine that can deal with discontinuous likelihood functions.

I use the simulated annealing algorithm developed by Corana, Marchesi, Martini and Ridella (1987) and slightly modified and extensively tested by Goffe, Ferrier and Rogers (1994). Simulated annealing is very robust. The algorithm starts by globally exploring almost the entire surface and subsequently closing in on the optimum. In this process, both changes in the parameter vector that increase (“uphill moves”) and decrease the likelihood function (“downhill moves”) are accepted. The algorithm is therefore capable of escaping local maxima, which is different from most traditional optimization routines. An additional benefit of this approach is that simulated annealing is practically insensitive to the starting parameters. Finally, simulated annealing can deal with functions that are not defined for certain parameters. Since few assumptions are made on the function to be evaluated, the algorithm is quite robust and hence very useful in the present challenging context. The benefits of the simulated annealing algorithm come at the cost of substantially longer computing times. However, in comparison to multiple runs of a traditional optimizer and in the light of the poor convergence, simulated annealing is attractive. I provide a brief description of the steps of the algorithm in the Appendix, but the interested reader is referred to Corana, Marchesi, Martini and Ridella (1987) and Goffe, Ferrier and Rogers (1994) for more details.

I initialize the parameters of the multivariate GARCH model with the parameter estimates from the univariate estimates. The first observation of the conditional covariance matrix is set to the unconditional sample covariance matrix. Although simulated annealing is not sensitive to the starting parameters, I re-estimate all models with the Broyden-Fletcher-Goldfarb-Shanno algorithm. Starting values are randomly selected within two times that standard errors from the multivariate model obtained with simulated annealing. None of the log likelihood values exceeds the original log likelihood value obtained with simulated annealing. It is therefore unlikely that the reported results reflect a local instead of a global maximum.

3.3. Diagnostics and hypothesis testing

Diagnostic checking for multivariate GARCH models is still at its infancy (see also Ding and Engle, 2001) and empirical studies do often not contain a discussion of the statistical properties of the model. I implement the diagnostic tests developed by Bollerslev and Wooldridge (1992) and Ding and Engle (2001). If the multivariate model is correctly specified with known parameters, the standardized residuals are conditionally i.i.d. with mean zero and the identity matrix as covariance matrix:

$$\mathbf{e}_t = \mathbf{H}_t^{-1/2} \boldsymbol{\varepsilon}_t \mid \Omega_{t-1} \sim N(\mathbf{0}, \mathbf{I}) \tag{8}$$

Although the errors may not be normally distributed, Bollerslev and Wooldridge (1992) show that quasi maximum likelihood is still consistent if the moment conditions that follow from (8) are satisfied. I therefore test:

$$E(e_{i,t}) = E\left(\frac{\varepsilon_{i,t}}{\sqrt{h_{ii,t}}}\right) = 0 \text{ and } E(e_{i,t} e_{j,t}) = E\left(\frac{\varepsilon_{i,t}\varepsilon_{j,t}}{h_{ij,t}}\right) = 1 \text{ for } i, j \in \{EQ, FI\}\,. \tag{9}$$

Each table shows tests of the moment conditions in (9) and checks whether the means of the standardized residuals are zero and means of the standardized products of residuals are one. Furthermore, I report Bollerslev and Wooldridge (1992) standard errors that are robust to departures from normality. I also check for serial correlation in standardized residuals and standardized products of the residuals up to the fifth lag using the Ljung-Box statistic.

Within this set-up, various hypotheses of interest can be tested. First, I test whether the multivariate model specification of stocks and bonds jointly is favored over the univariate models for stocks and bonds separately. I test this hypothesis for models with and without macroeconomic news. Second, I test whether all macroeconomic news is redundant and can be excluded from the model. This test compares univariate and multivariate models with and without news with each other. Third and finally, I test whether the aggregate impact of macroeconomic information over all announcement days is zero. The test is implemented by estimating a restricted model where the sum of the λ_{ij} −parameters in equation (4d) of the univariate model and equation (6e) of the multivariate model equals zero for stocks, bonds or their covariance. The hypothesis is tested for univariate as well as multivariate model specifications with news effects. In the multivariate context, restricted models are estimated three times where the restriction is imposed for stocks, bonds or the covariance between stocks and bonds. All tests take the form of a likelihood ratio test where the difference in restricted and unrestricted likelihoods is compared and tested for significance using a χ^2 - distribution.

4. Conditional variances, covariances and macroeconomic announcements

4.1. Univariate results

The results from the univariate GARCH(1,1)-models with and without macroeconomic news surprises are presented in Table 4. Specifically, the left panels show parameter estimates, standard errors, diagnostics and hypothesis tests for the model without news and the right panel with news. Panel A provides the results for the S&P 500 future and Panel B for the 10-year Treasury bond future.

[Insert Table 4 here]

Examination of Table 4 indicates that macroeconomic news has a significant impact on stocks as well as on bonds. Using a likelihood ratio test, the model without macroeconomic news is strongly rejected in favor of the model with macroeconomic news. The associated p-values are less than 1% for equities and fixed income.

Higher economic activity and lower consumer and producer prices lead on average to higher equity returns and lower bond returns. A one-standard deviation shock in aggregate macroeconomic indicator S_t leads, on average, to a return of 2.5 bp. in equities and -7.7 basis points in bonds. This effect is significant for bonds at the 1% level, but not for stocks. For stocks, the conditional variance on announcement days is a factor $1+\lambda_{Ann}=1.115$ higher than on non-announcement days. In other words, the conditional variance increases (an insignificant) 11.5% for stocks on announcement days and a significant 31.4% for bonds. The conditional variance on pre-announcement days is 13.2% higher for stocks, which is significant at the 10% level, and 3.6% for bonds, which is not significant. The strongest effect that both markets have in common is the reduction in conditional variances on post-announcement days. The higher conditional variance on announcement days is offset by a 21.9% lower conditional equity market variance and a 30.8% lower conditional bond market variance on post-announcement days. For both markets, these effects are significant at the 1% level. Both stock and bonds markets show significant relief after the announcement of macroeconomic news. This suggests that news on macroeconomic fundamentals resolves uncertainty among market participants, consistent with Pasquariello (2007).

Although there is substantial variation in conditional variances on and around announcement days, the overall impact of all announcements together is insignificantly different from zero for the equity market (p-value of the likelihood ratio test is 0.12) For the Treasury bond market, on the other hand, the aggregate impact is significantly different from zero (p-value of 0.005).

The conditional mean return is small, but significant for equities. The lagged equity return is negative, but not significantly from zero. For the bond market, on the other hand, the autoregressive term is positive and significantly different from zero at the 5% level. For equities and bonds, the ARCH- and GARCH-terms are both significantly different from zero. Parameter estimates of the constant and the lag in the mean equation only slightly change if macroeconomic news is included in the model. The same holds for the constant, ARCH- and GARCH-term in the conditional variance equations.

The models show no sign of misspecification: the Ljung-Box statistic for autocorrelation up to the 5th lag is insignificantly different from zero. This holds for both stocks and bonds and for normal residuals and squared residuals. This shows that the GARCH models are successful in removing serial correlation from the return and squared return series (see Table 1).

4.2. Multivariate results

Table 5 shows the estimation results for the multivariate models without (left panel) and with (right panel) macroeconomic announcements.

[Insert Table 5 here]

Macroeconomic news also matters significantly in the multivariate context. The null-hypothesis of redundancy of all macro-related variables is overwhelmingly rejected (p-value 0.00). The same general patterns are observable as in the univariate cases. In addition to conditional variances, the conditional stock-bond covariance shows significant post-macroeconomic-announcement relief. The conditional stock market variance drops 11.7%, the conditional bond market variance 28.6% and the conditional stock-bond covariance drops 23.4%. Announcement days typically have higher conditional variances (8.3% for stocks and 28.9% for bonds) and covariances (11.7%). The impact of announcement days on conditional second moments, however, is only statistically significant for the bond market. On pre-announcement days, conditional variances and covariances are on average somewhat higher, but not significantly so. Even more pronounced than in the univariate cases, the null hypothesis that the aggregate impact of news on and around announcement days is zero cannot be rejected. For equities, the associated p-value is 0.17, for bonds 0.15 and for the covariance 0.17. This suggests that the impact of news announcements on the covariance matrix is a transitory, rather than a permanent effect.

The effect on the mean return is comparable to the results from the univariate models: an insignificant and positive 2.6 bp. for equities and a significant -7.9 bp. for fixed income. This reiterates that good news is usually good news for stock returns, but bad news for bond returns during the sample period. It is comforting to see that the estimates of the remaining parameters are comparable for the multivariate models with and without news. In addition to that, the conclusions are very similar for the univariate and the multivariate models. The likelihood ratio test, however, strongly prefers the multivariate GARCH specification above the univariate GARCH specifications. Both for the model with and without news, the univariate models are significantly rejected in favor of the multivariate specifications at the 1% level of significance.

The diagnostic tests for the moment conditions of the standardized residuals and standardized (cross-) products of residuals show no apparent misspecifications. However, the mean of the standardized equity residual is significantly different from zero for both the model with and without macroeconomic news. Furthermore, the Ljung-Box test for autocorrelation up to the fifth lag is significant for the standardized bond residuals for the model without news and for the standardized equity residuals for the model with news (at the 10% level).

4.3. Key versus non-key announcements

Previous studies have primarily focused on releases of nonfarm payrolls, CPI, PPI and monetary policy. Jones, Lamont and Lumsdaine (1998) for example, study the effect of employment and producer prices announcements on Treasury bonds and Li and Engle (1998) examine the link between fixed income markets and CPI, PPI, and unemployment announcements. Christiansen (2000) assesses the impact of employment situation and PPI releases on the conditional covariance structure of Treasury bonds. Fleming and Remolona (1999) study CPI, PPI, and employment announcements in a high-frequency setting using Treasury bond tick data. Bomfim (2003) analyzes the impact of Fed funds target rate announcements on the stock market. More recently, Brenner, Pasquariello and Subrahmanyam (2008) examine the impact of Fed funds target rate, CPI and unemployment announcements on stocks, Treasury bonds, corporate bonds, and their interaction. In the interest of comparability with prior research and as a robustness check, I split the announcements in two groups. The first group contains the variables on which prior literature has focused primarily: nonfarm payrolls, CPI, PPI, and the Fed funds target rate. This is the group with key indicators. The second group contains the rest of the variables, i.e. consumer confidence, Chicago purchasing managers, ISM manufacturing, the trade balance, retail sales less autos, industrial production, housing starts, durable goods orders, and GDP announcements. This is the group with non-key indicators. I re-estimate the multivariate GARCH model with news effects from equations (6a) – (6e), but replace the aggregate pooled announcement series with the pooled indicator that consists of variables from either the first or the last group. This robustness exercise gives insight into whether the main conclusions from the previous paragraph are robust effects in the data or whether they are confined to certain (groups of) announcements. Table 6 shows the results.

[Insert Table 6 here]

The Table shows that the main findings for all variables are robust when they are partitioned in key- and non-key announcements. The likelihood ratio tests for redundancy of all macroeconomic variables are 91.3 for key announcements and 63.92 for non-key announcements with associated p-values less than 0.01 in both cases (full results not shown in the table).

Post-macroeconomic-announcement relief is also a robust feature of the data. Conditional stock market variances, conditional bond market variances, and conditional stock-bond covariances always decrease on post-announcement days. With the exception of the reaction of the equity market on days following non-key announcements, all estimates are statistically significant. The relief after key announcements, however, is in general stronger than after non-key announcements. For equities, the drop in conditional variances is 11.3% versus 8.7%, for bonds this amounts to -32.6% versus -10.8% and for the covariance it is -23.5% versus -18.7%. The announcement-day effect is positive, but insignificant for stock market variance (15.8% for key- and 3.3% for non-key announcements) and the stock-bond covariance (13.4% for key- and 18.1% for non-key announcements). Consistent with the previous findings, the conditional bond market variance is higher on announcement days: 35.9% for key- versus 20.4% for non-key announcements, both statistically significant. The impact of pre-announcement days is never statistically significant.

The effects of announcements on the mean return are also consistent with the results for all variables: positive, but insignificant for equities (5.7 bp. for key- and 1.9 bp. for non-key announcements) and negative and significant for bonds (-7.9 bp. for key- and -6.7 bp. for non-key announcements). The rest of the parameters (constant and lag in the mean equation and constant, ARCH, and GARCH in the variance equation) are very similar for both specifications. These estimates are also comparable to the estimates for the all variables together.

The diagnostic tests reveal no sign of serious misspecification and are very similar to the results for the multivariate model from Table 5 for all announcements. The mean of the standardized residuals is significantly different from zero for the models without and with news. The Ljung-Box test for autocorrelation is only marginally significant for standardized fixed income residuals and insignificant for the rest of the residuals and cross-products of residuals.

5. The impact of macroeconomic announcements on Value-at-Risk estimates

The previous sections argue that macroeconomic news releases significantly affect variances and covariances of stocks and bonds on announcement days, but also on days before and especially after scheduled news. This variation is potentially important for the risk management of asset portfolios, but somewhat surprisingly, previous literature has largely overlooked the practical implications of this finding. Studies in the GARCH literature typically use in-sample statistical loss functions for model evaluation purposes. Recently, De Goeij and Marquering (2008) extend the literature by examining the out-of-sample economic value of time-varying covariance predictions. They show that a mean-variance investor would be willing to pay significantly to use dynamic multivariate GARCH-based covariances instead of employing a passive strategy. They do not, however, take variation in covariances due to macroeconomic announcements into account. If the covariance structure of stocks and bonds changes around release dates of macroeconomic news, the risk characteristics of a portfolio consisting of stocks and bonds changes. This, in turn, affects the required capital for commercial banks and prop desks, for example. If conditional variances and covariances change predictably around days with macroeconomic news, incorporating this information may enhance predictions for risk management. In this section, I compare the forecasting performance of GARCH models with and without macroeconomic information using an out-of-sample Value at Risk application. I test whether VaR provisions are more precise when macroeconomic news is incorporated. This analysis gives insight into the economic significance of macroeconomic information in the context of risk management.

Starting at the end of 2000, I re-estimate the GARCH models at the end of each year with the information available up to that point. The annual frequency corresponds to roughly 250 observations, the number of observations recommended by the Basle Committee on Banking Supervision (1996). Based on the expanding parameter estimates, lagged returns, and dummies for announcement days, I produce conditional one-day-ahead VaR predictions. Since the surprise component of an announcement is not known to market participants a priori, I replace S_t in equations (4a) and (6a) with the announcement day dummy variable, D_t, when I form predictions for conditional variances and covariances. Note that VaR provisions change as a function of past returns and announcement dummies (which are known ex-ante), but that parameters are only re-estimated each year. Given the large number of parameters in the multivariate GARCH model, more frequent re-estimation is cumbersome. For each of the univariate GARCH models for stocks or bonds with and without macroeconomic information, I calculate the VaR provision as:

$$VaR = -N(level)^{-1} \times \sqrt{\hat{h}_{t|\Omega_{t-1}}}\ , \tag{10}$$

where $N(level)^{-1}$ is the inverse cumulative normal distribution function at the desired confidence level, and $\sqrt{\hat{h}_{t|\Omega_{t-1}}}$ is the conditional volatility estimate using information up to and including *t-1*[4]. I use the multivariate GARCH-model to form VaR provisions for a portfolio that consists of stocks and bonds:

$$VaR = -N(level)^{-1} \times \sqrt{\mathbf{w}'\hat{\mathbf{H}}_{t|\Omega_{t-1}}\mathbf{w}}\ , \tag{11}$$

where $-N(level)^{-1}$ is defined above, $\hat{\mathbf{H}}_{t|\Omega_{t-1}}$ is the predicted conditional covariance matrix at time *t* conditional on the information set up to *t-1*, and $\mathbf{w} = \begin{bmatrix} w^{EQ} \\ w^{FI} \end{bmatrix}$. In the empirical application, I use $w^{EQ} = w^{FI} = 0.5$, but varying the weights between 30% and 70% leads to the same general conclusions which are available upon request.

The Basel Committee on Banking Supervision (1996) has established a framework for evaluating internal market risk models. The framework distinguishes between three zones: green, yellow and red. The zones differ in the probability of erroneously accepting an incorrect risk model. The boundaries are determined using binomial probabilities. Following the definition of Value at Risk, one would expect the number losses greater than the VaR allowance to be equal to $(1-level)\times T$, where $level$ is the VaR confidence level (95% or 99%, for example) and T is the number of observations. In a sample of 100 trading days, for example, one would expect 1 (5) day(s) with a loss greater than the VaR provision at the 99% (95%) confidence level.

Some authors criticize the coverage tests for their poor performance in small samples, like for example the 250 days recommended by the Basel Committee on Banking Supervision (1996). There are two remedies to this problem: (1) use a lower confidence interval or (2) use more data. Both solutions increase the number of violations and hence the statistical accuracy. I use both recommended solutions and report results for the full out-of-sample period (2001 – 2007) and show results for the 99% and

[4] I.e., at the 99% (95%) confidence level, the conditional volatility is multiplied with 2.3263 (1.6449) to arrive at the VaR provision.

95% level of confidence. In addition, I report the accuracy of the various models for each out-of-sample year individually at the 99% level, which is closest to the original proposal of the Basle Committee on Banking Supervision. Furthermore, I implement three additional tests that have been developed in response to the original framework and which alleviate some of the statistical concerns. A more elaborate discussion of these tests to assess the accuracy of the various VaR predictions can be found in Jorion (2006). Pérignon, Deng and Wang (2008) use these tests to investigate whether the six largest Canadian commercial banks systematically overstate their Value-at-Risk. Martens and Poon (2001) also use a Value-at-Risk framework to analyze the accuracy of various volatility forecasting procedures. Their focus, however, is on the impact of different trading hours on volatility estimates and their models do not contain information other than past returns.

Kupiec (1995) develops a log-likelihood ratio test to formally test whether the actual number of violations differs significantly from the expected number (see also Jorion, 2006):

$$LR_{uncon} = -2\ln\left[(1-p)^{T-X} p^{X}\right] + 2\ln\left[\left(1-\frac{X}{T}\right)^{T-X}\left(\frac{X}{T}\right)^{X}\right], \tag{12}$$

where X is the number of violations, T is the number of observations and $p = (1 - level)$. The test statistic asymptotically follows a Chi-square distribution with one degree of freedom, $LR_{uncon} \sim \chi^2(1)$. It should be noted that this test is not defined if the number of VaR violations is zero. Jorion (2006) shows that this test is asymptotically equivalent to:

$$z_{uncon} = \frac{X - pT}{\sqrt{p(1-p)T}}, \tag{13}$$

where the symbols are as defined above. The test statistic follows a standard normal distribution. Although equivalence only holds asymptotically (and hence not necessarily in small samples), this alternative formulation is also defined if there are no VaR violations.

In addition to the unconditional tests, I also apply the conditional test developed by Christoffersen (1998). This test takes into account that VaR violations may exhibit clustering over time. The test statistic is:

$$LR_{con} = LR_{uncon} + LR_{ind} \tag{14a}$$

$$LR_{ind} = -2\ln\left[\left(1-\frac{X}{T}\right)^{T_{00}+T_{10}}\left(\frac{X}{T}\right)^{T_{01}+T_{11}}\right] + 2\ln\left[(1-\pi_0)^{T_{00}}\,\pi_0^{\,T_{01}}\,(1-\pi_1)^{T_{10}}\,\pi_1^{\,T_{11}}\right], \tag{14b}$$

where T_{xy} is the number of days where the state of the previous day is x and the state of the current day is y. $\pi_x = \dfrac{T_{x1}}{T_{x1}+T_{x0}}$ measures the probabilities of observing a VaR violation on a particular day while the previous day's state was x. The test statistic follows a χ^2–distribution with 2 degrees of freedom. Following Christoffersen (1998), I replace the second term of equation (14b) with $2\ln\left[(1-\pi_0)^{T_{00}}\,\pi_0^{\,T_{01}}\right]$ if there are no two subsequent violations in the period of analysis.

Table 7 summarizes the accuracy of out-of-sample Value-at-Risk provisions from the univariate and multivariate GARCH models with and without news. Panel A shows the evaluation of Value-at-Risk provision predictions at the 99% level, whereas Panel B shows the information at the 95% level. All results are for the full out-of-sample period that runs from 2001 to 2007.

[Insert Table 7 here]

Table 7 shows that the VaR provisions generated with the various GARCH models are of good quality. With the exception of fixed income at the 99% level, all variants fall in the green BIS-zone. In general, incorporating macroeconomic news decreases the number of VaR violations. The exception is the 99%-VaR provision for equities, where the number of violations for the model without news is 22 and 23 for the model with news. Both variants, however, are in the green zone. In all other cases, the models with macroeconomic information have a lower number of VaR violations than the models without macroeconomic information. At the 99% level, the number of violations reduces from 32 to 27 for fixed income and from 20 to 16 for the 50-50 portfolio. At the 95% level, the number of violations for equities drops from 90 to 86, from 88 to 82 for bonds and from 77 to 73 for the portfolio. The reduction from 32 to 27 violations for the 99%-VaR bonds models is, however, not enough to bring the model performance in the green zone, since only 17 violations are expected. The levels of significance for the various tests, however, fall from 1% for the model without news to 5% for the unconditional tests and 10% for the conditional test, respectively, for models with news. Overall, the reductions in VaR violations when incorporating macroeconomic fundamentals suggest that some VaR violations occur as a result of macroeconomic announcements. This is consistent with the evidence from the GARCH-model estimates reported in previous paragraphs.

Table 8 shows the accuracy results of the VaR provision estimates for each year in the out-of-sample period. The VaR confidence level is 99% and I use the original binomial test of the Basle Committee on Banking Supervision.

[Insert Table 8 here]

As would be expected from the previous discussion, the VaR provision is sufficient for the majority of years. In forming equity VaR provisions, incorporating macroeconomic announcement information helps to bring the performance of the model from the yellow to the green zone in 2001. The model's performance is in the green zone for each year in the period 2002 – 2006. In 2007, however, the accuracy of both model types (i.e. without and with macroeconomic information) is in the yellow zone. Turning to fixed income, the actual number of violations is somewhat too high in 2001, 2002, and 2003, bringing both model types in the yellow zone. But the performance of the model with macroeconomic information is again better: it reduces the number of violations from 8 to 6 in 2001 and from 8 to 7 in 2003. These reductions, however, are not enough to bring the performance of the model in the green zone during these years. The performance over the remaining period 2004 – 2007 is in the green zone for both model types. Interestingly, the VaR provisions for the portfolio of stocks and bonds using the multivariate GARCH model are more accurate than those for stocks and bonds individually using univariate GARCH models. Except for 2007, all out-of-sample years are in the green zone. In conclusion, the number of VaR violations decreases slightly when macroeconomic information is incorporated. Especially when accuracy is in the yellow zone, adding information about fundamentals helps in reducing the number of violations. These results are of practical importance for financial risk managers. Incorporating macroeconomic announcement information into VaR predictions results in less VaR provision violations, which enhances risk management.

6. Conclusion

Although the covariance matrix of asset returns plays a crucial role in asset allocation, risk management, and derivative pricing, why the covariance matrix changes over time remains elusive. This paper uses a novel dataset from Bloomberg with expectations and real-time macroeconomic announcements to study how stocks, bonds and their covariance changes around macroeconomic news release dates. Previous literature has primarily focused on a single asset class and a relatively small number of key announcements. I contribute to the literature by analyzing stocks and bonds simultaneously and for a broad set of 13 key and non-key macroeconomic announcements over the period 1996 – 2007. In addition, I explicitly consider the behavior of conditional moments on days surrounding macroeconomic announcements. Finally, I do not only analyze the statistical significance of macroeconomic fundamentals on the covariance structure of stocks and bonds, but also the economic significance in the context of risk management.

Summarizing, the results in this paper unambiguously show the importance of macroeconomic news in explaining conditional means, variances and covariances of stocks and especially bonds. Good economic news significantly decreases conditional bond returns and increases stock returns. The impact of news, however, is much stronger for the bond market than for the stock market. On average, conditional variances and covariances are higher on days when macroeconomic news is released. In contrast, financial markets show significant relief on days immediately following macroeconomic announcements. Conditional variances of stocks and bonds and the stock-bond covariance drop as much as 30% on post-announcement days. This decrease largely offsets the initial increase on announcement days and the results suggest that the aggregate impact of news around announcement days has a transitory, rather than permanent impact on the covariance structure of stocks and bonds during the sample period. The empirical results are very robust and show up in both samples when I partition the data in key and non-key announcements.

To study the economic significance of the findings, I compare the forecasting performance of GARCH models with and without macroeconomic information in an out-of-sample Value-at-Risk application. Although all GARCH specifications (univariate, multivariate, with and without news) provide VaR estimates of good quality, adding macroeconomic information enhances accuracy. Especially for periods when model performance is in the "yellow zone" as defined by of the Basle Committee on Banking Supervision (1996), models that incorporate macroeconomic information have lower prediction errors.

In conclusion, macroeconomic news matters in modeling changes in the covariance structure of stocks and bonds. There are pronounced patterns in conditional first and second moments around macroeconomic release dates. Furthermore, GARCH models with news effects are of practical relevance to financial risk managers. Incorporating macroeconomic news announcements results in more accurate out-of-sample Value-at-Risk predictions and less violations of the estimated VaR provision.

Appendix. The Simulated Annealing Algorithm

I employ the simulated annealing algorithm of Corana, Marchesi, Martini and Ridella (1987), slightly modified and extensively tested by Goffe, Ferrier and Rogers (1994)[5]. Simulated annealing originates from the field of thermodynamics and is best explained by briefly sketching the steps of the algorithm.

The algorithm optimizes the function f(vP), where vP is the vector with initial parameters. The user sets dT (the initial temperature) and vV (the step length for parameter vector vP). The algorithm starts by computing the function value f(vP) at starting parameters vP. Subsequently, a single parameter *i* of the parameter vector is changed: $vP_i^* = vP_i + dRnd_{[-1,1]} \cdot vV_i$, where $dRnd_{[-1,1]}$ is a uniformly distributed random number [-1, 1]. If the new function value at the new parameters, $f^*(vP^*)$, is higher than the former function value, the change in parameters is accepted and $vP = vP^*$. If the new function value at the new parameters is lower than the old function value, the Metropolis criterion is used. More specifically, if $\exp(\frac{f^*(vP^*) - f(vP)}{dT}) > dRnd_{[0,1]}$, the change in parameters is accepted and $vP = vP^*$. Else, the change in parameters is rejected. Note that this step leads the algorithm to accept "downhill" steps. Apart from a random element, downhill steps become less likely with lower values for the function at the new parameters and lower temperatures.

After iN_s changes in the elements of vP, the step length vV is adjusted in such a way that 50% of the proposed parameter changes are accepted. This guarantees that the function is evaluated at many different values for the parameter vector. After iN_T steps through the loops, the temperature dT is reduced: $dT^* = dT \cdot dr$, where dr lies between zero and one. As noted before, decreasing temperatures increase the number of rejections for parameter changes and decrease the step lengths. The algorithm thus closes in on the optimum of the function. The first function evaluation for each temperature reduction is recorded. During all steps of the algorithm, the overall optimum function value is stored.

The algorithm terminates if the last function value is within the set tolerance level from the optimum function value and the function values at each of the temperature reduction points. This evaluation helps to find the global, rather than a local, optimum.

[5] I use the implementation from Charles Bos in Ox, available from www.tinbergen.nl/~cbos/.

References

Andersen, T.G., Bollerslev, T., Diebold, F.X., Vega, C., 2003. Micro effects of macro announcements: real-time price discovery in foreign exchange. American Economic Review 93, 38 – 62.

Basle Committee on Banking Supervision, 1996, Supervisory framework for the use of “backtesting” in conjunction with the internal models approach to market risk capital requirements.

Bauwens, L., Laurent, S., Rombouts, J.V.K., 2006. Multivariate GARCH models: a survey. Journal of Applied Econometrics 21, 79 – 109.

Balduzzi, P, Elton, E.J., Green, T.C., 2001. Economic news and bond prices: evidence from the U.S. Treasury market. Journal of Financial and Quantitative Analysis 36, 523 – 543.

Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307 – 327.

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.

Bollerslev, T., Wooldridge, J.M., 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. Econometric Reviews 11, 143 – 172.

Bomfim, A.N., 2003. Pre-announcement effects, news effects, and volatility: Monetary policy and the stock market. Journal of Banking and Finance 27, 133 – 151.

Boyd, J.H., Hu, J., Jagannathan, R., 2005. The stock market’s reaction to unemployment news: why bad news is usually good for stocks. Journal of Finance 60, 649 – 672.

Brenner, M., Pasquariello, P., Subrahmanyam, M., 2008. On the volatility and comovement of U.S. financial markets around macroeconomic news announcements. Journal of Financial and Quantitative Analysis *forthcoming*.

Christiansen, C., 2000. Macroeconomic announcement effects on the covariance structure of government bond returns. Journal of Empirical Finance 7, 479 – 507.

Christoffersen, P.F., 1998. Evaluating interval forecasts. International Economic Review 39, 841 – 862.

Corana, A., Marchesi, M., Martini, C., Ridella, S., 1987. Minimizing multimodal functions of continuous variables with the ‘simulated annealing’ algorithm. ACM Transactions on Mathematical Software 13, p. 262 – 280.

Cutler, D.M., Poterba, J.M., Summers, L.H., 1989. What moves stock prices? Journal of Portfolio Management 15, 5 – 12.

De Goeij, P., Marquering, W., 2006. Macroeconomic announcements and asymmetric volatility in bond returns. Journal of Banking and Finance 30, 2659 – 2680.

De Goeij, P., Marquering, W., 2008. Stock and bond market interactions with level and asymmetry dynamics: an out-of-sample application. Journal of Empirical Finance, *forthcoming*.

Ding, Z. and R.F. Engle (2001), Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing, Academia Economic Papers 29, 157 – 184.

Ederington, L.H., Lee, J.H., 1995. The short-run dynamics of the price adjustment to new information. Journal of Financial and Quantitative Analysis 30, 117 – 134.

Ehrmann, M., Fratzscher, M., 2005. Exchange rates and fundamentals: new evidence from real-time data. Journal of International Money and Finance 24, 317 – 341.

Engle, R.F., Kroner, K.F., 1995. Multivariate simultaneous generalized ARCH, Econometric Theory 11, 122 – 150.

Flannery, M.J., Protopapadakis, A.A., 2002. Macroeconomic factors *do* influence aggregate stock returns. Review of Financial Studies 15, 751 – 782.

Fleming, M.J., Remolona, E.M., 1999. Price formation and liquidity in the U.S. Treasury market: The response to public information. Journal of Finance 54, 1901 – 1915.

Goffe, W.L., Ferrier, G.D., Rogers, J., 1994. Global optimization of statistical functions with simulated annealing. Journal of Econometrics 60, 65 – 99.

Jones, C.M., Lamont, O., Lumsdaine, R.L., 1998. Macroeconomic news and bond market volatility. Journal of Financial Economics 47, 315 – 337.

Jorion, P., 2006. Value at Risk: the new benchmark for managing financial risk. McGraw-Hill New York, NY, 3rd edition.

Kupiec, P.H., 1995. Techniques for verifying the accuracy of risk measurement models. Journal of Derivatives 3, 73 – 84.

Kuttner, K.N., 2001. Monetary policy surprises and interest rates: evidence from the Fed funds futures market. Journal of Monetary Economics 47, 513 – 545.

Li, L., Engle, R.F., 1998. Macroeconomic announcements and volatility of Treasury futures. University of California working paper 98-27.

Martens, M., Poon, S.-H., 2001. Returns synchronization and daily correlation dynamics between international stock markets. Journal of Banking and Finance 25, 1805 – 1827.

Pasquariello, P., 2007. Imperfect competition, information heterogeneity, and financial contagion. Review of Financial Studies 20, 391 – 426.

Pérignon, C., Deng, Z.Y., Wang, Z.J., 2008. Do banks overstate their Value-at-Risk? Journal of Banking and Finance, forthcoming.

Schwert, G.W., 1989. Why does stock market volatility change over time? Journal of Finance 44, 1115 – 1153.

Silvennoinen, A., Teräsvirta, T., 2007. Multivariate GARCH models, SSE/EFI working paper series in economics and finance 669.

Table 1
Summary statistics: equity and fixed income futures

	Equity		Fixed Income	
Mean	0.017		0.012	
Standard deviation	1.129		0.368	
Skewness	-0.208		-0.345	
Kurtosis	6.646		4.936	
Jarque-Bera	1726	***	524	***
Observations	3076		2977	
Covariance/*correlation*:				
Equity	1.293		-0.045	
Fixed Income	*-0.108*		0.134	
Ljung-Box statistics:				
Q(1)	2.75	*	3.87	**
Q(2)	5.23	*	9.21	***
Q(3)	9.30	**	14.30	***
Q(5)	10.60	*	15.10	***
$Q^2(1)$	125	***	16.8	***
$Q^2(2)$	247	***	41.6	***
$Q^2(3)$	331	***	48.9	***
$Q^2(5)$	416	***	112	***

The upper panel reports the mean, the standard deviation, skewness, kurtosis and Jarque-Bera test for non-normality of S&P 500 and 10-year Treasury bond futures returns. The sample period is 1/1/1996 – 12/31/2007. The middle panel shows the unconditional sample covariance matrix of the futures returns (with the correlation in the lower left corner). The lower panel provides the Ljung-Box Q-statistic for serial correlation in returns and squared returns up to lag 1, 2, 3 and 5.
* statistically significant at the 10% level.
** statistically significant at the 5% level.
*** statistically significant at the 1% level.

Table 2
Summary statistics: macroeconomic announcements

	First obs.	Last obs.	# of obs.	# of forec.	Release time
Consumer confidence	02/25/1997	12/27/2007	131	56	9:00[a]
Chicago Purchasing Managers	04/30/1997	12/28/2007	129	47	9:00[b]
ISM manufacturing	11/01/1996	12/03/2007	134	60	9:00[c]
Change in nonfarm payrolls	01/10/1997	12/07/2007	132	58	7:30
Trade balance	12/19/1996	12/12/2007	133	60	7:30
Producer price index	12/12/1997	12/13/2007	121	63	7:30
Retail sales less autos	06/13/2001	12/13/2007	79	65	7:30
Consumer price index	12/12/1996	12/14/2007	133	63	7:30
Industrial production	11/15/1996	12/14/2007	134	62	8:15
Housing starts	03/17/1998	12/18/2007	118	58	7:30
Durable goods orders	11/26/1997	12/27/2007	122	60	7:30[d]
GDP	02/27/1997	11/29/2007	43	65	7:30
Fed funds target rate	12/23/1998	12/12/2007	72	87	13:15

The table shows the names of the variables, dates of the first and last observation in the dataset (MM/DD/YYYY), the number of observations and the (median) number of forecasters during the sample. The final column shows the typical (Chicago) announcement time.
[a] The announcement time is 8:36 in July 2001 and 8:50 in January 2002.
[b] From January 2007, the time is 8:45.
[c] The announcement time is 8:26 in August 2000.
[d] Depending on the release of GDP, durable goods orders announcements are moved.

Table 3
Accuracy of monetary policy expectations from Bloomberg and the Fed funds futures market

All FOMC meetings (N = 72)			
	Bloomberg	**Fed funds fut.**	**Difference**
Mean abs. diff.	0.0245	0.0244	0.0001
Std. dev. abs .diff.	0.0558	0.0377	0.0355
Min. abs. diff.	0.0000	0.0000	-0.0900
Max. abs. diff.	0.2900	0.1938	0.1963
Target rate change (N = 35)			
	Bloomberg	**Fed funds fut.**	**Difference**
Mean abs. diff.	0.0432	0.0347	0.0085
Std. dev. abs. diff.	0.0747	0.0492	0.0474
Min. abs. diff.	0.0000	0.0000	-0.0900
Max. abs. diff.	0.2900	0.1938	0.1963
No target rate change (N = 37)			
	Bloomberg	**Fed funds fut.**	**Difference**
Mean abs. diff.	0.0067	0.0146	-0.0079
Std. dev. abs .diff.	0.0148	0.0176	0.0153
Min. abs. diff.	0.0000	0.0000	-0.0417
Max. abs. diff.	0.0720	0.0517	0.0316

The table reports the mean, the standard deviation, the minimum and the maximum of the absolute differences between the announced target rate and the predicted target rate as anticipated by the Bloomberg survey mean and the Federal funds future in percentages (%). The last column shows differences between the survey mean and the Fed funds future. The upper panel considers all FOMC meetings, the middle panel meetings with a change in the target rate and the lower panel the meetings without changes in the target rate.

Table 4
Univariate GARCH models: with and without macroeconomic news

	Without news			With news		
	Estimate		**Std. Err.**	**Estimate**		**Std. Err.**
Panel A: Equities						
μ	0.040	**	0.016	0.033	**	0.016
γ	-0.019		0.018	-0.018		0.018
β				0.025		0.029
α_0	0.015	**	0.007	0.014	**	0.006
α_1	0.081	***	0.016	0.081	***	0.017
α_2	0.909	***	0.018	0.908	***	0.022
$\lambda_{Ann.}$				0.115		0.097
$\lambda_{Ann.+}$				0.132	*	0.073
$\lambda_{Ann.-}$				-0.219	***	0.067
No. of obs.	3075			3075		
No. of ann.	-			1194		
$\ell(\hat{\theta})$	-4382.43			-4369.37		
LB Q(5)	7.34			8.00		
LB $Q^2(5)$	5.10			5.56		
Likelihood Ratio tests						
H_0: $\beta = \lambda_{Ann.} = \lambda_{Ann.+} = \lambda_{Ann.-} = 0$				26.12	***	
H_0: $\lambda_{Ann.} + \lambda_{Ann.+} + \lambda_{Ann.-} = 0$				2.40		
Panel B: Fixed Income						
μ	0.009		0.006	0.008		0.006
γ	0.043	**	0.019	0.038	**	0.019
β				-0.077	***	0.012
α_0	0.001	*	0.000	0.001	**	0.001
α_1	0.036	***	0.008	0.041	***	0.010
α_2	0.958	***	0.010	0.956	***	0.011
$\lambda_{Ann.}$				0.314	***	0.061
$\lambda_{Ann.+}$				0.036		0.047
$\lambda_{Ann.-}$				-0.308	***	0.044
No. of obs.	2976			2976		
No. of ann.	-			1193		
$\ell(\hat{\theta})$	-1110.80			-1057.10		
LB Q(5)	8.67			7.62		
LB $Q^2(5)$	2.70			2.65		
Likelihood Ratio tests						
H_0: $\beta = \lambda_{Ann.} = \lambda_{Ann.+} = \lambda_{Ann.-} = 0$				107.4	***	
H_0: $\lambda_{Ann.} + \lambda_{Ann.+} + \lambda_{Ann.-} = 0$				8.04	***	

$\ell(\hat{\theta})$ denotes the log likelihood at the estimated parameters and LB Q(5) (LB $Q^2(5)$) is the Ljung-Box Q-statistic for serial correlation in the residuals (squared residuals) up to the fifth lag. Robust standard errors are computed as in Bollerslev and Wooldridge (1992).
* statistically significant at the 10% level.
** statistically significant at the 5% level.
*** statistically significant at the 1% level.

Table 5
Multivariate GARCH model: with and without macroeconomic news

	Without news			With news		
	Estimate		**Std. Err.**	**Estimate**		**Std. Err.**
μ_{EQ}	0.050	***	0.016	0.047	***	0.015
γ_{EQ}	-0.029		0.018	-0.029		0.018
β_{EQ}				0.026		0.028
$\alpha_{EQ,0}$	0.016	**	0.007	0.017	**	0.007
$\alpha_{EQ,1}$	0.079	***	0.015	0.083	***	0.016
$\alpha_{EQ,2}$	0.911	***	0.016	0.897	***	0.021
$\lambda_{EQ,Ann.}$				0.083		0.077
$\lambda_{EQ,Ann.+}$				0.055		0.064
$\lambda_{EQ,Ann.-}$				-0.117	*	0.060
μ_{FI}	0.008		0.006	0.007		0.005
γ_{FI}	0.037	**	0.018	0.032	*	0.018
β_{FI}				-0.079	***	0.011
$\alpha_{FI,0}$	0.001	**	0.001	0.002	***	0.001
$\alpha_{FI,1}$	0.048	***	0.009	0.052	***	0.009
$\alpha_{FI,2}$	0.942	***	0.011	0.940	***	0.012
$\lambda_{FI,Ann.}$				0.289	***	0.054
$\lambda_{FI,Ann.+}$				0.017		0.043
$\lambda_{FI,Ann.-}$				-0.286	***	0.041
$\alpha_{EQ\ FI,0}$	0.000		0.000	0.000		0.000
$\alpha_{EQ\ FI,1}$	0.054	***	0.013	0.051	***	0.009
$\alpha_{EQ\ FI,2}$	0.938	***	0.015	0.949	***	0.012
$\lambda_{EQ\ FI,Ann.}$				0.117		0.116
$\lambda_{EQ\ FI,Ann.+}$				0.096		0.085
$\lambda_{EQ\ FI,Ann.-}$				-0.234	***	0.076
No. of obs.	2970			2970		
No. of ann.	-			1192		
$\ell(\hat{\theta})$	-5187.35			-5118.27		

Moment conditions

	T-stat		**LB Q(5)**		**T-stat**		**LB Q(5)**	
$e_{EQ,t}$	-2.70	***	8.94		-2.59	***	9.59	*
$e_{FI,t}$	0.65		9.60	*	0.71		9.09	
$e_{EQ,t}e_{EQ,t}$	-0.06		5.75		0.19		6.79	
$e_{FI,t}e_{FI,t}$	-0.03		2.22		0.30		2.46	
$e_{EQ,t}e_{FI,t}$	-0.24		0.00		1.32		0.04	

$\ell(\hat{\theta})$ denotes the log likelihood at the estimated parameters. The t-stat under "moment conditions" indicate whether the standardized residuals and standardized products of residuals are zero and one, respectively. The LB Q(5) columns show the Ljung-Box Q-statistic for serial correlation in the standardized residuals and standardized products of residuals up to the fifth lag. Robust standard errors are computed as in Bollerslev and Wooldridge (1992).

* statistically significant at the 10% level.
** statistically significant at the 5% level.
*** statistically significant at the 1% level.

Likelihood Ratio tests	No News		News	
H_0: univariate vs. multivariate	365.38	***	390.34	***
H_0: $\beta_i = \lambda_{i\ Ann.} = \lambda_{i\ Ann.+} = \lambda_{i\ Ann.-} = 0$ for i = EQ, FI, EQFI	-		138.16	***
H_0: $\lambda_{EQ\ Ann.} + \lambda_{EQ\ Ann.+} + \lambda_{EQ\ Ann.-} = 0$	-		1.84	
H_0: $\lambda_{FI\ Ann.} + \lambda_{FI\ Ann.+} + \lambda_{FI\ Ann.-} = 0$	-		2.12	
H_0: $\lambda_{EQ\ FI,Ann.} + \lambda_{EQ\ FI,Ann.+} + \lambda_{EQ\ FI,Ann.-} = 0$	-		1.88	

Table 6
Multivariate GARCH model: Nonfarm payrolls, CPI/PPI and Fed announcements versus other announcements

	NFP, CPI/PPI, Fed news			Non-key news		
	Estimate		**Std. Err.**	**Estimate**		**Std. Err.**
μ_{EQ}	0.051	***	0.015	0.048	***	0.016
γ_{EQ}	-0.029		0.018	-0.028		0.018
β_{EQ}	0.057		0.043	0.019		0.030
$\alpha_{EQ,0}$	0.017	**	0.007	0.016	***	0.006
$\alpha_{EQ,1}$	0.081	***	0.017	0.082	***	0.015
$\alpha_{EQ,2}$	0.905	***	0.019	0.902	***	0.018
$\lambda_{EQ,Ann.}$	0.158		0.146	0.033		0.075
$\lambda_{EQ,Ann.+}$	-0.035		0.100	0.071		0.077
$\lambda_{EQ,Ann.-}$	-0.113	*	0.068	-0.087		0.069
μ_{FI}	0.008		0.006	0.008		0.006
γ_{FI}	0.033	*	0.017	0.033	*	0.018
β_{FI}	-0.079	***	0.019	-0.067	***	0.012
$\alpha_{FI,0}$	0.002	**	0.001	0.002	***	0.001
$\alpha_{FI,1}$	0.055	***	0.009	0.050	***	0.009
$\alpha_{FI,2}$	0.939	***	0.012	0.940	***	0.012
$\lambda_{FI,Ann.}$	0.359	***	0.104	0.204	**	0.100
$\lambda_{FI,Ann.+}$	0.008		0.067	-0.089		0.064
$\lambda_{FI,Ann.-}$	-0.326	***	0.049	-0.108	*	0.062
$\alpha_{EQ\ FI,0}$	0.000		0.000	0.000		0.001
$\alpha_{EQ\ FI,1}$	0.051	***	0.010	0.060	***	0.013
$\alpha_{EQ\ FI,2}$	0.949	***	0.016	0.940	***	0.020
$\lambda_{EQ\ FI,Ann.}$	0.134		0.147	0.181		0.112
$\lambda_{EQ\ FI,Ann.+}$	0.031		0.124	-0.031		0.088
$\lambda_{EQ\ FI,Ann.-}$	-0.235	***	0.082	-0.187	**	0.091
No. of obs.	2970			2970		
No. of ann.	453			919		
$\ell(\hat{\theta})$	-5141.70			-5155.39		

Moment conditions

	T-stat		**LB Q(5)**		**T-stat**		**LB Q(5)**	
$e_{EQ,t}$	-2.76	***	8.80		-2.63	***	9.38	*
$e_{FI,t}$	0.57		10.10	*	0.63		9.36	*
$e_{EQ,t}e_{EQ,t}$	0.32		6.94		0.04		6.38	
$e_{FI,t}e_{FI,t}$	0.38		1.38		0.10		3.24	
$e_{EQ,t}e_{FI,t}$	1.08		3.17		-0.98		0.00	

$\ell(\hat{\theta})$ denotes the log likelihood at the estimated parameters. The t-stat under "moment conditions" indicate whether the standardized residuals and standardized products of residuals are zero and one, respectively. The LB Q(5) columns show the Ljung-Box Q-statistic for serial correlation in the standardized residuals and standardized products of residuals up to the fifth lag. Robust standard errors are computed as in Bollerslev and Wooldridge (1992).

* statistically significant at the 10% level.

** statistically significant at the 5% level.

*** statistically significant at the 1% level.

Table 7
Economic significance: Accuracy of out-of-sample Value-at-Risk estimates

Panel A:	GARCH without macroeconomic announcements			GARCH with macroeconomic announcements		
99% VaR level	EQ	FI	portfolio	EQ	FI	portfolio
No. of violations	22	32	20	23	27	16
Exp. no. of violations	17.89	17.36	17.31	17.89	17.36	17.31
LR_{uncon}	0.89	9.99 ***	0.40	1.35	4.62 **	0.10
Z_{uncon}	0.98	3.53 ***	0.65	1.21	2.33 **	-0.32
LR_{con}	2.10	10.20 ***	1.86	2.43	5.49 *	2.31
BIS-zone	Green	Yellow	Green	Green	Yellow	Green
Panel B:	GARCH without macroeconomic announcements			GARCH with macroeconomic announcements		
95% VaR level	EQ	FI	portfolio	EQ	FI	portfolio
No. of violations	90	88	77	86	82	73
Exp. no. of violations	89.45	86.80	86.55	89.45	86.80	86.55
LR_{uncon}	0.03	0.02	1.15	0.07	0.28	2.35
Z_{uncon}	0.17	0.13	-1.05	-0.27	-0.53	-1.49
LR_{con}	1.24	0.07	1.84	1.91	0.28	3.45
BIS - zone	Green	Green	Green	Green	Green	Green

The table shows the actual number of VaR violations, the number of expected VaR violations, the unconditional coverage tests (LR_{uncon} and Z_{uncon}) and the conditional coverage test (LR_{con}) at the 99%- (Panel A) and 95% VaR confidence level (Panel B). The expanding in-sample estimation period starts in 1/1/1996 and models are re-estimated at the end of each year starting in 12/31/2000. Out-of-sample, one-day-ahead VaR forecasts are generated by up-dating the models with the most recent daily data available. Univariate GARCH models are estimated to generate VaR estimates for equities (EQ) and fixed income (FI). The conditional variances and covariances of the multivariate GARCH model are used to form the VaR of a portfolio that consists of 50% equities and 50% fixed income.
* statistically significant at the 10% level.
** statistically significant at the 5% level.
*** statistically significant at the 1% level.

Table 8
Number of VaR violations for each out-of-sample year

	EQ		FI		portfolio	
	without	with	without	with	without	with
2001	6	4	8	6	4	4
2002	4	2	6	6	3	2
2003	1	4	8	7	2	2
2004	1	1	3	2	4	1
2005	1	1	2	2	0	0
2006	3	3	1	1	2	2
2007	6	8	4	3	5	5

The table shows the number of VaR violations for each out-of-sample year at the 99% VaR confidence level. The expanding in-sample estimation period starts in 1/1/1996 and models are re-estimated at the end of each year starting in 12/31/2000. Out-of-sample, one-day-ahead VaR forecasts are generated by up-dating the models with the most recent daily data available. VaR predictions in the EQ and FI columns are generated by the univariate GARCH models and by the multivariate GARCH model for equities and fixed income (portfolio) columns. Each left column ('without') shows the violations for the GARCH models without macroeconomic announcements and each right ('with') column with macroeconomic announcements. Shaded grey boxes indicate instances when the number of VaR violations falls within the yellow zone as stipulated by the Bank of International Settlements using binomial probabilities.