Do Macroeconomic Variables Predict Aggregate Stock Market Volatility?

Bradley S. Paye*
Jones Graduate School, Rice University

October 6, 2008

JEL Classification: G12, C22

Keywords: conditional volatility, volatility forecasting, fractional integration, spurious regression, realized variance.

*I am grateful to Graham Elliott, Jeff Fleming, Bruce Lehmann, Barbara Ostdiek, Dimitris Politis, Allan Timmermann, Halbert White and members of the finance group at Rice University for helpful comments and suggestions. I am particularly grateful to Tim Simin and James Weston, who provided detailed comments on an earlier draft of the paper. Heejoon Han provided excellent research assistance. Lastly, I thank James Davidson who graciously answered many questions regarding his fine TSM econometrics software. This paper is a substantially revised version of portions of Chapter 3 of the author's doctoral dissertation in economics (Paye (2004)) at the University of California, San Diego. Bradley Paye, Jones Graduate School of Management, Rice University, 6100 Main Street - MS531, Houston, TX 77005. Phone: (713) 348-4168, Fax: (713) 348-6296, E-mail:bpaye@rice.edu, URL: http://www.ruf.rice.edu/~bpaye.

Do Macroeconomic Variables Predict Aggregate Stock Market Volatility?

Abstract

The paper considers regressions of "realized variance" measures on macroeconomic predictors. The main contribution of the paper is to show that several well known inference issues documented in the literature on stock return predictability are also serious concerns in models that attempt to link aggregate stock return variance to macroeconomic factors. Specifically, we illustrate that spurious regression bias, finite sample bias due to lagged endogenous regressors, and omitted variables bias in a dynamic context are all potential sources of inference problems in realized variance regressions. We advocate a richly specified dynamic model for variance to minimize the impact of these inference problems. Empirically, the paper assesses the strength of evidence that macroeconomic variables forecast stock return variance using over a century of quarterly data. While there is limited evidence of predictive power in the latter portion of our sample period, our overall conclusion is that the extent to which common macroeconomic forecasting variables enhance our understanding of time-variation in volatility is rather limited.

Introduction

Recent empirical research suggests that many of the same variables that forecast expected stock returns also forecast the conditional *volatility* of stock returns. Campbell (1987), Breen, Glosten and Jagannathan (1989), Shanken (1990), Glosten, Jagannathan and Runkle (1993), Whitelaw (1994), Harvey (2001), Lettau and Ludvigson (2003) and Marquering and Verbeek (2005) all present evidence suggesting that some of the same variables popular in forecasting regressions for stock returns also forecast conditional volatility. This paper considers 'realized volatility regressions,' in which squared daily returns are used to build a proxy for unobserved volatility that is then subjected to time series regression analysis. A number of papers, including Schwert (1989a), Lettau and Ludvigson (2003) and Marquering and Verbeek (2005) have explored whether macroeconomic variables predict aggregate volatility within the context of realized volatility regressions.[1]

The question of whether macroeconomic variables predict volatility is an important issue in empirical finance. From an asset allocation perspective, variables that forecast either expected returns or volatility (or both) become state variables in the investor's portfolio optimization problem. From a risk-management perspective, understanding how future aggregate stock market volatility responds to changing macroeconomic conditions is critical to value-at-risk and stress-testing analysis at longer horizons. Finally, evidence of a strongly countercyclical Sharpe ratio is a challenge for modern asset pricing models. Understanding the appropriate conditioning instruments and functional specifications for both expected returns and volatility is thus critical for constructing the empirical 'stylized facts' against which asset pricing models are evaluated.

A large and active literature probes inference issues in the context of regressions of excess stock returns on lagged macroeconomic variables. This paper shows that several of the notable inference issues raised in the stock return predictability literature are also serious concerns in predictive regressions for (realized) variance. In particular, the spurious regression bias studied by Ferson, Sarkissian and Simin (2003a,b), the finite-sample bias studied by Stambaugh (1999),

[1]Another strand of literature examines the links between stock market volatility and the *volatility* of macroeconomic variables. See, for example, Schwert (1989a) and Beltratti and Morana (2002b). This paper focuses on the apparent forecasting power of the *level* of macroeconomic variables on return volatility.

and the omitted variables bias addressed in Butler, Grullon and Weston (2006) all have direct analogs in models that attempt to link macroeconomic variables to variation in stock return volatility.

A large existing literature documents the potential of spurious regression problems when both the dependent variable and regressor exhibit high 'persistence,' where such persistence may reflect unit root behavior, long memory, or even stationary, short memory dynamics that are difficult to distinguish from long memory in finite samples.[2] It is not surprising, then, that spurious regression is a concern in realized variance regressions since both variance and many macroeconomic forecasting variables are quite persistent.

In a set of Monte Carlo simulations, we demonstrate the less obvious result that standard hypothesis tests of the null of no predictability are likely to be significantly oversized in realized variance regressions if the dynamic model for variance is not sufficiently rich, even if the true variance process is stationary, and even if popular kernel-based HAC estimators are employed in constructing t-statistics. This finding is fairly invariant over sample sizes ranging from 100 to 500, which at the quarterly frequency reflect the size of datasets of practical interest.

The empirical evidence indicates that a number of popular variables in the macro-finance forecasting literature are contemporaneously correlated with stock return variance, and therefore these predictors are only weakly exogenous instruments. First, this immediately implies that slope coefficients are biased in finite samples as shown in the return regression setting by Stambaugh (1999). Furthermore, if the regression model for (log) variance is dynamically misspecified and incorporates a weakly exogenous macroeconomic forecasting variable, then the OLS slope coefficient on this forecasting variable is generally biased and inconsistent. In large samples, this slope coefficient will converge in probability to a value that minimizes the mean-square error in the population regression. Intuitively, when a macroeconomic predictor is correlated with volatility innovations, this variable provides useful forecasting information regarding omitted dynamic components. Thus, a spurious forecasting relationship will be uncovered, in the sense that under a correctly specified model the OLS coefficient would instead converge to its true population value of zero (under the null of no predictability). In our Monte

[2] Examples of papers that consider such situations include Ferson, Sarkissian and Simin (2003a,b), Granger, Hyung and Jeon (2001), Granger and Newbold (1974), Kirby (1997), Marmol (1998), Phillips (1986), Phillips (1988), Tsay and Chung (2000) and Valkanov (2003).

Carlo simulation analysis, we show that bias issues can be severe when insufficient dynamics are included in the predictive model for variance.

In addition to documenting inference issues in models linking stock return variance to macroeconomic factors, the paper also presents a reasonably comprehensive set of empirical results regarding predictive ability. This analysis is in the spirit of Goyal and Welch (2007), although we focus on in-sample, rather than out-of-sample, predictive ability. Our results illustrate that support for the predictive ability of macroeconomic variables, in both a statistical and economic sense, is fragile with respect to the assumed dynamics. For richer dynamic specifications that are less susceptible to the inference problems we document, we find significant evidence of predictability ability for relatively few variables, particularly when the sample period extends into the first half of the twentieth century. When we look at the sub-period following the Treasury Accord of 1951, the evidence for predictability is stronger. However, we illustrate that the *economic* significance of this predictability is weak, in the sense that: 1.) the fitted values from models that include the macroeconomic predictors differ little from those obtained under a pure univariate time series model; and 2.) the slope estimates for many forecasting variables translate to relatively little predictive power.

In addition to univariate models (in the sense of a single macroeconomic variable included), we also explore the predictive ability of specifications that include a number of macroeconomic variables together. This produces little substantiative change in the nature of our findings. As in the univariate case, there is only modest statistical support for predictive ability, and the economic significance of macro-predictors appears quite limited. Our main empirical findings are robust to a number of alternative schemes for computing realized variance and to using the level, as opposed to the logarithm, of realized variance as the dependent variable in the models.

Schwert (1989a) explores a variety of explanations for time-variation in volatility, including the potential that volatility may fluctuate with the level of economic activity. Schwert (1989a) does not find strong post-Treasury accord evidence of a relation between aggregate stock market volatility and dividend yields, price-earnings ratios and the default premium.[3] Subsequent research has identified what appear to be stronger relations between volatility and these and

[3]Interestingly, Schwert includes 12 lags of realized volatility in the regression specification, suggesting a keen sensitivity to the spurious regression danger highlighted in this paper. Schwert finds stronger evidence of predictability for the default premium over a longer sample period which includes pre-Treasury accord data.

other macroeconomic variables. The results in this paper largely concur with Schwert's findings, with the caveat that we do find some evidence of predictive ability in the post-Treasury Accord period.

Recent theoretical work by Antonio Mele (2005a,b) suggests that business-cycle variation in stock return volatility may be related to asymmetric variation in risk premia. In short, risk premia that move more violently in down times (recessions) relative to good times (expansions) can generate cyclical variation in volatility. This theory suggests that variables which forecast the degree of variation in risk premia related to business cycle fluctuations should forecast volatility over longer horizons.

To reconcile our empirical results with this theory, we first note that a number of the variables that do appear to have some predictive ability are 'real' economic variables such as Cochrane's investment to capital ratio. Such variables may proxy for the unobserved state of the economy. In terms of the relatively weak economic forecasting power that we document, the quarterly horizon we consider may not be of sufficient length to capture low-frequency predictable variation in volatility over the business cycle. Both Lettau and Ludvigson (2003) and Mele (2005b) find empirical evidence that macroeconomic variables forecast volatility over longer horizons of one or two years. Of course, for longer horizon regressions, the persistence issues addressed in this paper are compounded with strong induced serial correlation due to the use of overlapping data so that statistical inference is far from straightforward. Finally, we explore only a limited set of specifications in this paper. Other specifications, possibly permitting nonlinearities (as in Harvey (2001)) or time- variation in the effects of macroeconomic variables on volatility might yield substantially stronger economic effects.

The remainder of the paper proceeds as follows. Section 1 describes our data, the realized variance measures we employ, and characterizes the persistence of both volatility and macroeconomic variables. Section 2 describes the various distinct inference problems relevant in regressions of variance proxies on macro-factors. Section 3 presents our Monte Carlo simulation analysis. Section 4 presents empirical evidence regarding predictability in both univariate and multivariate linear regression models. Section 5 considers robustness issues associated with our results. Section 6 provides a summary discussion of the paper and potentially fruitful areas for future research.

1 Data Sources, Volatility Proxies, and Persistence in Variance and Macroeconomic Predictors

1.1 Data Sources

We collect daily total returns (capital gain plus dividends) on the S&P500 Index for the years 1885 - 2005. Daily returns for the period 1926-2005 are obtained from CRSP, while daily returns prior to 1926 are based on Schwert (1990).

Our macroeconomic predictors are sourced from Goyal and Welch (2007).[4] The book to market ratio (*b/m*) is the ratio of book value to market value for the Dow Jones Industrial Average. The consumption-wealth-income ratio (*cay*), proposed by Lettau and Ludvigson (2003), is the residual obtained from estimating a cointegrating relation between aggregate consumption, wealth, and labor income. The default return spread (*dfr*) is the difference between long-term corporate bond and long-term government bond returns. The default spread (*dfy*) is the difference between AAA and BAA-rated corporate bond yields. The dividend-price ratio (*dp*) is the log of dividends less the log of prices, while the dividend yield (*dy*) is the log of dividends less the lag of log prices. The earnings-price ratio (*ep*) is the log of earnings less the log of prices. The investment-to-capital ratio (*i/k*) proposed by Cochrane (1991) is the ratio of aggregate investment to aggregate capital for the entire economy. The rate of inflation (infl) is measured as the Consumer Price Index (All Urban Consumers). The long-term return (*ltr*) and long-term yield (*lty*) on government bonds are from Ibottson's Stocks, Bonds Bills and Inflation Yearbook. Net equity expansion (*ntis*) is defined as the ratio of the twelve-month moving sums of net issues by NYSE listed stocks divided by total end-of-year market capitalization of NYSE stocks. The short interest rate (*tbl*) is the secondary rate on the three-month bill. Finally, the term spread (*tms*) is defined as the difference between the long term yield on government bonds and the T-bill rate.[5]

Table **??** presents descriptive statistics for these variables, including the sample period over

[4]We thank Amit Goyal for making the Goyal and Welch (2007) data available for download from his website. Similarly, we downloaded stock returns prior to 1926 from William Schwert's website.

[5]Early in the sample period, some bond data are taken from alternative sources. Readers are referred to Goyal and Welch (2007) for additional details regarding data sources. The variable *ntis* is similar, but not identical, to the net payout measure considered by Boudoukh, Michaely, Richardson and Roberts (2005).

which each variable is available. Many of the 'financial ratio' variables such as *dp* and *ep* are available from 1885:1 onward, while other variables such as *ik* and *cay* are not available until significantly later.

1.2 Volatility Measurement

The conditional variance of a portfolio return is based on *ex ante* expectations and is fundamentally unobservable. The regression-based approach to modeling the conditional volatility of portfolio returns relies on an *ex post* measurement of variance. The time series of *ex post* volatility measurements is then amenable to standard time series regression and modeling techniques.

Following the approach of Taylor (1986), French, Schwert and Stambaugh (1987) and Schwert (1989a) we use squared daily returns on the Standard and Poor's (S&P) composite portfolio to construct an ex post measurement for the variance of excess returns on the S&P 500 portfolio.

A simple measure of the volatility of excess returns on the S&P 500 portfolio is constructed as

$$rvar_t = \sqrt{\sum_{i=1}^{N_t} R_{i,t}^2} \tag{1}$$

where N_t denotes the number of trading days in the t-th period (month or quarter in our empirical application) and $R_{i,t}$ indicates the daily excess return on the S&P 500 portfolio on the i-th trading day of the t-th period. We refer to the volatility measure $rvar_t$ computed in equation (1) as "realized variance." [6]

Recent research by Andersen, Bollerslev, Diebold and Labys (2003) and Barndorff-Nielsen and Shephard (2002) provides a theoretical underpinning for the use of volatility measures such as that described in equation (1). These studies show that the realized variance is a consistent and theoretically error-free estimator of the integrated variance of a frictionless, arbitrage-free asset price process. In practice, microstructure effects combine with fundamental limitations in sampling ability to limit the accuracy of realized variance. When daily squared returns

[6] The daily risk-free rate is computed by converting the 90 day Treasury bill rate for the corresponding month into a daily rate based on the number of trading days. Hence, we implicitly assume a constant risk-free rate within each quarter.

are summed at the quarterly horizon, the latent increment in integrated variance is certainly estimated with non-trivial error. Thus, the measure described by equation (1) is best described as a volatility "proxy." In this paper we abstract from measurement issues and implicitly treat the time series $rvar_t$ as observed volatility, although we do consider the robustness of our main findings to alternative realized variance proxies that incorporate the possibility of time-varying expected returns.

Table 1 presents summary statistics for *rvar* and for the logarithm of realized variance, denoted *lrvar*. The motivation for the logarithmic transformation stems from the heavily skewed and leptokurtotic nature of the realized variance. As Table **??** reveals, the logarithmic transformation removes most of the skewness and excess kurtosis in the series. As previously noted by Andersen, Bollerslev, Diebold and Labys (2003), the distribution of log realized volatility is approximately Gaussian. The plots in Figure 1 illustrate that the transformation dampens the influence of extremely large (absolute) return observations, such as the quarter associated with October 1987. Finally, the logarithmic transformation averts the potential of negative volatility forecasts. For these reasons, we focus on models with (*lrvar*) as the dependent variable in our empirical work. Nevertheless, in the robustness section of the paper we also use the level of variance as an alternative dependent variable and obtain qualitatively similar results.

1.3 Time Series Properties of Realized Variance and Macroeconomic Forecasting Variables

Inspection of the time series plots in Figure 1 suggest that both realized variance and the log of realized variance are persistent. Figure 2 displays correlograms for both *lrvar* and *rvar*. Autocorrelations appear to drop off rather quickly at first and then tend to stabilize around a positive value for longer lags. The fact that the autocorrelations at long lags remain substantially greater than zero suggests the possibility of long-range dependence or 'long memory' in (log) realized volatility. Loosely speaking, long memory processes exhibit a high degree of persistence, i.e., of autocorrelations that do not die out even at very long lags.[7]

[7] Ding, Engle and Granger (1993) document the slow decay rates in the correlogram for daily time series of absolute returns. Andersen, Bollerslev, Diebold and Ebens (2001) find evidence of long memory in realized volatility for stock returns at the daily horizon, where daily realized volatility estimates are constructed from high-frequency intradaily returns data.

In the time domain, a stationary discrete time series exhibits long memory if

$$\lim_{j\to\infty} \frac{\rho_j}{cj^{-\alpha}} = 1 \tag{2}$$

for some constants c and α such that $\alpha \in (0,1)$. The definition in (2) implies that the autocorrelations of the process decay hyperbolically to zero, rather than exponentially as in, for example, a stationary AR(1) time series process.

As an intermediate between I(0) and I(1) and processes, Granger(1980), Granger and Joyeux (1980) and Hosking (1981) introduced fractionally integrated time series, notated as

$$(1-L)^d Y_t = \varepsilon_t$$

where $L(\cdot)$ is the lag operator and ε_t is a white noise series, and d is a potentially fractionally valued quantity that captures the memory of the time series. A time series Y_t with zero mean follows an autoregressive fractionally integrated moving average process (ARFIMA) if

$$\Phi(L)(1-L)^d Y_t = \Theta(L)\varepsilon_t \tag{3}$$

The process Y_t is stationary and invertible if the roots of $\Theta(L)$ and $\Theta(L)$ lie outside the unit circle and the long memory parameter d satisfies $d < |0.5|$. The series displays long memory for positive d while the case of negative d was described by Mandlebrot as *antipersistence*. When $d \geq 0.5$, the time series is nonstationary. Obviously, unit root processes may be viewed as a special case of nonstationary ARFIMA processes for which $d = 1$.

1.4 Univariate Time Series Model Selection for (Log) Realized Volatility

We compare the fit of several standard univariate time series models for *lrvar*. Each of the models is a special case of (3). Table 2 reports parameters estimates, the adjusted-R^2 value, Akaike and Bayesian information criteria ('AIC' and 'BIC,' respectively) and Box-Pierce (Q) statistics for residual correlation based on 12 lags for each specification. Panel A reports results for the over the full sample (1885:2 - 2005:4) while Panel B restricts attention to the sub-period following the Treasury Accord (1951:2 - 2005:4). The latter sub-period is of interest both since US monetary and interest rate policy shifted following the Treasury Accord, and because the period excludes the late 1800s and Great Depression years, where Schwert (1990) argues that volatility differs.

The results in Table 2 illustrate that the richer dynamic specifications offered by the ARMA(1,1) and ARFIMA(0,d,0) models are strongly preferred to the AR(1) and AR(2) alternatives. The latter models achieve a higher adjusted-R^2 and are preferred by both information criteria. This is true over both sample periods examined and is therefore not driven by volatility behavior in the late 1800s or Great Depression. Estimates for the long memory parameter d are close to 0.5, suggesting that log realized volatility displays a degree of long memory that is near the borderline of stationary and nonstationary behavior.[8]

The univariate model selection results illustrate that the correlations implied by fitted AR(1) and AR(2) models inadequately capture the slower decay of the empirical autocorrelations for *lrvar*. The richer ARMA(1,1) and ARFIMA(0,d,0) specifications, by contrast, provide fairly good approximations.[9]

Analyzing the relative performance of the ARMA(1,1) and ARFIMA(0,d,0) specifications is more difficult, as these models perform comparably based on adjusted-R^2. The information criteria are split in terms of the favored model, with the more conservative BIC measure favoring the long-memory specification while AIC prefers the ARMA(1,1) model. We do not take a strong stand regarding which specification is superior. Indeed, in the sequel we will address both possibilities, which helps to illustrate that the main arguments of the paper do not explicitly depend upon the formal existence of long memory in variance.

2 Inference Issues in Predictive Regressions for Variance

A large and active literature probes inference issues in the context of regressions of excess stock returns on lagged macroeconomic variables. This literature is not easy to digest, in part because a number of distinct inference issues have been raised. In this section, we demonstrate that a

[8] Andersen, Bollerslev, Diebold and Ebens (2001) estimate a slightly lower value using daily realized volatilities constructed using high-frequency intradaily returns.

[9] These results are of course in-sample. West and Cho (1995) and Christoffersen and Diebold (2000) find that out-of-sample predictability of volatility for asset returns is limited to short horizons of roughly one to ten days. Recent papers such as Brandt and Jones (2002) and Guidolin and Timmermann (2005) find evidence of out-of-sample predictability at significantly longer horizons (up to two years).

number of the inference issues raised with respect to predictive regressions for stock returns are also relevant concerns in predictive regressions for variance. In particular, we discuss spurious regression bias (e.g., Ferson, Sarkisian and Simin (2003)), finite-sample bias (e.g., Stambaugh (1999)), and bias and inconsistency due to omitted variables in a dynamic context (e.g., Butler, Grullon and Weston (2006)). [10]

All of the inference issues we address may be illustrated in a simple dynamic regression model. Consider the model

$$Y_t = X_{t-1}\beta_0 + \varepsilon_t, \tag{4}$$

where for simplicity we assume that Y_t and X_t are scalar, zero-mean variables. Interest typically attaches to the null of 'no predictability,' that is, the null hypothesis that $\beta_0 = 0$ in the model (4). The OLS estimate of β is equal to

$$\hat{\beta} = \left(\sum_{t=1}^{t=T} X_{t-1}^2\right)^{-1} \left(\sum_{t=1}^{t=T} X_{t-1}Y_t\right). \tag{5}$$

We first summarize standard asymptotic results for the dynamic regression model. Subsequently, we describe potential inference problems in the dynamic regression model and point out how these inference issues are relevant in predictive regressions for variance.

2.1 Standard Inference Results

Prior to understanding what can go *wrong* in predictive regressions for variance, we briefly summarize a set of standard econometric results for dynamic regressions. Suppose that the model in (3) is correctly specified, the vector stochastic process $\{Y_t,\, X_t\}$ is jointly stationary and ergodic, X_{t-1} is predetermined in the sense that $E(X_{t-1}\varepsilon_t = 0)$, and $E(X_t^2 \equiv \sigma_X^2)$ is finite and positive. Then the OLS estimator $\hat{\beta}$ is consistent:

$$\hat{\beta} \xrightarrow{p} \beta_0 = 0. \tag{6}$$

When X_t is additionally *strictly exogenous* for Y_t, in the sense that $E(\varepsilon_t|\, \{X_t\}_1^T)$, then the OLS

[10] Although we do not specifically address long-horizon regressions using overlapping observations (e.g., Boudoukh et al (2005)) in this paper, similar inference issues exist in variance regressions, possibly exaggerated due to the persistence of variance if the regression model is dynamically misspecified.

estimator $\hat{\beta}$ is also unbiased for β_0.

If additional regularity conditions are satisfied such that the series $X_{t-1}\varepsilon_t$ obeys a central limit theorem, then the following asymptotic distribution and hypothesis testing results also hold:

$$\sqrt{T}(\hat{\beta} - \beta_0) \xrightarrow{d} N\left(0, \left(\sigma_X^2\right)^{-1} S \left(\sigma_X^2\right)^{-1}\right) \tag{7}$$

$$t_{\hat{\beta}} = \frac{\hat{\beta}}{\sqrt{\frac{1}{T}\left(\hat{\sigma_X^2}\right)^{-1} \hat{S} \left(\hat{\sigma_X^2}\right)^{-1}}} \xrightarrow{d} N(0,1), \tag{8}$$

where $S \equiv \sum_{-\infty}^{\infty} E(\varepsilon_t X_t X_{t-j}\varepsilon_{t-j})$, $\hat{S}$ is a consistent estimator for S and $\hat{\sigma_X^2}$ is a consistent estimator of σ_X^2, e.g., the sample variance of X_t. In practice, the nonparametric heteroskedasticity and autocorrelation consistent ('HAC') estimator $\hat{S}$ proposed by Newey and West (1987) is often used to construct t-statistics.

2.2 The Danger of Spurious Regressions for Variance

As is apparent from the autocorrelations displayed in Table **??**, the macroeconomic forecasting variables display marked persistence. Among the macroeconomic predictor variables considered, 10 of the 15 first order autocorrelation estimates at the quarterly horizon are in excess of 0.8. The strong persistence of the macroeconomic variables, coupled with the persistence and slow decay in the autocorrelation structure of (log) realized variance, raises the concern of *spurious regression*.

Spurious regression refers to the danger of finding a correlation between unrelated random variables. The classic reference regarding spurious correlation in time series is Yule (1926). The introduction of spurious regression into econometrics is due to Granger and Newbold (1974), who illustrate that when the regressor and regressand in a univariate regression follow stochastically independent random walks there is a possibility of a spurious finding of a relationship between the two variables.

Suppose that X_t is in fact independent of Y_t. Two variations of spurious regression might occur. The first variation occurs when the standard asymptotic results summarized in (4)-

(7) fail to obtain, and in particular in situations where the statistic $t_{\hat{\beta}}$ diverges to infinity as the sample size grows. We refer to this case as 'asymptotic spurious regression.' The second variation of spurious regression refers to instances in which the standard asymptotic results remain valid, but the finite-sample distribution of t-statistics differs significantly from the limiting case, despite the fact that a HAC estimator of the long-run covariance matrix is employed. We will refer to this case as 'virtually-spurious' regression to capture the notion that, without an exceedingly large sample of data, the null of no predictability will be rejected too frequently despite the fact that a consistent estimator of the long-run covariance matrix is employed. As discussed below, either variety of spurious regression may be relevant in the context of predictive regressions for variance.

2.2.1 Asymptotic Spurious Regression

Phillips (1986) shows that when the Y_t and X_t follow independent random walks with no drift $\hat{\beta}$ converges not to zero but rather to a random variable. In addition, $t_{\hat{\beta}}$ does not converge to the standard normal limiting distribution but instead diverges as the sample size grows.

Time series need not exhibit nonstationary behavior (as in the random walks studied by Phillips (1986)) in order for spurious regression to obtain. Particularly relevant to the variance regression setting are results by Tsay and Chung (2000), who show that nonsense regressions can occur when two series are independent, stationary, fractionally integrated series.[11] Letting d_Y and d_X indicate the fractional order of integration of Y_t and X_t, Tsay and Chung (2000) show that whenever $d_Y + d_X > 0.5$ a regression of Y_t on X_t will be spurious in the sense that $t_{\hat{\beta}}$ diverges asymptotically. If (log) realized variance follows a stationary fractionally integrated variance, and an econometrician runs a predictive regression of *lrvar* on a lagged macroeconomic predictor that is I(1), then the results of Tsay and Chung (2000) show that the resulting t-statistics are divergent.

[11]Marmol (1998) develops theoretical results for spurious regression with fractionally integrated nonstationary processes. Granger, Hyung and Jeon (2001) illustrate the potential of spurious regression in regressions with stationary series including long moving averages and positively autocorrelated autoregressive series.

2.2.2 Virtually-Spurious Regression

Suppose that Y_t and X_t are short-memory processes that satisfy mild assumptions so that the asymptotic results in (5) hold. Finite sample inference may still be plagued by persistence that is difficult to capture or correct for using standard HAC covariance estimators.

Ferson, Sarkissian and Simin (2003a,b; hereafter jointly referenced as "FSS") illustrate the potential of virtually-spurious regression in predictive regressions for stock returns. FSS assume that the unobserved expected return follows a persistent, but stationary and short-memory, AR(1) process. They assume that the forecasting variable also follows a persistent but stationary AR(1) process. When the forecasting variable is independent of returns, so that there is no true predictability, FSS demonstrate via simulation that tests of the null of no predictability are oversized even when HAC estimators are employed and even for sample sizes that are very large in a practical sense.[12]

Table 2 illustrates that log realized variance at the quarterly frequency is well-described by a stationary, short memory ARMA(1,1) process with a correlogram that resembles that of a long-memory ARFIMA(0,d,0) process. Assuming a given macroeconomic forecasting variable follows a persistent, but stationary, AR(1) process, then regressions of log variance on lagged macroeconomic variables are likely to lead to spurious findings of predictability, even though standard dynamic regression assumptions are formally satisfied. Therefore spurious regression of the variety studied by FSS is certainly a concern in variance regressions.

Similarly, and more subtly, we show in the next section of the paper that virtually-spurious regression continues to obtain when mis-specified dynamic models that include one or two lags of the dependent variable are estimated. Indeed, these latter cases are more similar to the setting studied by FSS, where a persistent component in returns is obscured by the low signal-to-noise ratio in returns. Similarly, when one or two lags of variance are included in regressions, the errors may not display 'obvious' signs of correlation, but if variance is truly a long memory process then a persistent component lurks in the error term.

[12]FSS also show that when regressors are chosen via a data mining process the size distortions associated with the spurious regression problem increase substantially, because highly persistent spurious predictors are most likely to be selected as the result of a specification search. Although data mining is not the focus of this paper, data mining would exacerbate the spurious regression problems documented here in exactly the same fashion.

2.3 Weakly Exogenous Regressors and the Stambaugh Bias

The previous section considered inference in a setting where the forecasting variable X_t is independent of Y_t. Consequently, X_t is strictly exogenous and the OLS estimator $\hat{\beta}$ is unbiased.

In the context of models of the conditional *mean* of stock returns, Stambaugh(1999) points out that the strict exogeneity assumption is strongly violated by 'financial ratio' predictors such as *dp* or *ep*, since such variables are are highly contemporaneously correlated with shocks to returns. When the strict exogeneity assumption is violated, $\hat{\beta}$ will be biased in finite samples. As Stambaugh (1999) illustrates, the magnitude of the bias can be large since variables like *dp* are highly persistent while, at the same time, the contemporaneous correlation between shocks to excess returns and *dp* is large in magnitude.

Unfortunately, the strict exogeneity assumption is also dubious in predictive regressions for *variance*. Table 3 presents two measures of correlation between macroeconomic predictors and realized variance. Panel A of Table 3 presents estimates of contemporaneous correlations between *lrvar* and each of the macroeconomic forecasting variables. It is clear that several of the macroeconomic predictors are contemporaneously correlated with stock return variance. Most notably, the contemporaneous correlation between *dfy* and *lrvar* is 0.62 over the full sample, although this value drops to around 0.39 over the post-Treasury Accord sub-sample. Over the post-Treasury Accord period, estimates of the contemporaneous correlation with *lrvar* are in excess of 0.20 (in absolute value) for 8 of the 15 forecasting variables considered.

Panel B of Table 3 presents estimates of correlations between *shocks* to *lrvar* and shocks to the macroeconomic predictors. These estimates are obtained from maximum likelihood estimation of the following bivariate restricted VARMA specification:

$$\begin{aligned} (1-\phi L)\, Y_t &= (1-\theta L)\, \epsilon_t \\ (1-\phi_X L)\, X_t &= \eta_t \end{aligned} \tag{9}$$

where we assume that the error vector is i.i.d. Gaussian:

$$(\epsilon, \eta_t)' = N\left(0, \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}\right) \tag{10}$$

The estimates of ρ vary among the forecasting variables. Some, including those for *ik*, *ltr* and *tbl*, are close to zero, suggesting that the strict exogeneity assumption may not be a severe departure from reality. For other forecasting variables, the strict exogeneity assumption is clearly unsatisfactory. This is true in particular for a number of the 'financial ratio' predictors, including *bm*, *dp* and *ep*, for which the correlation estimates are in the 0.2 - 0.4 range. These conclusions do not depend on the logarithmic transformation applied to realized variance, as similar estimates are obtained using *rvar* (not shown in the table).

When ρ is nonzero, the regressor X_t is only 'weakly exogenous,' in the sense that $E(\varepsilon_t|\{X_t\}_1^T) \neq 0$. As a consequence, the OLS estimator $\hat{\beta}$ will generally be biased in finite samples. Thus, the same bias documented by Stambaugh(1999) is relevant for predictive regression for variance. This bias is likely to be less severe in the context of realized variance regressions, though, since the correlations between *lrvar* and the forecasting variables tend to be significantly smaller than the correlation between, e.g., stock returns and the dividend yield.

2.4 Weakly Exogenous Regressors and Dynamic Misspecification: Inconsistent Estimation

When X_t is only weakly exogenous for ε_t, misspecification of the regression Y_t on X_t in the direction of omitted dynamics can lead to severe consequences. In particular, under the null of no-predictability, the OLS estimator $\hat{\beta}$ is typically inconsistent for the true value β_0.

Domowitz and White (1982) show that, under mild conditions, the estimator $\hat{\beta}$ will be consistent for a "pseudo-true" value β^* that delivers the best approximation to Y_t in a mean square error sense:

$$\hat{\beta} \xrightarrow{p} \beta^* \equiv \text{argmin}_\beta \; E((Y_t - \beta X_{t-1})^2) \tag{11}$$

As a simple example, assume that Y_t and X_t follow the restricted VARMA process described in (9) and (10). Additionally assume that both series have zero mean and unit variance for simplicity. Now suppose that the regression model (4) is estimated using OLS. The regression model is misspecified since the ARMA dynamics of Y_t are omitted.

By (11), the OLS estimate $\hat{\beta}$ converges in probability to the value β^* that minimizes the population mean-square error of the regression. Differentiating the condition in (11) with respect to β and setting the result to zero yields:

$$\begin{aligned}\beta^* &= E(Y_t X_{t-1}) \\ &= \frac{(\phi - \theta)\rho}{(1 - \phi\phi_X)},\end{aligned} \tag{12}$$

where the second expression follows algebraically upon substituting in the $MA(\infty)$ representations for Y_t and X_{t-1} in the previous line. The formula (12) reveals that $\hat{\beta}$ will be inconsistent for the true value of zero unless either: 1.) $\rho = 0$, which is the special case where X_{t-1} is strictly exogenous; or 2.) $\phi = \theta$, when the AR and MA terms for Y_t cancel out so that Y_t becomes a white noise series. The results in Table 2 show that for log realized variance $\phi - \theta > 0$. This implies that the direction of the asymptotic bias in $\hat{\beta}$ is driven by the sign of ρ. In particular, when $\rho > 0$, as appears to be the case for a number of 'financial ratio' variables, then $\hat{\beta}$ will be asymptotically upward biased.

The results discussed above were developed in a simple setting where a scalar Y_t was regressed on a lagged scalar predictor X_{t-1}. Similar insights obtain in more complicated settings, although explicit results such as (12) are more tedious to obtain. In particular, suppose that the regression model in (4) is augmented to include one or two lagged values of Y_t. Since the model remains misspecified, the OLS slope estimator will generally remain inconsistent for the true slope when the forecasting variable is only weakly exogenous. The added lags will tend to reduce the severity of the asymptotic bias, though, since X_t will play a reduced role in correcting for the omitted dynamics.

In the stock return predictability literature a close analog of the dynamic misspecification issue addressed here is studied by Butler, Grullon and Weston (2006; 'BGW'). BGW find evidence of a structural change in the level of excess bond returns in the 1980s that coincides with a shift in managers' bond issuance behavior. When excess bond returns are forecast using managers' issuance activity without including the structural break, the resulting regression model is misspecified. BGW point out that this leads to biased and inconsistent estimation of the slope coefficient for issuance activity in the bond return regression. In both the variance regression setting addressed here and the bond return regression setting studied by BGW, the root source of the inference problem is an omitted explanatory variable in a dynamic regression model.

3 Simulation Evidence

The previous section describes several distinct sources of inference issues in predictive regressions for realized variance. These issues were presented in a very simple, abstract, dynamic linear regression setting. In this section of the paper we describe a series of Monte Carlo experiments that support the relevance of the various inference issues for realistic data generating processes that match key features of realized variance and macroeconomic predictors.

In the predictive regression literature, forecasting variables such as the dividend yield and nominal interest rates are commonly been modeled as persistent AR(1) processes that satisfy the stationarity condition. The univariate time series modeling results discussed earlier suggest that *lrvar* may be modeled using either an ARMA(1,1) or ARFIMA(0,d,0) process. For simplicity, and to emphasize that finite sample inference issues arise even if variance is formally a short-memory process, we assume the ARMA(1,1) specification in the simulations that follow.

3.1 Simulation Design

Data are randomly generated according to the restricted bivariate VARMA process described by (9) and (10), calibrated to empirical estimates of the system for different macroeconomic predictors. Although the simulated Y_t values follow an ARMA(1,1) process tuned to *lravr*, we assume that the econometrician estimates one of the following predictive regression models for Y_t using ordinary least-squares:

$$\begin{array}{lc} \text{Model 1} & Y_t = \alpha + \beta X_{t-1} + \varepsilon_t \\ \text{Model 2} & Y_t = \alpha + \beta X_{t-1} + \phi Y_{t-1} + \varepsilon_t \\ \text{Model 3} & Y_t = \alpha + \beta X_{t-1} + \phi_1 Y_{t-1} + \phi_2 Y_{t-1} + \varepsilon_t \end{array} \tag{13}$$

The regressions in (13) are misspecified, specifically with respect to the dynamic behavior assumed for Y_{t+1}. Model 1 includes no dynamics whatsoever. Models 2 and 3 attempt to handle the apparent persistence in Y_t by including one and two lags of the dependent variable in the regression specification, respectively.

The first set of simulations focus on the potential of spurious regression bias, and therefore we set $\rho = 0$. Otherwise, we calibrate the system using the forecasting variable *ik*, for which the point estimate of ρ is very close to zero. The second set of simulations focus on the case where the forecasting variable is contemporaneously correlated with (log) variance. For this case, we calibrate the system using the forecasting variable *dp*, which exhibits the highest absolute estimate of ρ, namely a value of 0.45.

In all cases, results are based on 10,000 Monte Carlo simulations. We compute results for sample sizes of 100, 250 and 500, which correspond to empirically plausible sample sizes ranging from 25-125 years at the quarterly frequency. Results are reported both for the case where t-statistics are computed using the classical covariance matrix estimator and for the case where they are computed using the nonparametric HAC estimator suggested by Newey and West (1987).

3.2 Simulation Results

Table 4 presents the simulation results. The table reports the bias, standard deviation, skewness and kurtosis of t-statistics along with the proportion of the simulations in which the absolute value of the t-statistic exceeds 1.96, corresponding to rejection of the null at the 5 percent level. Figures 3 and 4 display density plots for the Monte Carlo t-statistics for various sample sizes and dynamic specifications.

Panels A and B of Table 4 present results for the case with $\rho = 0$ for classical and Newey-West standard errors, respectively. For the intercept only model ('CONS'), the t-tests based on classical standard errors in Panel A are dramatically over-sized for all three sample sizes considered, with a standard deviation of nearly three rather than one, and an empirical rejection rate of nearly 50 percent. The extremely poor performance of the t-tests is unsurprising given the strong serial correlation in both the dependent variable and forecasting variable.

More interestingly, the analogous results in Panel B, where the HAC estimator is employed, continue to exhibit strikingly oversized behavior. The performance is poorest for the smallest sample size of 100, where the test rejects nearly 30 percent of the time. Even for the largest sample size of 500, the test rejects nearly 15 percent of the time, illustrating a case of virtually spurious regression.

When one or two lagged dependent variables are included in the specification, the severity of the spurious regression problem diminishes significantly as the moments of the t-statistics are much closer to those of the standard normal. Still, even in the case where two lags of the dependent variable are included in the model, the tests remain significantly oversized, with rejections occurring approximately 9 percent of the time for a sample size of 500 and more often for smaller sample sizes. It is interesting to note that the performance of the Newey-West covariance matrix estimator and classic covariance estimator are nearly identical in this case.

Finally, Figure 3, along with the mean values in Table 4, shows that the t-statistics appear to be unbiased for all of the cases where $\rho = 0$. This is expected since in this case the regressor is strictly exogenous so that OLS parameters estimates are unbiased.

Panels C and D, along with Figure 4, present results for the case where X_t and Y_t are contemporaneously correlated due to the fact that $\rho = 0.45$. As expected based on the discussion in the preceding section, the t-statistics in this case are biased as clearly indicated by the density plots in Figure Y. This is particularly true for the model that omits any dynamics. Note also that in this case the t-statistics are biased upward, as expected based on (12) since $\rho > 0$.

Adding lags of the dependent variable reduces, but does not eliminate, the bias and oversized behavior for the t-statistics. When two lags of the dependent variable are included in the model, the empirical rejection rate is again roughly 8-9 percent, similar to the strictly exogenous case previously examined.

Overall, we conclude that standard methods of robustifying inference to serial correlation and heteroskedasticity in the regression errors do not appear to offer adequate insurance against misspecification of the realized variance dynamics in the direction of ignoring a long memory component. Further, including one or two lags in the regression leads only to modest improvement in the size properties of the tests. Although not reported in Table 4, simulation experiments using a correct ARMA(1,1) dynamic specification for Y_t show that, as expected, correct model specification leads to well-behaved hypothesis tests.

4 Empirical Regression Results

4.1 Assessomg the Ability of Macroeconomic Variables to Predict Volatility

Mindful of the potential of spurious regression and bias driven by omitted dynamics, we turn to an analysis of the empirical evidence regarding the ability of macroeconomic variables to forecast volatility at a quarterly horizon. Table 5 presents estimation results for each macroeconomic predictor in turn. For comparative purposes, a range of dynamic specifications are estimated, including a model with no dynamics, models with AR(1) and AR(2) dynamics, a model with ARMA(1,1) dynamics and a model with ARFIMA(0,d,0) dynamics. All of these may be viewed as special cases of the following ARFIMAX model:

$$(1-\phi L)(1-L)^d lrvar_t = c + X_t\beta + (1-\theta L)\varepsilon_t, \tag{14}$$

Based on the results and discussion in the preceding sections of the paper, regression results under the richer ARMA(1,1) and ARFIMA(0,d,0) specifications are likely to be substantially more reliable indications of the true forecasting ability of macroeconomic variables for *lrvar*.

For each specification, Table 5 reports estimates of the slope coefficient $\hat{\beta}$, the t-statistic for a two-sided test of the null that $\beta = 0$ based on Newey-West standard errors, and the regression R^2. Results are reported over both the full sample period, which varies over the different macroeconomic predictors as indicated in Table 1, and also over the post-Treasury accord period of 1951:2 - 2005:4.

4.1.1 Full Sample Results

Over the full sample period (Panel A of Table 5), there is only limited evidence regarding the ability of macroeconomic predictors to forecast volatility. Specifically, for the preferred ARMA(1,1) and ARFIMA(0,d,0) specifications, the slope coefficients on lagged macroeconomic predictors are insignificant at conventional levels for all variables save the investment to capital ratio *ik*. The predictive power of *ik* is robust across the various dynamic specifications, whereas the apparent predictive ability of the default yield *dfy* evaporates when richer dynamic specifications are entertained. It is notable that the coefficient on *dfy* changes substantially moving from a constant-only specification to an ARFIMA(0,d,0) specification. This also occurs for several other forecasting variables and in some cases the point estimate of the slope changes sign. Finally, it is interesting that the long-memory specification appears to strengthen the evidence for predictability for the *tbl*; although this variable is still not significant at the 5 percent level.

4.1.2 Post-Treasury Accord Results

The macroeconomic variables fare better over the post-Treasury Accord period. Roughly half of the 15 forecasting variables are statistically significant based on a regression with ARMA(1,1) dynamics. Statistically significant variables include *de*, *dfr*, *dy*, *ik*, *lty* and *tbl*. Several other variables are borderline significant. While the evidence for predictability is stronger, it is important to note that several other variables would appear to be significant based on regression models that include no dynamics, or only a single lag of *lrvar*. Also, there are still large differences in the estimated coefficient values across the various dynamic specifications. So, while there is some evidence that macroeconomic variables predict (log) variance at a quarterly horizon during the post-Treasury Accord period, the nature of inference remains sensitive to the dynamic specification employed in the regression.

4.1.3 Economic Significance

It is clear from a cursory examination of Table 5 that the magnitude of the regression coefficients on the macroeconomic predictors varies substantially over the different dynamic specifications considered. Often, the magnitude of the coefficients drops in moving from an intercept-only or AR(1) specification for realized (log) variance to the richer ARFIMA(0,d,0) specification. A particularly striking example is the default yield variable, for which the absolute value of the coefficient plummets steadily as richer dynamic specifications are entertained for the dependent variable.

While it would be overly tedious to provide an explicit interpretation and discussion of economic significance for each forecasting variable considered, several examples are worth mention. The coefficients in regressions of *lrvar* on (lagged) *dp*, *de*, *ep*, *dy* and *bm* are particularly easy to interpret, as these are simply estimates of the elasticity of variance forecasts with respect to, e.g., the lagged dividend-price ratio. When the model includes no dynamics these estimated elasticity values vary widely in sign and magnitude across the different 'ratio' variables. For example, the estimated elasticity of variance forecasts with respect to the dividend yield is 0.11, implying that a 10 percent increase in the yield results in a 1.1 percent increase in volatility. The estimated elasticity for the dividend-earnings ratio, on the other hand, is 0.73, suggesting a rather large 7.3 percent increase in the forecasted variance for a 10 percent increase in this ratio. Under a richer AR(2) or ARMA(1,1) specification, on the other hand, the estimated elasticity values are typically below 0.1 in absolute value, suggesting a change of less than one percent in forecasted variance for a 10 percent change in the ratio.

4.2 Multivariate Regressions

The analysis up to this point has focused on the ability of individual macroeconomic variables to explain time variation in stock return volatility. To obtain a better sense of the overall predictive strength of macroeconomic variables for stock return variance we run multivariate predictive regressions over the sample period 1927:1 - 2005:4 and 1952:1 - 2005:4. The latter period essentially corresponds to the post-Treasury Accord period, adjusted by two quarters to sync up with the onset of *cay* data. The former period corresponds to the CRSP coverage period.

In each regression, we include most of the available macroeconomic predictors. The predictors *dp*, *ep*, *bm* and *dy* are highly correlated, and so we include only *dp* from this set (arbitrarily). Similarly, *lty* and *tbl* are highly correlated and we include only *tbl*. Table 6 displays the estimated coefficients and standard errors, as well as information criteria, adjusted R^2 and Box Pierce (Q) diagnostic statistics for residual correlation.

We also include results for an ARMA(1,1) model that excludes the macroeconomic predictors as a reference. One way to assess the overall strength of the predictive power of the macroeconomic variables is via a comparison of the improvement in fit offered by adding the macroeconomic variables to the basic ARMA(1,1) model. Only a handful of the slope coefficients on macroeconomic variables are statistically significant at conventional level in either sub-period. The variable *dfr* is the only statistically significant variable over both periods examined. The information criteria and adjusted R^2 values indicate that adding the full set of macroeconomic predictors leads to only a modest improvement in model fit (an adjusted R^2 improvement of roughly 2 percent). The information criteria are split regarding whether they favor the more heavily parameterized model that includes lagged macroeconomic data (favored by AIC) versus the simpler univariate model (favored by BIC). Time series plots of fitted variances based on the two models over the period 1927:1 - 2005:4 (not explicitly shown) are very similar, again suggesting that the economic significance of predictability afforded by the macro variables is limited.

5 Robustness and Extensions

5.1 Alternative Variance Proxies

The realized variance measure *rvar* implicitly assumes that the mean excess return is constant and equal to zero. Both assumptions are likely invalid. To explore the consequence of relaxing these assumptions, we computed a number of alternative realized variance measures with different assumptions regarding the unconditional and conditional mean of excess returns. Empirical evidence suggests that daily excess returns are positively autocorrelated. In this case the sum of squared daily returns is a downward-biased estimator for the true integrated variance. Following French, Schwert and Stambaugh (1987) we construct an alternative measure of realized variance as:

$$\sum_{i=1}^{N_t} R_{i,t}^2 \left[1 + \frac{2}{N_t} \sum_{j=1}^{N_t - 1} (N_t - j)\, \hat{\rho}^j \right]. \tag{15}$$

As in French, Schwert and Stambaugh (1987), the within-month returns are assumed to follow an AR(1) process which gives rise to the bracketed bias correction term in (15). The parameter $\hat{\rho}$ is an estimate of the autocorrelation in excess daily returns. Other alternative estimators were constructed by first estimating a predictive regression for stock returns to obtain a fitted series of time-varying expected returns. These were subsequently converted to the daily frequency and subtracted from realized daily returns prior to computing the realized variance. Variations of this sort had very little effect on our results as the realized variance series were very similar. This accords with the notion that expected return variation is an order of magnitude smaller than unexpected return variation.

5.2 Modeling the Level of Variance

We have focused on models in which the dependent variable is the logarithm of realized variance. As a robustness check on our findings, we repeated the empirical analysis using the level of realized variance, *rvar*, as the dependent variable. To conserve space, these results are not explicitly reported in the paper, but are available from the author upon request. Qualitatively, the findings are similar, in the sense that point estimates and inference results regarding the predictive power of macroeconomic variables are highly sensitive to the dynamic specification employed, with richer dynamic models corresponding to weaker evidence of predictability in both a statistical and economic sense.

6 Discussion

The paper examines linear regressions of 'realized volatility' measures on macroeconomic variables at the monthly horizon. The paper notes the potential of finite sample inference problems in regressions of realized volatility measures on persistent macroeconomic predictors such as the dividend yield, nominal interest rates, default premium and term spread. If realized volatility follows a simple long memory process then regressions of realized volatility on persistent macroeconomic predictors may be subject to a spurious regression bias similar to that illustrated by Ferson, Sarkissian and Simin (2003a,b) when the dynamic model for realized variance is misspecified in the direction of omitting the long memory component. Our simulation results illustrate that the practice of including one or two lags of the dependent variable in the realized volatility model may not be sufficient to control spurious regression bias when realized volatility exhibits long range dependence. This is true even if common kernel-based HAC covariance estimators are employed.

Some of the macroeconomic predictors we consider are contemporaneously correlated with stock return variance. In this the forecasting variable is only weakly exogenous in the regression model, which raises the spectre of several additional inferences problems. In particular, slope coefficient estimates will be biased in finite samples as shown by Stambaugh (1999), and mispecification of the dynamic structure in volatility may lead to inonsistent estimation of the true slope. A simulation analysis suggests that the inference issues raised in this paper are of concern in a realistic setting.

Turning to the empirical evidence of the forecasting ability of macroeconomic variables for stock return variance, we find that the case for predictive ability for many of these variables is far from compelling. Evidence of predictability is strongest in the latter portion of our sample; however, even in these cases the economic signficance of the macroeconomic factors for volatility seems to be quite weak.

In this paper we have operated under the assumption that volatility is in fact observed. This is clearly not strictly valid. Without question some measurement error remains when daily squared returns are used to construct volatility proxies at monthly and quarterly horizons. As a consequence, the results reported in this paper must be interpreted with some caution. Our inability to observe volatility without error presumably translates into reduced power to detect true, valid forecasting relationships as well as underestimation of the economic significance of such relationships. High-frequency intraday returns data provide a means to obtain more accurate volatility proxies. However, such data are only available since the mid 1980s for the S&P 500. Given the high persistence in macroeconomic quantities such as the interest rate and dividend yield, the effective sample size in regressions based on such data would be very small. For example, the dividend yield largely trends downward during the post-1985 period, so it is unclear to what extent one might learn about the ability of the dividend yield to forecast volatility in such a regression. Nevertheless, this represents a possible area for future work, particularly as the available time series of high frequency intraday returns gets longer.

In recent research, Han and Park (2005a,b) provide a theoretical foundation for the ability of macroeconomic variables to generate the type of long memory behavior typically observed in volatility. This is obviously intriguing as it suggests the possibility of a strong connection between the dynamics in volatility and observable macroeconomic quantities. In a way, it is somewhat disappointing that when we simultaneously include a long memory component and lagged macroeconomic quantities in our volatility regressions that the macroeconomic variables only 'explain away' a relatively small fraction of the long memory behavior. However, two things are important to keep in mind. First, we consider only linear functions of macroeconomic predictors. Nonlinear specifications may increase the explanatory ability of macroeconomic predictors in our regressions and this is an interesting area for future research. Perhaps more importantly, we consider only a relatively small subset of the vast quantity of macroeconomic information that is publicly available. It is possible that by including more macroeconomic information we might ultimately be able to further reduce the residual degree of long memory behavior. Obviously, data mining becomes a concern in this context. One idea might be to construct indices or factors that are linear combinations of a large number of macroeconomic quantities.

A host of intriguing questions remain for future research. The present paper focuses entirely on a full sample or in-sample analysis of predictability. There is relatively little existing work that addresses the out-of-sample forecasting ability of models linking stock return volatility to macroeconomic conditions.[13] There has been much recent activity in the econometrics literature related to the out-of-sample evaluation of forecasting models. A thorough out-of-sample examination of the effectiveness of incorporating macroeconomic variables in longer horizon volatility forecasting models would be of great interest.[14]

Finally, while this paper documents evidence of long range dependence in stock return volatility at the monthly horizon, questions remain regarding the source of this behavior and in particular whether structural instability might be the root cause. Diebold and Inoue (2001) illustrate that regime switching can give rise to long memory behavior. Banergee and Urga (2004) review the literature on breaks, long memory and the connection between the two. Empirical evidence suggests the possibility of breaks in both volatility and macroeconomic variables such as interest rates. It would be interesting to consider inference problems that might arise from ignoring such breaks in the type of regressions considered in this paper.

[13] Some limited results presented by Marquering and Verbeek (2005) are not particularly encouraging.

[14] Inoue and Killian (2004), however, argue that in-sample tests of the variety discussed in this paper may be preferable to out-of-sample tests of the null of no predictability.

References

Andersen, T.G., T. Bollerslev, F.X. Diebold and H.. Ebens, 2001, The distribution of realized stock return volatility, *Journal of Financial Economics*, 61, 43-76.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys, 2003, Modeling and forecasting realized volatility, *Econometrica*, 71, 529-626.

Andrews, D.W.K. and J.C. Monahan, 1992, An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator, *Econometrica*, 60, 953-966.

Ang, A. and G. Bekaert, 2004, Stock return predictability: Is it there? Manuscript, Columbia University, USC and NBER.

Baillie, R.T., T. Bollerslev and H.O. Mikkelsen, 1996, Fractionally Integrated Generalized Autoregressive Conditional Heteroskedastiticy, *Journal of Econometrics*, 74, 3-30.

Banergee, A. and G. Urga, 2005, Modelling structural breaks, long memory and stock return volatility, *Journal of Econometrics*, Forthcoming.

Barndorff-Nielsen, O.E. and N. Shephard, 2002, Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society*, Series B, 64, 253-280.

Beltratti, A. and C. Morana, 2002a, Structural change and long range dependence in volatility of exchange rates: either, neither, or both? Manuscript, Bocconi University.

Beltratti, A. and C. Morana, 2002b, Breaks and persistency: macroeconomic causes of stock market volatility, Manuscript, Bocconi University.

Bossaerts, P. and P. Hillion, 1999, Implementing statistical criteria to select return forecasting models: What do we learn? *Review of Financial Studies* 12, 405-428.

Brandt, M.W. and C. Jones, 2002, Volatility forecasting with range-based EGARCH models, Manuscript, The Wharton School, University of Pennsylvania.

Breen, W., L. Glosten and R. Jagannathan, 1989, Economic significance of predictable variations in excess stock returns, *Journal of Finance*, 44, 1177-1189.

Butler, A., G. Grullon and J.P. Weston, 2006, Can managers successfully time the maturity of their debt issues?, *Journal of Finance*, 61, 1731 - 1758.

Campbell, J.Y., 1987, Stock returns and the term structure. *Journal of Financial Economics*, 18, 373-399.

Campbell, J.Y., 2003, Consumption-based asset pricing. Chapter 13 in George Constantinides, Milton Harris and Rene Stulz, eds. *Handbook of the Economics of Finance* Vol. 1B. North-Holland, Amsterdam, 803-887.

Campbell, J.Y. and M. Yogo, 2004, Efficient tests of stock return predictability. Working paper, Harvard University.

Christoffersen, P. and F.X. Diebold, 2000, How relevant is volatility forecasting for financial risk management? *Review of Economics and Statistics*, 82, 12-22.

Davidson, J., 2005. Time Series Modelling Version 4.11 (University of Exeter).

Den Haan, W.J. and A. Levin, 1997, A practitioner's guide to robust covariance matrix estimation, in G.S. Maddala and C. R. Rao, ed.: *Handbook of Statistics*, 15, Chapter 12, 291-341.

Den Haan, W.J. and A. Levin, 2000, Robust covariance matrix estimation with data-dependent VAR prewhitening order, Technical Working Paper 255, NBER.

Diebold, F.X. and A. Inoue, 2001, Long memory and regime switching, Journal of Econometrics, 105, 131-159.

Ding, Z., C.W.J. Granger and R.F. Engle, 1993, A long memory property of stock market returns and a new model, *Journal of Empirical Finance*, 1, 83-106.

Domowitz, I. and H. White, 1982, Misspecified models with dependent observations, *Journal of Econometrics*, 20, 35-50.

Doornik, J.A., 2002, Object-oriented Matrix Programming Using Ox (3rd ed. London: Timberlake Consultants and Oxford: www.huff.ox.ac.uk/Users/Doornik).

Doukhan, P. G. Oppenheim and M.S. Taqqu, 2003, Theory and Applications of Long-Range Dependence (Birkhauser Press).

Ferson, W., S. Sarkissian and T.T. Simin, 2003a, Spurious regressions in financial economics? *Journal of Finance*, 58, 1393-1414.

Ferson, W., S. Sarkissian and T.T. Simin, 2003b, Is stock return predictability spurious? *Journal of Investment Management*, 3, 10-19.

Foster, F.D., T. Smith and R.E. Whaley, 1997, Assessing goodness-of-fit of asset pricing models: The distribution of the maximal R-squared, *Journal of Finance*, 52, 591-607.

French, K.R., G.W. Schwert and R.F. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics*, 19, 3-29.

Geweke, J. and S. Porter-Hudak, 1983, The estimation and application of long-memory time series models, *Journal of Time Series Analysis*, 4, 221-238.

Glosten, L., R. Jagannathan and D. Runkle, 1993, On the relation between the expected value and the volatility of the nominal excess retrun on stocks, *Journal of Finance*, 48, 1779-1801.

Goyal, Amit, and I. Welch, 2007, A comprehensive look at the empirical performance of the equity premium prediction. *Review of Financial Studies*, forthcoming.

Granger, C.W.J., 1980, Long memory relationships and the aggregation of dynamic models, *Journal of Econometrics*, 14, 227-238.

Granger, C.W.J., N. Hyung and Y. Jeon, 2001, Spurious regression with stationary series, *Applied Economics*, 2001, 33, 899-904.

Granger, C.W.J. and R. Joyeux, 1980, An introduction to long-memory time series models and fractional differencing, *Journal of Time Series Analysis*, 1, 15-29.

Granger, C.W.J. and P. Newbold, 1974, Spurious regressions in Economics, *Journal of Econometrics*, 4, 111-120.

Guidolin, M. and A. Timmermann, (2004), Term structure of risk under alternative specifications, *Journal of Econometrics*, Forthcoming.

Han, H. and J.Y. Park, 2005a, Time series properties of ARCH processes with persistent covariates, Working Paper, Rice University.

Han, H. and J.Y. Park, 2005b, ARCH models with nonstationary covariates, Working Paper, Rice University.

Harvey, C.R., 2001, The specification of conditional expectations, *Journal of Empirical Finance*, 8, 573, 637.

Hodrick, R.J., 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies*, 5, 357-386.

Hosking, J.R.M., 1981, Fractional differencing, *Biometrika*, 68, 165-176.

Inoue, A. and L. Kilian, 2004, In-sample or out-of-sample tests of predictability: Which one should we use? *Econometric Reviews*, 23, 371-402.

Kirby, Chris, 1997, Measuring the predictable variation in stock and bond returns, *The Review of Financial Studies*, 10, 579-630.

Lettau, M. and S. Ludvigson, 2001, Consumption, aggregate wealth and expected stock returns, *Journal of Finance*, 56, 815-849.

Lettau, M. and S. Ludvigson, 2005, Measuring and modeling variation in the risk-return trade-off, Manuscript, New York University and NBER, Forthcoming in Y. Ait-Sahalia and L.P. Hansen, ed.: Handbook of Financial Econometrics (Elsevier).

Lo, A. and A.C. MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies*, 3, 431-468.

Marquering, W. and M. Verbeek, 2005, The economic value of predicting stock index returns and volatility, *Journal of Financial and Quantitative Analysis*, Forthcoming.

Marmol, F., 1998, Spurious regression theory with nonstationary fractionally integrated processes, *Journal of Econometrics*, 84, 233-250.

Mele, A., 2005a, Asymmetric stock-market volatility and the cyclical behavior of expected returns, Working paper, London School of Economics.

Mele, A., 2005b, Understanding stock market volatility, London School of Economics Financial Market Groups Review, 67, 10-15.

Newey, W.K. and K.D. West, 1987, A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55, 703-708.

Paye, B.S. and A. Timmermann, 2006, Instability of Return Prediction Models, *Journal of Empirical Finance*, Forthcoming.

Phillips, P.C.B., 1986, Understanding spurious regressions in Econometrics, *Journal of Econometrics*, 33, 311-340.

Phillips, P.C.B., 1988, Regression theory for near-integrated time series. *Econometrica*, 56, 1021-1044.

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

Schwert, G.W., 1989b, Tests for unit roots: A Monte Carlo investigation, *Journal of Business and Economic Statistics*, 7, 147-159.

Shanken, J.A., 1990, Intertemporal asset pricing: an empirical investigation, *Journal of Econometrics*, 45, 99-120.

Stambaugh, R.F., 1999, Predictive regressions, *Journal of Financial Economics*, 54, 375-421.

Sullivan, R., A. Timmermann, and H. White, 1999, Data-Snooping, Technical Trading Rule Performance, and the Bootstrap, *Journal of Finance*, 54, 1647-1691.

Taylor, S.J., 1986, Modeling Financial Time Series (John Wiley and Sons, Ltd.).

Tsay, W.J. and C.F. Chung, 2000, The spurious regression of fractionally integrated processes, *Journal of Econometrics*, 96, 155-182.

Valkanov, R., 2003, Long horizon regressions: theoretical results and applications, *Journal of Financial Economics,* 68, 201-232.

West, K.D. and D. Cho, 1995, The predictive ability of several models of exchange rate volatility, *Journal of Econometrics*, 69, 367-391.

Whitelaw, R.F., 1994, Time variations and covariations in the expectation and volatility of stock market returns, *Journal of Finance*, 49, 515-541.

Yule, G.U., 1926, Why do we sometimes get nonsense correlations between time series? A study in sampling and the nature of time series, *Journal of the Royal Statistical Society*, 89, 1-64.

Table 1: Descriptive Statistics

The table presents descriptive statistics for the realized volatility measures and macroeconomic forecasting variables. Data for excess returns and macroeconomic forecasting variables are based on Goyal and Welch (2008) and are sourced from Amit Goyal's webpage. Data on realized volatility measures are constructed using daily total return data on the S&P 500 from CRSP from 1926 - 2005 and daily total return data provided by Schwert (1990) prior to 1926.

		Sample Period	mean	sd	skew	kurt.	ρ_1	ρ_2	ρ_3
rvar	Realized Variance	1885q1 - 2005q4	0.0073	0.0119	4.85	30.61	0.6531	0.6195	0.5517
lrvar	Log Realized Variance	1885q1 - 2005q4	-5.4208	0.8778	0.83	4.30	0.6934	0.6094	0.5479
exret	Excess Return	1885q1 - 2005q4	0.0515	0.1033	2.40	26.13	-0.0364	0.0002	0.187
bm	Book to Market	1921q1 - 2005q4	0.6049	0.2586	0.76	5.33	0.9311	0.8929	0.8455
cay	Consumption-Wealth-Income	1951q4 - 2005q4	0.0000	0.0121	0.00	2.68	0.8412	0.7258	0.6401
de	Dividend Payout Ratio	1885q1 - 2005q4	-0.5411	0.2728	0.51	3.79	0.9607	0.8897	0.8049
dfr	Default Return Spread	1926q1 - 2005q4	0.0010	0.0169	-0.89	6.34	-0.0981	-0.0664	-0.0058
dfy	Default Yield Spread	1919q1 - 2005q4	0.0120	0.0073	2.07	9.56	0.9143	0.8887	0.8314
dp	Dividend Price Ratio	1885q1 - 2005q4	-3.1950	0.4043	-0.84	4.32	0.963	0.9252	0.8854
dy	Dividend Yield	1885q1 - 2005q4	-3.1833	0.3968	-0.95	4.46	0.9633	0.9271	0.8923
ep	Earnings Price Ratio	1885q1 - 2005q4	-2.6539	0.3628	0.05	3.11	0.9493	0.8877	0.8239
i/k	Investment-Capital Ratio	1947q1 - 2005q4	0.0358	0.0035	0.41	2.42	0.964	0.8984	0.8177
infl	Inflation Rate	1913q2 - 2005q4	0.0082	0.0163	0.49	7.31	0.5647	0.4372	0.3658
ltr	Long Term Bond Return	1926q1 - 2005q4	0.0143	0.0431	1.06	7.80	-0.051	0.0296	0.1124
lty	Long Term Bond Yield	1919q1 - 2005q4	0.0528	0.0278	1.05	3.62	0.9873	0.9777	0.9674
ntis	Net Equity Expansion	1926q3 - 2005q4	0.0213	0.0233	2.01	11.80	0.9061	0.7984	0.6562
tbl	Treasury-Bill Rate	1920q1 - 2005q4	0.0378	0.0300	1.05	4.53	0.9585	0.9269	0.9141
tms	Term Spread	1920q1 - 2005q4	0.0151	0.0127	-0.20	3.34	0.8431	0.7157	0.662

Table 2: Univariate Time Series Model Section

The table presents estimation results for a univariate time series analysis of the log of realized variance (*lrvar*). Five models are estimated, including an intercept-only model, an AR(1) model, an AR(2) model, an ARMA(1,1) model and an ARFIMA(0,d,0) model. Panel A displays results for the full sample period 1885q1 - 2005q4, while Panel B presents results for the post-Treasury Accord period from 1952q2 - 2005q4. Standard errors are displayed in parentheses below coefficient estimates.

Panel A: Full Sample (1885q1:2005q4)

	INTERCEPT	AR1	AR2	ARMA(1,1)	ARFIMA(0,d,0)
AR		0.69	0.52	0.90	
		(0.04)	(0.05)	(0.03)	
AR2			0.25		
			(0.05)		
MA				0.45	
				(0.07)	
d					0.51
					(0.04)
AIC	-622.91	-465.03	-449.93	-444.98	-446.26
SIC	-625.00	-469.21	-456.19	-451.25	-450.44
Adj-R2	0.00	0.48	0.51	0.52	0.52
SIG-RES	0.88	0.63	0.61	0.61	0.61
Box-Pierce(12)	1204.86***	47.45***	8.70	6.74	7.21

Panel B: Post Treasury Accord (1951q2:2005q4)

	INTERCEPT	AR1	AR2	ARMA(1,1)	ARFIMA(0,d,0)
AR		0.65	0.52	0.86	
		(0.05)	(0.06)	(0.06)	
AR2			0.21		
			(0.07)		
MA				0.39	
				(0.10)	
d					0.50
					(0.05)
AIC	-259.85	-200.65	-196.87	-195.29	-195.74
SIC	-261.55	-204.04	-201.95	-200.37	-199.13
Adj-R2	0.00	0.42	0.44	0.45	0.45
SIG-RES	0.79	0.60	0.59	0.59	0.59
Box-Pierce(12)	357.96***	17.46	9.45	6.53	6.38

Table 3: Correlations

The table presents evidence of correlations between the log of realized variance (*lrvar*) and various macroeconomic forcasting variables. Panel A presents estimates of the contemporaneous correlation between *lrvar* and each macroeconomic predictor over the full sample of available data (see Table 1) and over the post-Treasury Accord period. Panel B presents the estimated correlation between shocks to *lrvar* and shocks to the corresponding macroeconomic predictors based on maximum likelihood estimation of the restricted VARMA(1,1) system described in equations (9) and (10) of the paper.

Panel A: Contemporaneous Correlations

	1885q1:2005q4	1951q2:2005q4
	lrvar	*lrvar*
bm	0.199	-0.060
cay	-0.057	-0.057
de	0.244	-0.280
dfr	-0.008	-0.075
dfy	0.617	0.392
dp	0.140	-0.203
dy	0.074	-0.254
ep	-0.027	-0.076
ik	0.198	0.200
infl	-0.087	0.173
ltr	0.077	0.153
lty	-0.067	0.302
ntis	0.014	-0.204
tbl	-0.143	0.240
tms	0.191	0.092

Panel B: Residual Correlation - Restricted VARMA(1,1)

	1885q1:2005q4	1951q2:2005q4
	lrvar	*lrvar*
bm	0.21	0.315
cay	0.228	0.228
de	0.043	-0.010
dfr	-0.140	-0.241
dfy	0.181	0.172
dp	0.373	0.452
dy	0.005	0.060
ep	0.304	0.379
ik	0.060	0.004
infl	0.046	0.024
ltr	0.041	0.104
lty	-0.025	-0.053
ntis	0.158	0.051
tbl	-0.017	0.040
tms	-0.008	-0.082

Table 4: Simulation Evidence for t-statistics

The table presents results for Monte Carlo simulation experiments based on random draws from the restricted VARMA(1,1) system described in (9) and (10) of the paper calibrated to alternative macroeconomic predictors. For each set of simulated data, regressions of Y on lagged X are run with alternative dynamic specifications. The table presents the first four moments for the *t*-statistics associated with the slope coefficient on lagged X along with the percentage of simulations in which the null hypothesis of no predictability was rejected at the 5% significance level. In Panels A and B the VARMA(1,1) system is calibrated using the log realized variance (as Y) and the investment to capital ratio ik (as X). In Panels C and D the system is calibrated using the log realized variance (as Y) and the dividend-price ratio (as X). The t-statistics in Panels A and C are computed using classic standard errors while those in panels B and D are computed using the HAC estimator of Newey and West (1987).

Panel A: Strictly Exogenous Regressor tuned to ik; classic errors									
	N = 100			N = 250			N = 500		
	CONS	*AR1*	*AR2*	*CONS*	*AR1*	*AR2*	*CONS*	*AR1*	*AR2*
mean	-0.013	-0.002	0.005	0.008	0.004	0.010	0.033	0.013	0.013
sd	2.749	1.430	1.245	2.850	1.391	1.167	2.867	1.371	1.129
skew	-0.025	-0.016	0.002	0.014	-0.007	-0.021	0.017	0.013	0.023
kurtosis	3.207	2.952	2.936	3.274	3.116	3.019	3.055	2.993	2.947
% |t| > 1.96	46.26%	17.07%	11.60%	48.30%	15.55%	9.16%	49.42%	15.25%	8.35%

Panel B: Strictly Exogenous Regressor tuned to ik; NW errors									
	N = 100			N = 250			N = 500		
	CONS	*AR1*	*AR2*	*CONS*	*AR1*	*AR2*	*CONS*	*AR1*	*AR2*
mean	-0.011	-0.006	0.006	0.002	0.005	0.012	0.014	0.012	0.015
sd	1.940	1.545	1.345	1.568	1.406	1.223	1.362	1.285	1.170
skew	-0.008	-0.027	0.017	-0.006	-0.001	-0.009	0.013	0.021	0.038
kurtosis	3.919	3.146	3.193	3.750	3.195	3.038	3.387	3.069	2.976
% |t| > 1.96	28.21%	20.20%	14.06%	19.29%	15.71%	10.45%	14.63%	12.78%	9.54%

Table 4: Simulation Evidence for t-statistics (Cont.)

Panel C: Predetermined Regressor tuned to dp; classic errors									
	N = 100			N = 250			N = 500		
	CONS	*AR1*	*AR2*	*CONS*	*AR1*	*AR2*	*CONS*	*AR1*	*AR2*
mean	1.951	-0.077	-0.308	3.009	0.145	-0.195	4.221	0.405	-0.054
sd	2.879	1.376	1.230	3.019	1.356	1.155	3.101	1.344	1.123
skew	0.081	0.047	0.055	0.052	0.042	0.053	0.075	0.035	0.022
kurtosis	3.349	3.092	3.014	3.063	3.038	2.987	3.141	3.105	3.083
% |t| > 1.96	57.26%	15.22%	12.35%	68.26%	14.98%	9.22%	78.44%	15.50%	8.57%

Panel D: Predetermined Regressor tuned to dp; NW errors									
	N = 100			N = 250			N = 500		
	CONS	*AR1*	*AR2*	*CONS*	*AR1*	*AR2*	*CONS*	*AR1*	*AR2*
mean	*1.299*	*-0.098*	*-0.336*	*1.538*	*0.133*	*-0.207*	*1.854*	*0.378*	*-0.058*
sd	1.973	1.490	1.329	1.610	1.392	1.209	1.418	1.289	1.163
skew	0.361	0.012	0.008	0.307	-0.006	0.037	0.324	0.002	0.001
kurtosis	4.242	3.163	3.298	3.694	3.110	3.058	3.698	3.219	3.115
% |t| > 1.96	38.22%	18.43%	15.05%	38.27%	16.06%	11.26%	45.89%	13.81%	9.44%

Table 5: Predictive Regressions for Variance

The Table presents estimation results for predictive regressions of the log of realized variance for the S&P500 Index on lagged macroeconomic forecasting variables. Five different dynamic specifications are entertained, including a model with no dynamics, and models with AR(1), AR(2), ARMA(1,1) and ARFIMA(0,d,0) dynamics. The table reports the OLS estimate of the slope coefficient on the corresponding lagged macroeconomic variable, the t-stat based on Newey-West (1987) standard errors, and the R^2 -value for the regression.

Panel A: Full Sample (first observation period - 2005:4)

	INTERCEPT			AR1			AR2			ARMA(1,1)			ARFIMA(0,d,0)		
	β	*t*-tstat	R^2	β	*t*-tstat	R^2	β	*t*-tstat	R^2	β	*t*-tstat	R2	β	*t*-tstat	R^2
bm	0.578	0.869	0.025	0.038	0.226	0.550	-0.008	-0.053	0.592	-0.045	-0.618	0.601	-0.102	-0.815	0.575
cay	-8.844	-1.154	0.018	-6.336	-2.081	0.432	-5.551	-1.945	0.455	-3.486	-1.840	0.466	-4.254	-1.441	0.419
de	0.730	1.340	0.051	0.197	1.416	0.485	0.111	1.001	0.514	0.045	0.691	0.524	-0.195	-1.299	0.501
dfr	-0.304	-0.093	0.000	0.051	0.024	0.560	-2.053	-1.077	0.606	-2.720	-1.338	0.612	-1.341	-0.604	0.581
dfy	74.572	6.395	0.335	25.569	3.779	0.570	16.249	2.684	0.598	6.935	1.339	0.598	-5.311	-0.809	0.574
dp	0.196	0.562	0.008	-0.014	-0.186	0.481	-0.029	-0.415	0.514	-0.014	-0.372	0.524	-0.126	-1.720	0.504
dy	0.119	0.355	0.003	0.007	0.100	0.481	-0.033	-0.509	0.514	-0.018	-0.508	0.524	-0.114	-1.675	0.504
ep	-0.166	-0.835	0.005	-0.120	-1.708	0.484	-0.091	-1.288	0.515	-0.040	-1.071	0.524	-0.049	-0.723	0.504
ik	56.851	2.191	0.067	30.076	3.044	0.430	31.118	3.191	0.453	18.204	3.416	0.463	37.155	3.336	0.436
infl	-5.947	-0.787	0.011	-2.277	-1.239	0.531	-1.163	-0.664	0.574	-0.886	-0.657	0.584	-0.696	-0.444	0.548
ltr	0.443	0.418	0.000	-0.861	-1.058	0.562	-0.925	-1.159	0.606	-1.096	-1.450	0.612	-0.993	-1.206	0.582
lty	-2.006	-0.553	0.003	-0.311	-0.336	0.546	-0.026	-0.031	0.589	-0.024	-0.040	0.595	1.939	1.785	0.578
ntis	0.605	0.103	0.000	0.195	0.110	0.562	1.093	0.838	0.605	1.124	1.341	0.614	-1.610	-1.138	0.588
tbl	-3.720	-0.957	0.014	-0.387	-0.380	0.546	0.318	0.447	0.590	0.361	0.675	0.599	1.966	1.805	0.569
tms	11.251	1.233	0.023	0.797	0.318	0.546	-1.825	-0.698	0.590	-2.198	-1.480	0.601	-2.394	-0.855	0.569

Panel B: Post-Treasury Accord (1951q2:2005q4)

	INTERCEPT			AR1			AR2			ARMA(1,1)			ARFIMA(0,d,0)		
	β	*t*-tstat	R^2	β	*t*-tstat	R2	β	*t*-tstat	R^2	β	*t*-tstat	R2	β	*t*-tstat	R^2
bm	-0.287	-0.555	0.008	-0.159	-0.966	0.425	-0.128	-0.833	0.449	-0.075	-0.848	0.457	-0.085	-0.561	0.445
cay	-8.844	-1.154	0.018	-6.336	-2.081	0.432	-5.551	-1.945	0.455	-3.486	-1.840	0.466	-4.254	-1.441	0.419
de	-1.312	-2.584	0.095	-0.559	-2.360	0.439	-0.496	-2.420	0.460	-0.308	-2.028	0.467	-0.417	-1.910	0.456
dfr	-4.584	-1.776	0.010	-2.293	-1.162	0.425	-4.085	-2.032	0.455	-4.786	-2.295	0.466	-3.638	-1.970	0.449
dfy	62.183	3.639	0.109	16.774	1.932	0.430	9.181	0.972	0.449	3.502	0.519	0.456	2.598	0.275	0.445
dp	-0.484	-1.836	0.064	-0.237	-2.616	0.438	-0.202	-2.247	0.458	-0.119	-1.892	0.464	-0.155	-1.654	0.454
dy	-0.533	-2.092	0.078	-0.231	-2.533	0.437	-0.216	-2.397	0.460	-0.126	-2.008	0.466	-0.170	-1.871	0.455
ep	-0.240	-0.715	0.014	-0.137	-1.282	0.428	-0.110	-1.111	0.451	-0.063	-0.902	0.458	-0.079	-0.763	0.449
ik	60.179	2.250	0.072	32.227	3.246	0.443	33.329	3.407	0.469	21.279	3.602	0.483	48.967	4.972	0.467
infl	19.873	2.320	0.046	9.931	2.669	0.434	9.086	2.376	0.457	6.216	1.991	0.464	9.851	2.445	0.456
ltr	1.099	1.016	0.005	-0.495	-0.608	0.424	-0.542	-0.662	0.449	-0.773	-0.907	0.458	-0.662	-0.817	0.445
lty	9.206	3.532	0.101	3.887	2.893	0.439	3.362	2.596	0.460	1.883	1.978	0.464	3.217	2.241	0.455
ntis	-10.890	-2.557	0.045	-4.114	-1.867	0.429	-3.080	-1.315	0.451	-2.139	-1.431	0.459	-1.726	-0.731	0.445
tbl	7.525	2.848	0.075	3.421	2.732	0.438	3.285	2.696	0.461	1.982	2.509	0.468	3.968	2.853	0.462
tms	3.235	0.456	0.003	-0.195	-0.068	0.423	-1.713	-0.631	0.448	-2.087	-1.254	0.459	-4.442	-1.588	0.449

Table 6: Multivariate Regressions for Log Realized Variance

The table presents estimation results for ARMA(1,1) models of log realized variance *(lrvar)*. A simple ARMA(1,1) specification is compared to the fit of an ARMAX(1,1) that employs lagged macroeconomic forecasting variables. The variables *bm*, *dy* and *ep* are omitted due to high correlation with the included variable *dp*. Similarly, *lty* is omitted due to strong correlation with the included variable *tbl*. Results are presented for the sample period 1927q1 - 2005q4 and alternatively for 1952q1 - 2005q4. The estimated intercept is omitted to conserve space. Standard errors based on the Newey-West (1987) HAC estimator are presented in parenteses below coefficient estimates.

	1927q1 - 2005q4		1952q1 - 2005q4	
AR	0.9255***	0.83663***	0.86274***	0.80874***
	(0.03)	(0.06)	(0.06)	(0.08)
MA	0.44933***	0.44254***	0.39888***	0.4177***
	(0.07)	(0.09)	(0.10)	(0.12)
cay(-1)				-2.48
				(2.87)
de(-1)		0.08		0.15
		(0.12)		(0.22)
dfr(-1)		-5.12**		-6.11**
		(2.13)		(2.52)
dfy(-1)		17.79**		4.20
		(7.84)		(13.00)
dp(-1)		-0.147		-0.13
		(0.07)		(0.12)
ik(-1)				8.26
				(12.37)
infl(-1)		-0.134		3.74
		(2.25)		(5.30)
ltr(-1)		-1.783		-1.94
		(1.04)		(1.32)
ntis(-1)		1.448		-2.40
		(1.52)		(2.15)
tbl(-1)		0.056		1.26
		(0.97)		(2.01)
tms(-1)		-2.767		2.00
		(2.48)		(2.76)
AIC	-294.843	-291.163	-193.737	-193.047
SIC	-300.476	-313.698	-198.8	-216.674
Adj-R2	0.6089	0.6284	0.4515	0.4811
SIG-RES	0.6122	0.5968	0.5893	0.5732
Box-Pierce(12)	9.9303	10.934	6.7785	6.9772

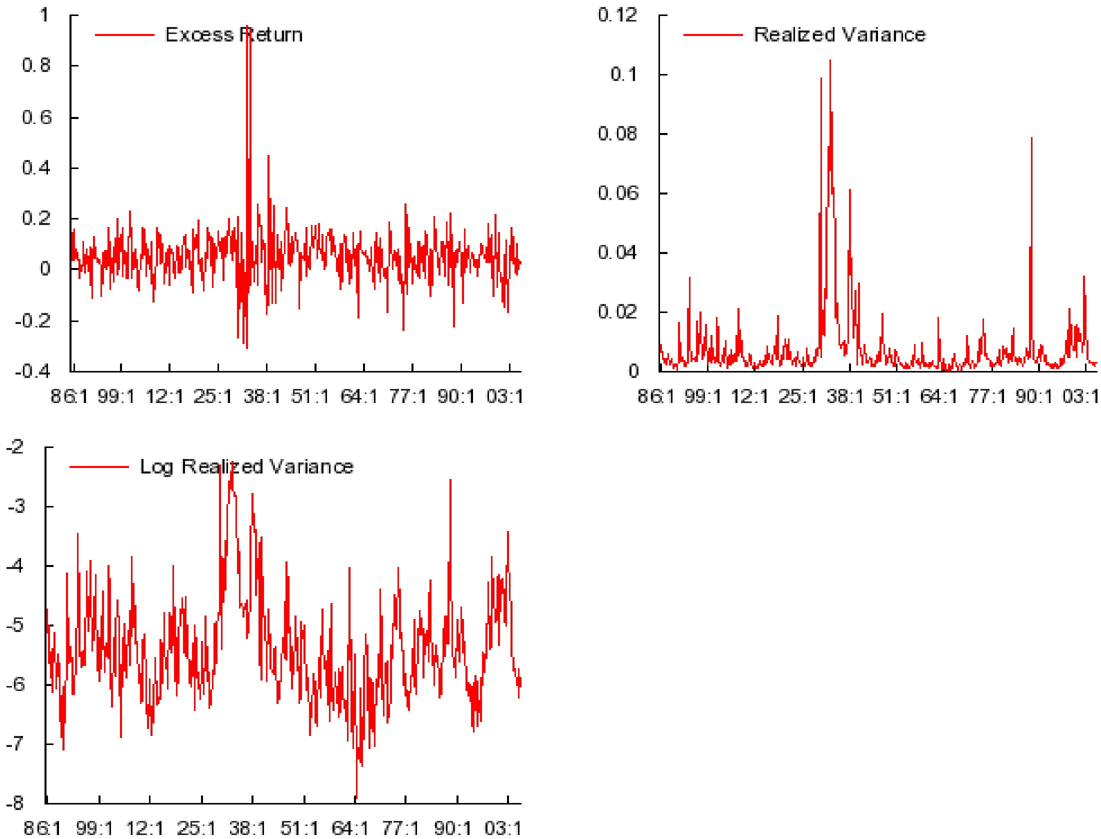

Figure 1: Time series plots of excess returns (top left), realized variance (top right) and the logarithm of realized variance (bottom left) over the sample period 1885:1 – 2005q4.

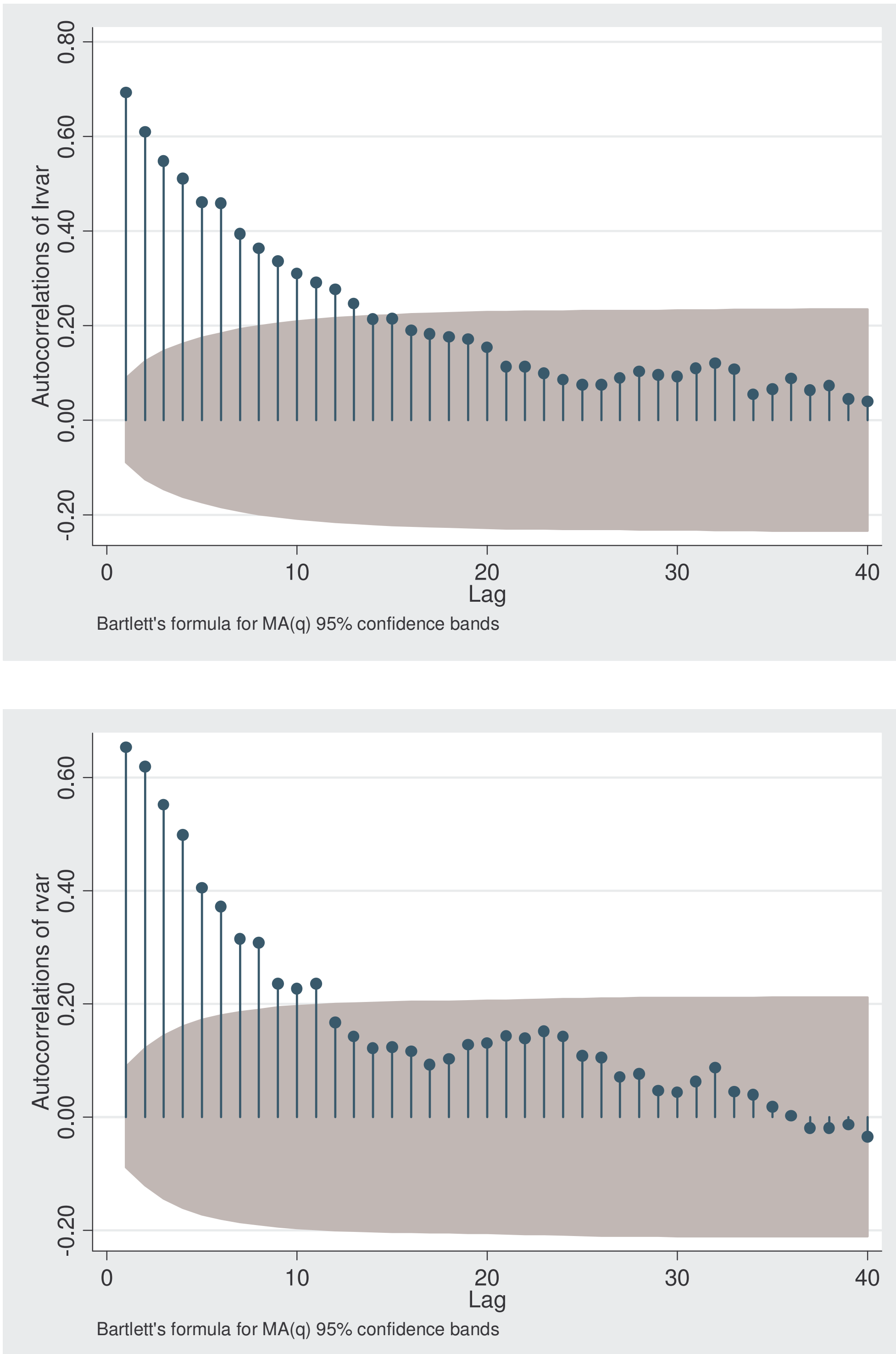

Figure 2: Correlograms for the logarithm of realized variance (*lrvar* – top panel) and the realized variance (*rvar* – bottom panel) along with 95% confidence bands.

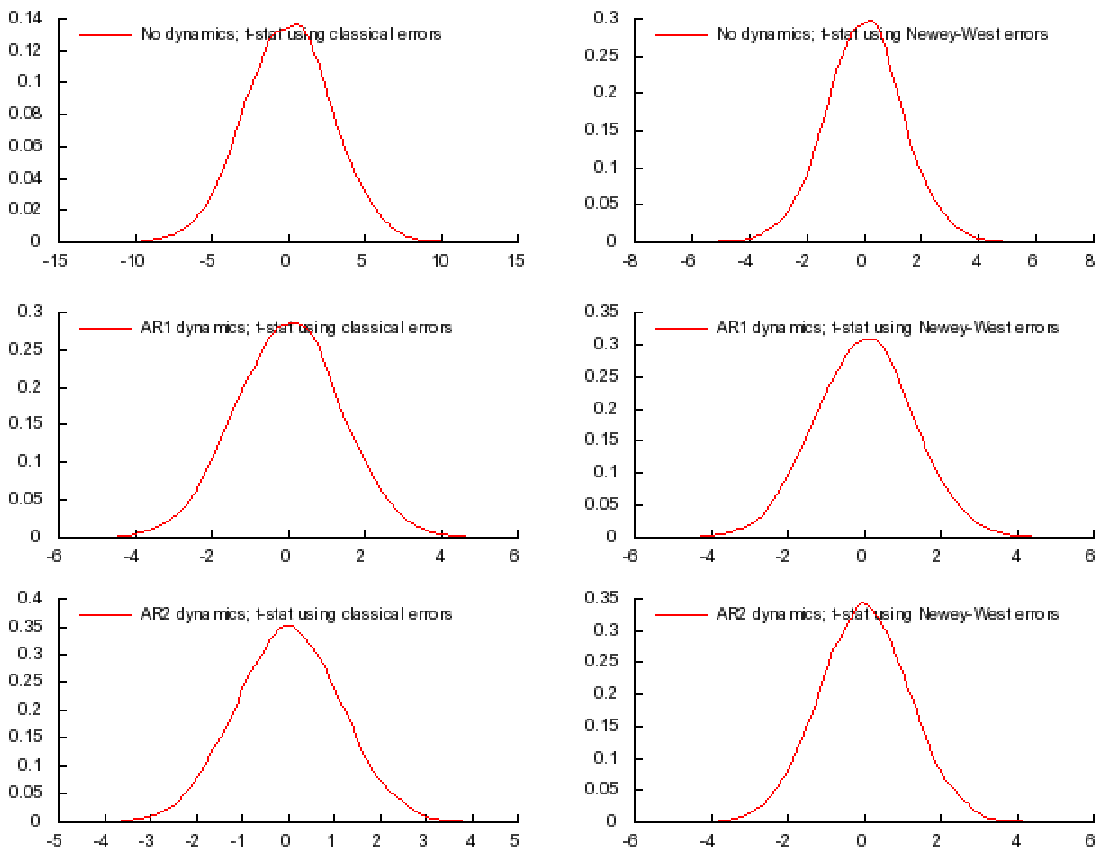

Figure 3: Density plots of *t*-statistics from regressions of Y_t on X_{t-1} over 10,000 Monte Carlo simulations. The simulations are based on equations (9) and (10) of the paper calibrated using the log realized variance (as *Y*) and the investment to capital ratio (as *X*). The correlation of shocks is set exactly to zero so that the forecasting variable is strictly exogenous. The sample size is 500 in all cases. In the top row the model includes no dynamics. In the second row the model includes one lag of the dependent variable, while in the third row the model includes two lags of the dependent variable.

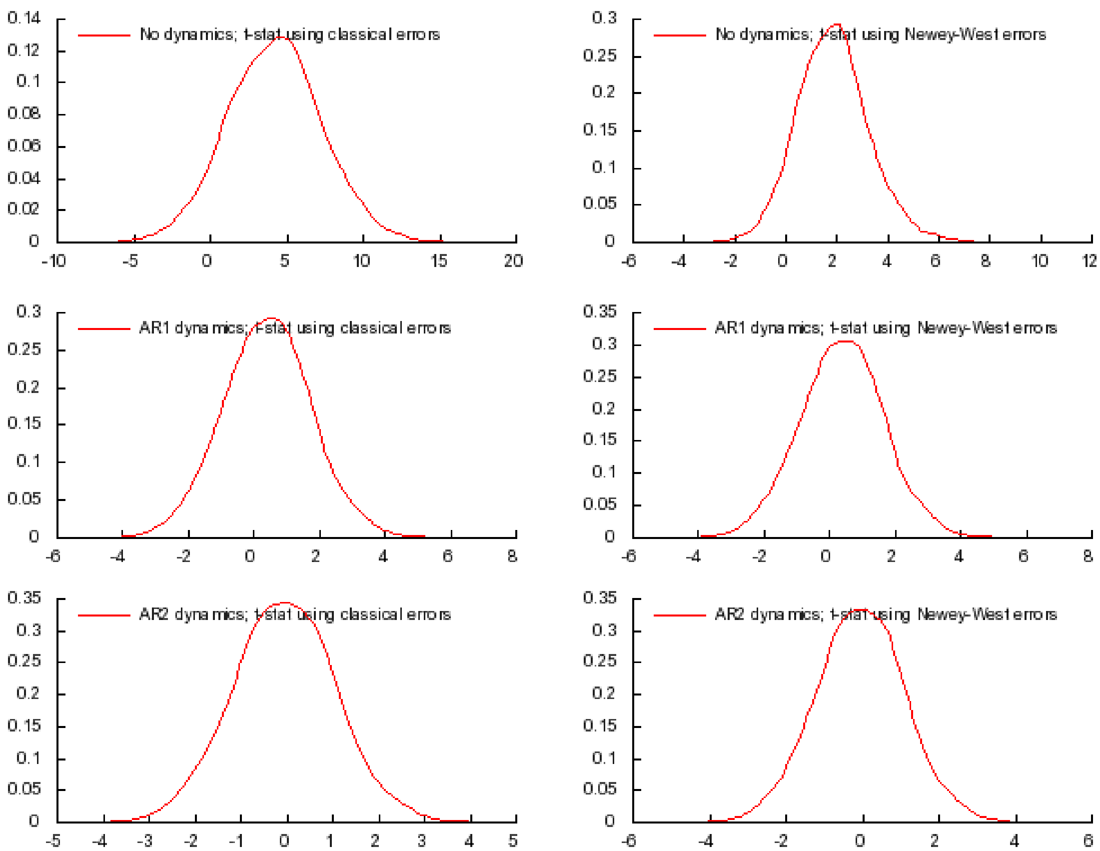

Figure 4: Density plots of *t*-statistics from regressions of Y_t on X_{t-1} over 10,000 Monte Carlo simulations. The simulations are based on equations (9) and (10) of the paper calibrated using the log realized variance (as *Y*) and dividend-price ratio (as *X*). The correlation of shocks is set to the empirical value of 0.45 over the post-Treasury accord sub-period. The forecasting variable is therefore only weakly exogenous. The sample size is 500 in all cases. In the top row the model includes no dynamics. In the second row the model includes one lag of the dependent variable, while in the third row the model includes two lags of the dependent variable.