The Impact of Sector and Market Variance on Individual Equity Variance[*†]

Haoming Wang[‡]

Professor George Tauchen, Faculty Advisor

[*] The Duke Community Standard was upheld in the completion of this report
[†] Honors thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University
[‡] Duke University; Durham, North Carolina, 2009

Acknowledgements

I would like to thank my advisor, Professor George Tauchen, for his help and support during the completion of this paper. I would also like to thank Professors Bjorn Eraker and Tim Bollerslev for their help with my research. The comments and suggestions of Brian Jansen, Allison Keane, Zed Lamba, Michael Schwert, and Caleb Seeley were also integral to the completion of this thesis.

Abstract

This paper investigates how changes in measures of sector and market variance affect equity variance by examining forecasts of equity variance over 1, 5, and 22 day time horizons. These forecasts were generated using heterogeneous autoregressive regressions that included measures of sector and market variance. The results demonstrate that sector and market variance both play an important role in determining equity variance. Further, the inclusion of measures of sector and market variance improves goodness of fit and decreases forecasting errors. These results imply that the inclusion of these measures could improve predictive models of equity variance.

1. Introduction

The variance of equity returns plays an important role in modern finance. Essential financial concepts such as derivatives pricing and risk management all depend on forecasts of equity variance. Recent large intra-day stock price swings have also demonstrated the importance of using high frequency financial data to calculate variance. This paper will investigate determinants of equity variance and attempt to create a model that generates more accurate forecasts of equity variance.

The availability of high frequency financial data has led to the development of predictive models of variance which utilize intra-day data. One such model is the heterogeneous autoregressive realized variance (HAR-RV) model developed in Corsi (2003). This model is used and expanded upon in later sections. The HAR-RV model specifies a stock's future variance solely as the function of measures of averages of the stock's past variances. The tendency of equity variance to form clusters of high and low variance provides justification for this specification. This paper proposes that additional information could be realized and the base HAR-RV model improved by including measures of sector and market variance.

The intuition for this addition is as follows: from popular models such as the Capital Asset Pricing Model (CAPM), we expect a company's stock performance to be affected by factors outside of its own corporate performance. In the case of the CAPM, instead of factors such as the company's earnings or its stock's dividend yield, returns on a market portfolio, the correlation between the stock and the market portfolio, and the risk free rate determine the returns on a company's stock. Even though Fama and French (2004) document the failings of the CAPM when applied to empirical data, its central tenets still hold: non-company-specific factors often play a role in determining the returns on a stock. For example, a hedge fund that is facing redemptions due to losses in the equity markets might have to sell stock to generate the cash necessary for meeting redemption requests. This process would depress returns for that stock regardless of the company's corporate performance. This leads to a natural extension to equity variance: if a stock's returns can be impacted by non-company-specific factors, we expect the variance of returns to also be impacted by non-company-specific factors. Two such factors are considered in this paper: the variance of a portfolio constructed from stocks in a company's sector, the sector variance, and the variance of a market portfolio, in this case the S&P 500.

The justification for the effect of sector variance is that a company participates in the same market as their competitors. Thus, it is expected that on average, an increase in the variance of the sector as a whole, *ceteris paribus*, leads to an increase in the variance of the stock. These changes in the variance of stock returns could reflect changing market conditions or any factor that extends beyond the individual stock and causes uncertainty about the sector. These changes in market conditions will most likely already be included in longer term measures of the stock's variance. However, in the short term it is reasonable to assume that there are sector specific factors in play because sector wide effects have not yet been fully incorporated into stock returns. Further, when market-wide variance is high, investors often exit the equity markets and instead invest in less volatile securities. Thus, we would expect measures of market variance to also impact stock returns and stock variance.

This paper examines the informational content of measures of sector and market variance in two specific industries: the pharmaceuticals industry and the banking industry. These two industries were chosen because they have disparate correlations, or beta factors, with the broader market and because of the potential to isolate the companies into specific industries. First, in sections two and three, the model of asset price movements that is used throughout this paper and methods to estimate equity variance are introduced. In section three, the base HAR-RV model is expanded to include measures of sector and market variance. In section four, the data used throughout the paper is examined and in section five, results of the model with sector and market variances are described. Section five also introduces re-specification of the model and discusses the new results that arise.

2. Model of Equity Returns and Variance

2.1. Diffusive and Jump-Diffusive Models

The foundation for our model of equity returns is the following stochastic differential equation which expresses the continuous logarithmic price path of a stock as

$$dp(t) = \mu(t)dt + \sigma(t)dW(t), \qquad 0 \leq t \leq T. \qquad (1)$$

In this model, $\mu(t)$ represents the time-dependent drift component of the asset price, $\sigma(t)$ represents the time dependent volatility of the price, and $dW(t)$ is a standard Brownian motion. However, Merton (1976) argued that intuitively, stock price dynamics cannot be represented by a stochastic process with a purely continuous path because we can often observe large discontinuous stock price movements which appear to be “jumps”. Due to this flaw in the pure continuous model, Merton added a jump component to the standard diffusive model. Empirical data support Merton’s assertion: the presence of large, discontinuous price movements in financial markets caused by new information is an oft observed phenomenon and is discussed in works such as Andersen, Bollerslev, and Diebold (2006). Under Merton’s jump-diffusion model, the path of the logarithmic price is expressed as

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) + \kappa(t)dq(t), \qquad 0 \leq t \leq T. \qquad (2)$$

The first two terms of the right side of Equation 2 represent the continuous part of the log price process and are exactly the same as in the pure diffusive model. The third term accounts for jumps: *q(t)* is a counting process (usually assumed to be the Poisson arrival process) and *k(t)* is the size of the corresponding discrete jump in the otherwise continuous price process.

2.2. Measures of Variance

The availability of high frequency intra-day financial data has allowed the development of measures of variance that better reflect the amount of variation in a stock's price. The importance of intra-day data is best seen with an example: on October 23rd, 2008, the S&P 500 opened the trading session at 902.99 and closed at 908.11 for a gain of 0.56%. However, during the day, the S&P reached a low of 865.24 and a high of 919.45, a trading range of 6%. Using price data at the daily frequency, or even low frequency intraday data, would lead to us to miss important information about how volatile the S&P 500 was on October 23rd.

An important measure of the variance in our model of equity returns is quadratic variation; however, quadratic variation is a continuous time calculation, and thus must be estimated instead of directly observed because changes in stock prices happen in discrete time intervals. One consistent estimator of quadratic variation is realized variance, which is a daily calculation. We let *t* be the day and *M* be the sampling frequency. Throughout the paper samples are taken at a 5 minute interval, which corresponds with *M*, the number of daily observations, equal to 78. Then intraday geometric returns are defined as

$$r_{t,j} = p\left(t - 1 + \frac{j}{M}\right) - p\left(t - 1 + \frac{j-1}{M}\right), \quad j = 1,2,3 \ldots M \qquad (3)$$

where *p* is the log price. The realized variance is then defined as

$$RV_t = \sum_{j=1}^{M} {r_{t,j}}^2 \qquad (4),$$

which is the sum of squared intra-day log returns. Andersen and Bollerslev (1998) emphasize that according to the theory of quadratic variation, the above converges to the integrated variance plus the jump component as the time between observations approaches zero. That is,

$$\lim_{M\to\infty} RV_t = \int_{t-1}^{t} \sigma^2(s)ds + \sum_{t-1<s\leq t} \kappa^2(s)\,. \qquad (5)$$

The first term corresponds to the integrated variance of the continuous process and the second term corresponds to the squared discrete jumps. This limit is also the definition of quadratic variation, which shows that realized variance is a consistent estimator of quadratic variation.

While the realized variance measures both the variation of the continuous process and the jump process, the realized bipower variation only measures the variation of the continuous process. The realized bipower variation is defined as

$$BV_t = {\mu_1}^{-2}\left(\frac{M}{M-1}\right)\sum_{j=2}^{M}|r_{t,j-1}||r_{t,j}| = \frac{\pi}{2}\left(\frac{M}{M-1}\right)\sum_{j=2}^{M}|r_{t,j-1}||r_{t,j}| \qquad (6)$$

with

$$\mu_a = \mathrm{E}(|Z|^a),\ \ Z\sim N(0,1), a > 0. \qquad (7)$$

By multiplying adjacent returns, the effect of returns that are jumps is mitigated. Bipower variation is thus a consistent estimator of the integrated variance of the continuous price process, which means that

$$\lim_{M\to\infty} BV_t = \int_{t-1}^{t} \sigma^2(s)ds. \qquad (8)$$

3. Methods

3.1. Background for HAR-RV Type models

The tendency of equity variance to form into clusters of low and high levels is a common characteristic of the financial markets and can be seen clearly in Figure 1. Because of this clustering, even though it is nigh-impossible to predict equity returns, the variance of equity returns can be forecasted. The generalized autoregressive heteroskedasticity model (GARCH) of Bollerslev (1986) and the HAR-RV model are two different ways to model variance clustering. Both models attempt to predict future equity variance based on historical measures of variance. However, only the HAR-RV model utilizes the additional information found in intraday price data.

GARCH forecasts variance as a function of past squared residuals and past predictions of variance. The most popular GARCH specification is the GARCH(1,1) model, which forecasts today's variance as a weighted average of yesterday's variance forecast and yesterday's squared residual term. On the other hand, the HAR-RV model predicts that future variance is a function of past averaged realized variances at different frequencies. The interpretation of this is that each historical measure of averaged realized variance will play a role in determining expectations of future variance. Changes in expectations of future variance will in turn cause traders to act, influencing future realized variance.

Empirical evidence from Andersen, Bollerslev, Diebold, and Labys (2003) has shown that linear models with historical intra-day variance measures as regressors, such as the HAR-RV model, are better predictors of variance than GARCH type models. This is because the GARCH model only makes use of daily return data, and isn't able to take advantage of the additional information found in the intra-day data. Due to the superior performance of HAR-RV type models, as well as the availability of high frequency intraday data, this paper will use the HAR-RV model as a foundation for further work.

3.2. HAR-RV Model Specification

HAR-RV type regressions make use of average realized variance over daily, weekly, and monthly periods. The average realized variance over some *h* discrete periods is

$$RV_{t,t+h} = \frac{RV_{t+1} + RV_{t+2} + \cdots + RV_{t+h}}{h} \qquad (9)$$

with $h = 1$ corresponding to daily periods, $h=5$ corresponding to weekly periods, and $h=22$ corresponding to monthly periods. Thus the standard HAR-RV model is expressed as:

$$RV_{t,t+h} = \beta_0 + \beta_D RV_{t-1,t} + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} + \varepsilon_{t+1}. \qquad (10)$$

The model is used to predict values for day-ahead, 5-days ahead, and 22-days ahead realized variances. These time horizons correspond to day-ahead, week-ahead, and month-ahead predictions of average realized variance. Throughout this paper, Newey-West covariance matrix estimators with a lag of 60 days are used to calculate standard errors. The Newey-West method for estimating standard errors is a common statistical procedure designed to account for auto-correlation in the residuals of time series data. The Newey-West method assumes that the correlation between residuals approaches zero as the time between observations approaches infinity. The covariance of lagged residuals is also weighted by an inputted maximum lag time horizon, in this case 60 days. Thus, greater weight is placed on the covariance of observations that are closer together. Newey-West estimates of standards errors are necessary given the time-varying nature of equity variance.

In order to determine the effect of sector and market variance on individual equity variance, two extra sets of regressors are added to the base HAR-RV model in Equation 10. First, the sector realized variance measure is constructed as an average of the realized variance of all companies in the S&P 100 within that sector. For example, the pharmaceuticals sector realized variance was constructed as the average of the realized variance of Abbott Labs (ABT), Bristol Myers Squibb (BMY), Johnson & Johnson (JNJ), Merck (MRK), and Pfizer (PFE). The market realized variance was calculated using S&P 500 futures data, also taken at five minute intervals. Thus, the first model introduced is of the following form:

$$\begin{aligned} RV_{t,t+h} = \; & \beta_0 + \beta_{CD} RV_{t-1,t} + \beta_{CW} RV_{t-5,t} + \beta_{CM} RV_{t-22,t} + \beta_{SD} RV_{sector,t-1,t} \\ & + \beta_{SW} RV_{sector,t-5,t} + \beta_{SM} RV_{sector,t-22,t} + \beta_{SPD} RV_{market,t-1,t} \\ & + \beta_{SPW} RV_{market,t-5,t} + \beta_{SPM} RV_{market,t-22,t} + \varepsilon_{t+1} \qquad (11) \end{aligned}$$

with *C* denoting the individual company, *S* denoting sector, and *SP* denoting market. While this model is the starting point for this paper, reduced forms of this model will also be examined in later sections. This model is referred throughout the paper as the *Full* model.

3.3. Jump Test Statistic

A part of recent literature on financial markets has been focused on developing tools to detect intraday jumps in asset prices. In this section of the paper we examine statistics that test for jumps that are taken from Barndorff-Nielsen and Shephard (2004). We can see from the definition of realized variance and bipower variation that

$$\lim_{m\to\infty}(RV_t - BV_t) = \sum_{t-1<s\leq t} \kappa^2(s), \qquad (12)$$

which is the jump component of asset price variance. Thus, a measure of the percentage of total variance caused by the jump process is the relative jump,

$$RJ_t = \frac{RV_t - BV_t}{RVt}. \qquad (13)$$

In order to standardize the relative jump into units of standard deviation, any test statistic also needs to incorporate an estimate of integrated quarticity. One such estimator that is commonly used is the Tri-Power Quarticity, which is defined as:

$$TP_t = M\mu_{\frac{4}{3}}^{-3}\left(\frac{M}{M-2}\right)\sum_{j=3}^{M}\left|r_{t,j-2}\right|^{4/3}\left|r_{t,j-1}\right|^{4/3}\left|r_{t,j}\right|^{4/3}. \qquad (14)$$

Huang and Tauchen (2005) discovered that in simulations, the best performing test statistic for jump detection is the Z-Tripower-Max test statistic, which is defined as:

$$Z_{TP,rm,t} = \frac{RJ_t}{\sqrt{(v_{bb} - v_{qq})\frac{1}{M}\max(1, \frac{TP_t}{BV^2})}}, v_{bb} - v_{qq} = \left(\frac{\pi}{2}\right)^2 + \pi - 5, \qquad (15)$$

where $Z_{TP,rm,t}$ has a standard normal distribution. Values of the test statistic that correspond to p-values at the 0.001 level are flagged as days with jumps.

4. Data

High frequency data for 11 individual stocks in two sectors as well as S&P 500 futures were examined. The data were purchased from an online data vendor, price-data.com, and include the pharmaceuticals: Abbott Labs (ABT), Bristol Myers Squibb (BMY), Johnson & Johnson (JNJ), Merck (MRK), and Pfizer (PFE) as well as the banks: Bank of America (BAC), Bank of New York (BK), Citigroup (C), JP Morgan (JPM), US Bancorp (USB), and Wells-Fargo (WFC). Data for the pharmaceuticals sector were obtained for 10 years, from April 1997 until the end of October 2007. This corresponds to a total of 2,682 trading days of price data. In-sample results for the model were calculated from the beginning of the dataset until October 2006, in order to have a year's worth of data to test the model out-of-sample. Data for the banking sector are from August 1997 until October 2007. This corresponds to a total of 2,595 trading days of price data. Similar to the pharmaceutical stocks, the last year's worth of data was used for out-of-sample testing. In-sample results were again generated using data from August 1997 to October 2006. For all datasets, non-full trading days were removed so that each day would have the full complement of 78 5-minute observations which were then used to calculate multiple measures of variance.

5. Results

The preceding sections have been concerned with creating a predictive model of equity variance incorporating sector and market variances. In the following sections, results for the *Full* model are examined. In the course of this examination, it becomes apparent that the *Full* model needs to be pared down. After a discussion of the results for the *Full* model, reduced forms of the *Full* model are introduced and discussed.

5.1. In-sample Results (*Full* Model)

By examining the R-squared measures of both the standard HAR-RV and the *Full* model, one can see whether measures of sector and market variance include additional information about individual equity variance. Since the R-squared is the proportion of variation in the dependent variable explained by variation in the regressors, models with higher R-squared are considered to be superior. Further, by examining the coefficients of the regression models, we can better understand what determines the variance of individual equities, and how that effect might change across companies and industries. Coefficient estimates for the pharmaceuticals and R-squared measures for both industries can be seen in tables 2 through 7.

Both the HAR-RV and the *Full* model were tested over three time spans: day-ahead predictions, 5 days-ahead predictions, and 22 days-ahead predictions. As in previous papers such as Fradkin (2007), we see the highest R-squared for 5 days-ahead predictions, followed by 22 days, and then finally 1 day ahead regressions. This makes sense: the model should be better at explaining the measures of realized variance that have been smoothed out over 5 or 22 days because realized variance can undergo large day-to-day changes.

The in-sample results for both the pharmaceuticals and the banks suggest that the inclusion of measures of realized variance for both the sector and the market provides additional information for predictions of equity realized variance. Across the board, goodness of fit for the models is increased. There is an average increase in R-squared of 5.5% for day-ahead predictions, 5.95% for 5 days-ahead predictions, and 5.85% for 22 days-ahead predictions. Further, there is a wide range in improvements in R-squared, with increases ranging from 17% to 1.6%. The *Full* model was also separated into two versions: one version that just included the sector regressors and one version that just included the market regressors. On average, the R-squared increase from a model incorporating only the stock and the market to the *Full* model is greater than the increase from a model incorporating only the stock and the sector to the *Full* model. This might suggest that measures of sector realized variance include more unique information than the market realized variance. We can also conclude that both the sector and the market realized variance measures contain unique information because models incorporating only one or the other have lower explanatory power than models that contain both. One important note is that we see these improvements in explanatory power for stocks in both the banking and pharmaceuticals sectors. For two such structurally different industries to exhibit the same benefits from the inclusion of measures of sector and market variance suggests that these are wide ranging results that can be applied to most equities.

Further, the coefficients provide some interesting insights: for almost all stocks, the one day lagged sector realized variance is statistically significant, positive, and also relatively large. From this we can conclude that there are clear sector effects that determine equity variance. Again this makes intuitive sense, we would expect there to a positive relationship between sector variance and equity variance because we expect some spill-over from factors that would influence the entire sector. The importance of the lagged daily sector realized variance can be seen in both the banks and the pharmaceuticals. For such a consistent result to be seen across these two very different industries, it would suggest that the previous day's sector variance plays an important role in determining the variance of most equities. This result provides some evidence for sector variance influencing stock variance.

However, in general, careful interpretation of these coefficient estimates is required. Some coefficient estimates are easily interpretable, with the daily lags decreasing in importance as the predictive time horizon increases and the weekly and monthly lags increasing in importance. Unfortunately, many of the coefficients are not statistically significant, which is especially troubling for the lagged values of the realized variance of the stock itself. The coefficient estimates for the *Full* model also consistently exhibit a negative coefficient for the lagged monthly market realized variance. This means that increases in the level of the average market variance over the last month actually imply lower levels of future equity variance. One possible reason why this occurs is because the regression coefficient is defined as the predicted impact on the stock's realized variance when the market realized variance rises, while keeping stock and sector-specific realized variance constant. Thus, this negative coefficient could demonstrate the tendency of market participants to rotate away from volatile stocks and into more stable stocks.

Interestingly, this effect is different across the banking and pharmaceutical sectors. While this negative relationship is statistically significant across the pharmaceuticals, it is no longer statistically significant with the bank stocks. Further, the absolute values of the negative coefficients are lower, and in some cases are even positive. This could be because the beta factors, a measure of the correlation of stock returns to market returns, of both industries are significantly different. The beta factor of the banking sector is 1.15 while that of pharmaceuticals is 0.70, meaning that returns from bank stocks are much more correlated with the market than returns from pharmaceutical stocks. This could lend more evidence to the view that this negative coefficient is tied to investors rotating their holdings into certain industries. The negative correlation between equity returns and equity variance is a well-known characteristic of stock markets and is the reason for implied volatility skew on equity options. Thus, if the variance of the market portfolio increases, this could mean that on average, the returns on the market portfolio are decreasing. Because banking stocks are pro-cyclical, returns on banking stocks are likely to be depressed as well. This leads to selling, which could mitigate the impact of the rotation interpretation. However, another possibility is that our model needs to be re-specified, a subject that will be discussed in the next section.

5.2. Model Shortcomings

One specific aspect of the *Full* model that makes interpretation of coefficient estimates difficult is the high correlation between some regressors. There is high correlation between market and sector variables: the correlation between the S&P monthly realized variance and the sector monthly realized variance is 0.9298 for the banks and 0.9127 for the pharmaceuticals. Further, there is high correlation between the sector and market weekly realized variance for the banks and the pharmaceuticals. There is also high correlation between daily sector realized variance and daily market realized variance for the pharmaceuticals. Finally, the data also exhibit high correlation between lagged monthly realized variance for the equity and for the market in both sectors.

This high correlation between the regressors is known as collinearity. While the presence of collinearity does not affect the explanatory ability of a regression, it makes it so that coefficient estimates become unpredictable. This occurs since it is impossible to determine the individual effects of the regressors on the dependent variable because the regressors are so highly correlated. Since part of the goal of this paper is to investigate the dynamics of what determines individual equity variance, collinearity and the associated unreliable coefficients pose a problem. Because there are highly correlated regressors, re-specifying the original model with new regressors might solve this collinearity problem.

However, collinearity is not the only reason why re-specification of the *Full* model might be necessary. Harvey (1980) and Granger and Newbold (1974) show that in general, regressions involving economic variables in levels can be misleading and could cause problems with interpretations. One possible solution suggested by both papers is to use the first differences between variables instead of the levels. Because we have coefficients whose values seem difficult to interpret, taking differences between related variables could yield more interpretable results.

5.3. Model Re-specification

The presence of collinearity and the difficult to interpret results of the *Full* model motivate the need for re-specification of the *Full* model. The first step is to determine which variables are significant in the original model. Multiple F-tests were performed on the *Full* model to determine which sets of regressors were jointly insignificant and could thus be dropped from the model. One pattern emerged: for almost every stock and time horizon combination the coefficients of the lagged weekly and monthly sector realized variance, as well as the lagged weekly market realized variance, were statistically insignificant. The p-values of the coefficients for these regressors can be seen in Table 11. Further, because of the high correlation between the lagged monthly market and equity realized variances, a new regressor was constructed as the difference between the lagged daily market realized variance and the monthly market realized variance. This new regressor has a clear interpretation: it is the daily deviation from the average realized variance over the last month. A new *Parsimonious* model was created by taking the *Full* model and dropping the lagged weekly and monthly sector realized variance, as well as the lagged weekly market realized variance. Finally, the lagged daily and monthly market realized variances were replaced with the measure of their difference. The model is defined as:

$$RV_{t,t+h} = \beta_0 + \beta_{CD} RV_{t-1,t} + \beta_{CW} RV_{t-5,t} + \beta_{CM} RV_{t-22,t} + \beta_{SD} RV_{sector,t-1,t} + \beta_{SPDiff} (RV_{market,t-1,t} - RV_{market,t-22,t)} + \varepsilon_{t+1}. \quad (16)$$

The intuition behind this model is that the determinants of equity realized variance are the historical measures of the stock's realized variance as well as yesterday's sector variance and the daily deviation from the monthly average market realized variance.

While this *Parsimonious* model is mostly robust through both the day-ahead predictions and the week-ahead predictions, it struggles to match the performance of the *Full* model for month-ahead predictions. For the updated month-ahead predictive model, the *Parsimonious* model is taken; however, the daily deviation from the monthly mean of market realized variance is replaced with the lagged monthly level of market realized variance. The lagged monthly level of sector realized variance is also added back to the model. This replacement causes the problem of collinearity to reappear, but the difference in explanatory power is too pronounced to ignore. The third and final model that we use will be referred as the *Reduced Monthly* model and is:

$$RV_{t,t+h} = \beta_0 + \beta_{CD} RV_{t-1,t} + \beta_{CW} RV_{t-5,t} + \beta_{CM} RV_{t-22,t} + \beta_{SD} RV_{sector\,,t-1,t} + \beta_{SM} RV_{sector\,,t-22,t} + \beta_{SPM} RV_{market\,,t-22,t} + \varepsilon_{t+1}. \qquad (17)$$

This model suggests that average monthly equity variance is driven by its own lagged realized variances, the level of lagged daily and monthly sector realized variance, and the level of the lagged monthly market realized variance. This need for re-specification makes some sense: we would expect that the level of these longer time horizon realized variance measures to have greater impact on the longer time horizon predictions than on the shorter time horizon predictions.

For the day-ahead predictions, there is an average decrease in R-squared from the *Full* model by using the *Parsimonious* model of only 0.56%. For week-ahead predictions, the loss in explanatory power is higher, at 1.04%. Using the *Parsimonious* model for the month-ahead predictions, there is a decrease in R-squared of 2.3%. However, using the *Reduced Monthly* model, the percent decrease in R-squared is only 0.56%. The full R-squared results can be seen in Tables 4-6.

Interpretation of the coefficient estimates of the *Parsimonious* model is also much clearer than with the *Full* model. All coefficients are positive and we have mitigated the problems of collinearity by introducing the market difference as a regressor. As before, there is a clear pattern that the coefficient estimate for the lagged daily sector realized variance is both relatively large as well as statistically significant across almost every time horizon. This also holds in the *Reduced Monthly* model. It appears that sector variance is an important factor in individual equity variance. The coefficient estimate for the daily deviation in market realized variance from the last month's average is also always positive. The interpretation of this result is that an increase in market variance relative to the last month's average has a positive impact on equity variance.

The *Parsimonious* model is remarkable because the *Full* model has been distilled into a streamlined version that highlights the critical factors that drive equity variance. This result provides further evidence for the intuitive guess that sector and market variance impact stock variance. Moreover, the *Parsimonious* model highlights that these effects mainly occur in the short-run. They act in the one day and one week time horizons through yesterday's sector variance and yesterday's deviation in market variance from the average over the last month. That they act in the short run is to be expected: it seems intuitive that the information contained in averages of weekly and monthly sector and market variances has already been incorporated into stock returns. Complete results for the *Parsimonious* and *Reduced Monthly* models can be seen in Tables 11-13.

By examining both the *Full* model and these two new models, we can conclude that there are sector and market factors which play a role in equity variance. Moreover, the *Full* model, which includes all measures of sector and market variance, has been streamlined into more intuitive versions. This was accomplished by using F-tests to discover coefficients that were consistently insignificant. It is also striking that across all models and all time horizons, the lagged daily sector realized variance plays a significant role in determining the variance of a stock. This demonstrates that while new information about the sector or the market is often already realized in longer term measures of equity variance, in the short run, there is unique information contained in the variance of a sector or the variance of the market. Finally, the inclusion of measures of sector and market realized variances clearly provide more accurate results when forecasting equity variance.

5.4. Out-of-sample Results

The evolution of realized variance over the full sample period is especially striking, as seen in Figure 1; there are clearly distinct low variance and high variance regimes. In order to examine how the models investigated in this paper perform in these two regimes, all models were tested out-of-sample using data from November 2006 to November 2007. This date range is particularly interesting because there are two distinct periods of low and high variance. The first half of the sample is a period of rather low equity variance; however, starting in the summer of 2007, we see marked increases in equity variance due to the subprime mortgage crisis. This will allow testing of each model in two very distinct environments, similar to the circumstances that we see in the full sample. Out-of-sample testing was performed separately using each model for all of November 2006 to November 2007, the first half of the sample, and the second half of the sample.

One clear pattern is that for the earlier sampling period of relatively low equity variance, models that incorporate sector and market realized variances consistently outperform the base HAR-RV model. Further, there is a clear difference between how the models perform for the first half of the data and for the second half: model performance is consistently better during the lower variance regime. The benefits of including sector and market variance measures in the model also are much greater in the earlier part of 2007. This effect is especially pronounced for the banking sector. For the pharmaceuticals sector, on average the mean-squared error (MSE) of the *Full* model is 19% lower than that of the base HAR-RV model for the early period while the *Full* MSE is only 10% lower for the later period. For the banking sector, on average the MSE of the *Full* model is again 20% lower than of that of the base model; however, for the later period, the MSE of the *Full* model is actually higher by about 1%. The poor performance of the *Full* model is troubling. Since on average the banking sector was more volatile during the credit crisis than the pharmaceuticals sector and the *Full* model for the banking sector performed worse than the base model, this would suggest that the performance of the *Full* model suffers in periods of high variance.

This difference in performance is apparent when looking at Figures 2-4, which plots the residuals and out-of-sample predictions for JP Morgan. There's a marked increase in the level of the residuals once the subprime mortgage crisis begins and we enter the high variance regime. Further, we see one specific spike in variance that the models incorporating market and sector variance dramatically overestimate. While in-sample performance of the *Parsimonious* model was relatively similar to that of the *Full* model, the out-of-sample performance of the *Parsimonious* model is significantly worse, offering much less improvement over the base HAR-RV model. Even though there is the problem of collinearity with the *Full* model, it seems that its explanatory ability is much stronger. However, the *Reduced Monthly* model does perform better on average.

One interesting dynamic is that for the first half of the year, during the low variance regime, the models all consistently over-predict equity realized variance. This is especially obvious for the week and month-ahead forecasts. This result can also be seen in Fradkin (2007), which attributes this to the presence of a constant used to fit the model to the early part of the sample which causes model performance to suffer during periods of lower variance. Once we enter the second half of the testing period, this bias decreases as we enter a period of higher overall variance. For some stocks, there is even a period where the models under-predict equity realized variance. Similar to how the constant was too high in the earlier part of our sample, the constant term which caused a bias towards over-predicting equity variance in periods of lower market stress is now too low to accurately predict the level of variance. This is especially true for the bank stocks because during the second half of the sample, the financial sector was hit the hardest by the turmoil in the mortgage backed securities market. With the recent turmoil in the financial markets, it seems that authors of further work on HAR-RV type models using data from 2008 have to be especially cognizant of the constant term's impact on the accuracy of predictions of realized variance.

The performance of the models introduced in this paper in periods of high volatility also requires greater investigation. Conveniently we are in the middle of one such period: since September 2008, six of the ten largest daily percentage gains and five of the ten largest daily percentage losses in the S&P 500 have occurred. Unfortunately, datasets used for this paper only included price data up to 2007 but it is still possible to discuss what results we would expect had more recent data been included. Since the summer of 2007, volatility in the stock prices of banks has been driven by uncertainty over how to value the assets held on their balance sheets. This uncertainty has increased the volatility of stocks in other sectors for a variety of reasons that range from concerns about how declining credit availability will impact corporate earnings to institutions selling stock in order to meet demands for cash. We would expect that for banking stocks, sector variance effects would become more prominent because there is sector wide uncertainty about the health of American financial institutions. In other words, investors are less concerned about asynchronous company risk and more concerned about sector wide risk. Conversely, we would expect that for pharmaceutical stocks, the market variable has a larger role in determining equity variance because broader market events are playing a greater role in determining returns on pharmaceutical stocks. For example, it could be that market participants are simply unwilling to stomach the large swings in the stock market and instead choose to hold other assets, or that funds are forced to sell their holdings of pharmaceutical stocks in order to meet investor redemption requests. Overall we would also expect the results to be less accurate because the model was calibrated using data from a period of much lower volatility.

5.5 Jump Test Results

In addition to investigating sector specific effects on stock variance through the HAR-RV model, possible sector specific effects on stock price jumps were also investigated. This section of the paper attempts to find sector effects for equity variance by examining periods of shared jump days between companies in the same industry and attempting to link these shared jumps to a possible news event. Jump days were examined for the pharmaceuticals industry in 2007 and the Wall Street Journal was used to determine if there were any shared jump days that could be attributed to sector-related news events. Summary statistics for the pharmaceuticals sector are provided in Table 14.

Because pharmaceutical companies are often in direct competition with each other, we might expect that a significant jump in one stock is associated with significant jumps in other stocks in the sector. Three clusters of shared jump days were found: one from January 29th-31st, another in February 14th, and one last cluster October 16th-17th. In each case, it was difficult to discern any news that would likely affect the entire industry. Two events could have motivated increases in sector wide jump measures: for the first cluster of jumps, the Thai government ruled that they would sell special generic versions of drugs made by Abbot Labs and Bristol-Myers. During that period, Abbot Labs, Merck, and Pfizer all had significant jumps. For the second cluster, a European pharmaceuticals company released earnings and there were rumors that Bristol-Myers would be acquired. On that day, both Bristol-Myers and Pfizer underwent significant jumps.

The results show that it is difficult to investigate sector wide effects using the flagged jump days. First, statistically significant jumps should reflect new and important information entering the market, or else jumps would already be priced in. It is difficult to determine to what extent these news items were surprises for the market. It is also difficult to isolate exactly what motivates these jumps by simply looking for Wall Street Journal articles from that day. Finally, there could be sector wide effects; however, they might not be strong enough to cause statistically significant jumps. This qualitative process is neither elegant nor mathematically satisfying, as we cannot quantify these sector effects on equity variance.

6. Conclusion

This paper has examined the inclusion of sector and market realized variance measures in the HAR-RV model in order to better understand the influence of sector and market variance on a specific stock's variance. The intuition behind this was provided by popular models such as the CAPM, which states that a stock's returns are affected by non-company specific factors. This paper has determined that the variance of a stock's returns is also affected by non-company specific factors. These factors were identified by first creating a model that included all relevant measures of sector and market variance, then paring down that model using relevant statistical methods. The results demonstrate that sector and market variances do play a role in determining a stock's variance, especially within the one week time horizon. Yesterday's sector variance and the daily deviation in the market variance from the previous month's average both consistently impact a stock's variance.

In-sample and out-of-sample results also show that models which incorporate measures of sector and market realized variance do a better job of forecasting variance in normal market conditions. The inclusion of sector and market variance measures lead to consistent increases in R-squared and decreases in mean squared errors. Overall, these results hold across both the banking sector and the pharmaceuticals sector. However, in periods where there are high levels of stock volatility, such as the past few months, these results might not hold. While the HAR-RV model is a tremendously useful tool to forecast equity variance, this paper has demonstrated that the basic model can generally be improved by the inclusion of measures of sector and market variance. This is an important development because currently most models that forecast equity variance only focus on measures of equity variance. These results suggest that there needs to be additional research done on incorporating measures of sector and market variance into models of equity variance. Further, the performance of predictive models of equity variance in periods of high market stress also requires additional investigation.

7. Tables

Tables 1 – 3 correspond to the *Full* model and equation 11

Table 1: Coefficient Estimates and Significance Levels Day-Ahead

	ABT	BMY	JNJ	MRK	PFE
βCD	0.1664***	0.0422	0.1046*	0.1556***	0.1328**
βCW	0.2039**	0.0453	-0.0826	0.0315	0.0055
βCM	0.5038***	0.5735***	0.4859**	0.1563	0.5300***
βSD	0.2420*	0.5398***	0.3010***	0.1822*	0.2283***
βSW	-0.059	-0.0076	0.1991*	0.1216	0.0945
βSM	-0.0535	-0.1063	-0.1956	0.1979	-0.0844
βSPD	0.2561	0.1185	0.1228	0.3043*	0.3788***
βSPW	0.0392	0.3303	0.5201	-0.0572	0.2057
βSPM	-0.4811**	-0.5143	-0.5261*	-0.3814*	-0.6540*
β0	.0000198	0.000005	-0.0000009	0.000041**	0.00004***

* p<0.05, **p<0.01, ***p<0.001

Table 2: Coefficient Estimates and Significance Levels Week-Ahead

	ABT	BMY	JNJ	MRK	PFE
βCD	0.0902**	-0.0141	0.0493	0.0789**	-0.00046
βCW	0.1854	0.0724	-0.0733	-0.04395	0.1308
βCM	0.6272***	0.6081**	0.5064*	0.0820	0.4728**
βSD	0.1036	0.2397**	0.1907***	0.0454	0.1990**
βSW	0.0473	0.2747	0.3081*	0.3411	0.0496
βSM	0.0214	-0.0382	-0.1060	0.3472	0.1411
βSPD	0.1492*	0.2371*	0.2697*	0.1441*	0.2328*
βSPW	0.0766	0.0110	0.0545	0.0294	0.1481
βSPM	-0.6822**	-0.6222	-0.4817*	-0.5288*	-0.7767**
β0	0.00004*	0.0000196	-0.000009	0.00006***	0.00006***

* p<0.05, **p<0.01, ***p<0.001

Table 3: Coefficient Estimates and Significance Levels Month-Ahead					
	ABT	BMY	JNJ	MRK	PFE
βCD	0.0473**	-0.0099	0.0115	0.0208	0.0078
βCW	0.1083*	0.1307	-0.0596	-0.0220	0.1779*
βCM	0.8062***	0.4909*	0.4526	-0.1349	0.2707*
βSD	0.0627	0.1771***	0.1190***	0.0766*	0.1113***
βSW	0.0634	0.2053	0.2549	0.2130	-0.0524
βSM	0.1242	0.2219	0.1615	0.7081**	0.6033**
βSPD	0.0678	0.1242*	0.1110*	0.0785*	0.1046*
βSPW	0.0074	-0.0965	-0.0545	-0.0577	0.0003
βSPM	-1.1189***	-0.9008***	-0.6644**	-0.7346*	-1.064***
β0	0.000053**	0.0000562*	0.000033*	0.00009***	0.0000841
* p<0.05, **p<0.01, ***p<0.001					

Table 4: In-sample R-Squared (1 day ahead)												
R-Squared		ABT	BAC	BK	BMY	C	JNJ	JPM	MRK	PFE	USB	WFC
HAR-RV		0.5133	0.5178	0.3331	0.4264	0.4788	0.4538	0.4739	0.2578	0.3511	0.5149	0.4139
S&P Only		0.5285	0.5219	0.3439	0.4525	0.4909	0.5222	0.5413	0.2766	0.3911	0.588	0.4733
Sector Only		0.528	0.5363	0.3496	0.4696	0.49	0.5407	0.5525	0.2792	0.3799	0.5941	0.5167
Full		0.5359	0.5403	0.3513	0.4719	0.4954	0.5563	0.5636	0.2884	0.3985	0.6142	0.525
Parsimony		0.5352	0.5342	0.3522	0.4706	0.4919	0.5417	0.5474	0.2871	0.3975	0.6112	0.5102
% Increase	Base to Full	2.26%	2.25%	1.82%	4.55%	1.66%	10.25%	8.97%	3.06%	4.74%	9.93%	11.11%
	S&P Only	1.52%	0.41%	1.08%	2.61%	1.21%	6.84%	6.74%	1.88%	4.00%	7.31%	5.94%
	Sector Only	1.47%	1.85%	1.65%	4.32%	1.12%	8.69%	7.86%	2.14%	2.88%	7.92%	10.28%
% Loss	Full to Pars	-0.07%	-0.61%	0.09%	-0.13%	-0.35%	-1.46%	-1.62%	-0.13%	-0.10%	-0.30%	-1.48%

Table 5: In-sample R-Squared (Week ahead)												
R-Squared		ABT	BAC	BK	BMY	C	JNJ	JPM	MRK	PFE	USB	WFC
HAR-RV		0.6209	0.6366	0.5187	0.604	0.532	0.5225	0.4353	0.3483	0.4684	0.622	0.5958
S&P Only		0.6332	0.6413	0.5274	0.6336	0.5497	0.6	0.5602	0.3672	0.5015	0.6493	0.6206
Sector Only		0.628	0.6641	0.5355	0.6399	0.5558	0.6195	0.5872	0.4008	0.4987	0.6637	0.642
Full		0.6369	0.6681	0.5355	0.6481	0.5623	0.6427	0.6141	0.4095	0.5161	0.668	0.6463
Parsimony		0.632	0.6576	0.5358	0.6435	0.5532	0.6333	0.5792	0.3843	0.5112	0.6654	0.6376
% Increase	Base to Full	1.60%	3.15%	1.68%	4.41%	3.03%	12.02%	17.88%	6.12%	4.77%	4.60%	5.05%
	S&P Only	1.23%	0.47%	0.87%	2.96%	1.77%	7.75%	12.49%	1.89%	3.31%	2.73%	2.48%
	Sector Only	0.71%	2.75%	1.68%	3.59%	2.38%	9.70%	15.19%	5.25%	3.03%	4.17%	4.62%
% Loss	Full to Pars	-0.49%	-1.05%	0.03%	-0.46%	-0.9%	-0.94%	-3.49%	-2.5%	-0.49%	-0.26%	-0.87%

Table 6: In-sample R-Squared (Month ahead)												
R-Squared		ABT	BAC	BK	BMY	C	JNJ	JPM	MRK	PFE	USB	WFC
HAR-RV		0.5992	0.5853	0.482	0.5927	0.4984	0.4624	0.4269	0.3437	0.4609	0.6295	0.6514
S&P Only		0.6305	0.5886	0.4914	0.6117	0.5129	0.4919	0.5042	0.354	0.4771	0.646	0.6656
Sector Only		0.6036	0.5916	0.5006	0.6115	0.5244	0.5234	0.4971	0.4368	0.472	0.6428	0.666
Full		0.6364	0.6023	0.5016	0.6277	0.5338	0.5482	0.5184	0.4663	0.511	0.6515	0.6764
Parsimony		0.603	0.5898	0.4989	0.6106	0.516	0.5186	0.5007	0.3996	0.4754	0.6442	0.6642
New 22		0.6355	0.6004	0.5007	0.6252	0.5221	0.5383	0.4992	0.4605	0.5084	0.6472	0.6746
% Increase	Base to Full	3.72%	1.70%	1.96%	3.50%	3.54%	8.58%	9.15%	12.26%	5.01%	2.20%	2.50%
	S&P Only	3.13%	0.33%	0.94%	1.90%	1.45%	2.95%	7.73%	1.03%	1.62%	1.65%	1.42%
	Sector Only	0.44%	0.63%	1.86%	1.88%	2.60%	6.10%	7.02%	9.31%	1.11%	1.33%	1.46%
% Loss	Full to Pars	-3.34%	-1.25%	-0.27%	-1.71%	-1.78%	-2.96%	-1.77%	-6.67%	-3.56%	-0.73%	-1.22%
	Full to Red 22	-0.09%	-0.19%	-0.09%	-0.25%	-1.17%	-0.99%	-1.92%	-0.58%	-0.26%	-0.43%	-0.18%

Table 7: Out-of-sample Mean Squared Errors 1 Day Ahead												
		ABT	BAC	BK	BMY	C	JNJ	JPM	MRK	PFE	USB	WFC
Full	Base	0.808	1.33	8.8	1.75	**2.89**	**0.146**	**4.44**	0.918	0.924	**2.12**	2.29
	Full	0.733	**1.17**	8.87	**1.54**	2.9	0.285	4.72	**0.843**	0.925	2.46	2.3
	Pars	**0.734**	1.18	8.87	1.56	2.96	0.217	4.95	0.85	**0.924**	2.34	**2.24**
Early	Base	0.455	0.534	0.528	1.08	0.621	**0.114**	0.526	0.528	1.47	**0.263**	0.292
	Full	**0.427**	0.436	**0.411**	**0.982**	**0.544**	0.124	0.422	**0.457**	**1.34**	0.337	0.255
	Pars	0.429	**0.431**	0.415	0.985	0.56	0.115	**0.415**	0.477	**1.34**	0.338	**0.242**
Late	Base	1.16	2.15	**17.2**	2.42	**5.19**	**0.177**	**8.41**	1.3	**0.384**	**4.02**	4.31
	Full	**1.03**	**1.92**	17.5	**2.09**	5.3	0.444	9.08	**1.22**	0.517	4.61	4.37
	Pars	1.04	1.93	17.5	2.13	5.39	0.318	9.55	**1.22**	0.517	4.38	**4.28**
All values are to the 10^-8 power												

Table 8: Out-of-sample Mean Squared Errors 5 Days Ahead												
		ABT	BAC	BK	BMY	C	JNJ	JPM	MRK	PFE	USB	WFC
Full	Base	0.535	0.936	5.32	0.882	**1.66**	**0.129**	**2.33**	0.56	0.487	1.53	1.42
	Full	**0.473**	0.909	**5.19**	**0.746**	1.67	0.183	2.26	**0.435**	**0.452**	1.52	1.4
	Pars	0.501	**0.905**	5.21	0.748	1.75	0.182	2.28	0.463	0.499	**1.48**	**1.37**
Early	Base	0.302	0.361	0.518	0.428	0.596	0.124	0.723	0.365	0.638	**0.29**	0.199
	Full	**0.281**	0.305	**0.457**	**0.322**	**0.382**	**0.101**	0.417	**0.233**	**0.553**	0.312	0.174
	Pars	0.293	**0.273**	0.465	0.334	0.455	0.102	**0.342**	0.288	0.564	0.308	**0.156**
Late	Base	0.766	1.52	10.2	1.33	**2.73**	**0.135**	**3.96**	0.753	**0.339**	2.8	2.65
	Full	**0.661**	**1.52**	9.99	**1.16**	2.98	0.264	4.14	**0.634**	0.352	2.75	2.64
	Pars	0.705	1.55	**10**	**1.16**	3.06	0.261	4.24	0.635	0.436	**2.67**	**2.6**
All values are to the 10^-8 power												

Table 9: Out-of-sample Mean Squared Errors 22 Days Ahead												
		ABT	BAC	BK	BMY	C	JNJ	JPM	MRK	PFE	USB	WFC
Full	Base	0.588	1.37	6.19	0.976	**2.84**	0.252	3.21	0.616	0.579	2.04	2.1
	Full	**0.45**	1.37	**5.86**	**0.769**	2.98	0.168	**3**	**0.457**	**0.445**	**1.94**	2.06
	Pars	0.572	**1.35**	5.95	0.873	2.89	0.264	3.04	0.512	0.601	1.96	**2.06**
	Red. 22	0.454	**1.35**	5.88	0.772	3.05	**0.156**	3.07	0.465	**0.445**	1.96	**2.05**
Early	Base	0.338	0.452	1.01	0.421	0.878	0.252	1.27	0.457	0.669	0.561	0.188
	Full	**0.279**	0.416	0.93	0.304	**0.399**	**0.155**	0.804	**0.191**	**0.577**	0.603	0.186
	Pars	0.336	**0.371**	0.947	0.323	0.651	0.198	**0.761**	0.335	0.64	**0.548**	**0.159**
	Red. 22	0.277	**0.408**	**0.927**	**0.302**	0.407	**0.155**	0.816	**0.191**	0.581	0.612	0.189
Late	Base	0.834	**2.3**	11.5	1.52	**4.83**	0.252	**5.18**	0.773	0.491	**3.55**	4.04
	Full	**0.617**	2.34	**10.9**	**1.23**	5.6	0.182	5.23	**0.718**	0.316	**3.29**	**3.95**
	Pars	0.805	2.35	11	1.41	5.17	0.328	5.36	**0.687**	0.563	3.39	4
	Red. 22	0.627	**2.3**	**10.9**	1.24	5.74	**0.157**	5.37	**0.735**	**0.312**	3.34	**3.95**
All values are to the 10^-8 power												

Table 10: F-test P-value for Sector Weekly, Monthly and Market Weekly					
	Days			Days	
ABT	1	0.8995	JPM	1	0.0117
	5	0.7719		5	0.3126
	22	0.7116		22	0.5838
BAC	1	0.7636	MRK	1	0.072
	5	0.796		5	0.0055
	22	0.3077		22	0.0005
BK	1	0.8014	PFE	1	0.6807
	5	0.693		5	0.3267
	22	0.5433		22	0.0128
BMY	1	0.1968	USB	1	0.2501
	5	0.4718		5	0.8061
	22	0.2776		22	0.5671
C	1	0.074	WFC	1	0.1319
	5	0.2568		5	0.7033
	22	0.0408		22	0.8724
JNJ	1	0.0205			
	5	0.0295			
	22	0.0767			

Tables 11 – 12 correspond to the *Parsimonious* model and equation 16

Table 11: Coefficient Estimates and Significance Levels Day-Ahead (Parsimonious Model)

	ABT	BMY	JNJ	MRK	PFE
βCD	0.1829***	0.0464	0.1152*	0.1227*	0.1183**
βCW	0.2020***	0.1003	0.1958*	0.0648	0.1173*
βCM	0.3689***	0.4219***	0.1914**	0.3268***	0.3933***
βSD	0.1968*	0.5258***	0.3681***	0.2794***	0.2573***
βSPDiff	0.2793	0.2019*	0.2165*	0.2742*	0.4282***
β0	0.0000194*	-0.000002	-0.000005	0.00005***	0.00004***
* p<0.05, **p<0.01, ***p<0.001					

Table 12: Coefficient Estimates and Significance Levels 5 Days Ahead (Parsimonious Model)

	ABT	BMY	JNJ	MRK	PFE
βCD	0.0845**	-0.0367	0.0312	-0.0081	-0.0159
βCW	0.2629***	0.2266***	0.1620	0.1187	0.2066**
βCM	0.4340***	0.4798***	0.3238**	0.3225***	0.4038***
βSD	0.1040*	0.3055***	0.2852***	0.2750***	0.2209**
βSPDiff	0.1773*	0.2615**	0.2697*	0.1178*	0.2710*
β0	0.000036**	0.000025**	0.000015	0.00007***	0.00006***
* p<0.05, **p<0.01, ***p<0.001					

Table 13 corresponds to the *Reduced Monthly* model and equation 17

Table 13: Coefficient Estimates and Significance Levels 22 Days Ahead (Reduced Monthly Model)

	ABT	BMY	JNJ	MRK	PFE
βCD	0.0473**	-0.0367	-0.0107	0.0019102	0.0137
βCW	0.1520***	0.2189***	0.1132*	0.0854	0.1490**
βCM	0.7654***	0.4259*	0.2889	-0.2257	0.2902*
βSD	0.0997**	0.2836***	0.2204***	0.1529**	0.1366**
βSM	0.1556	0.3354	0.3276	0.8596**	0.5318**
βSPM	-1.053***	-0.8973**	-0.6170**	-0.7440*	-0.9651***
β0	0.000053*	0.000056*	0.000033*	0.00009***	0.00008***
* p<0.05, **p<0.01, ***p<0.001					

Table 14: Pharmaceutical Jump Days (1997 - 2007)					
	ABT	BMY	JNJ	MRK	PFE
Mean RV ($x10^{-4}$)	2.8404	3.3421	1.8252	2.429	2.8078
Mean BV ($x10^{-4}$)	2.603	2.9993	1.6801	2.232	2.6035
Number of Jump Days	110	137	114	86	95

8. Figures

Figure 1: JP Morgan (JPM) Daily Realized Variance July 98 – Jan 08

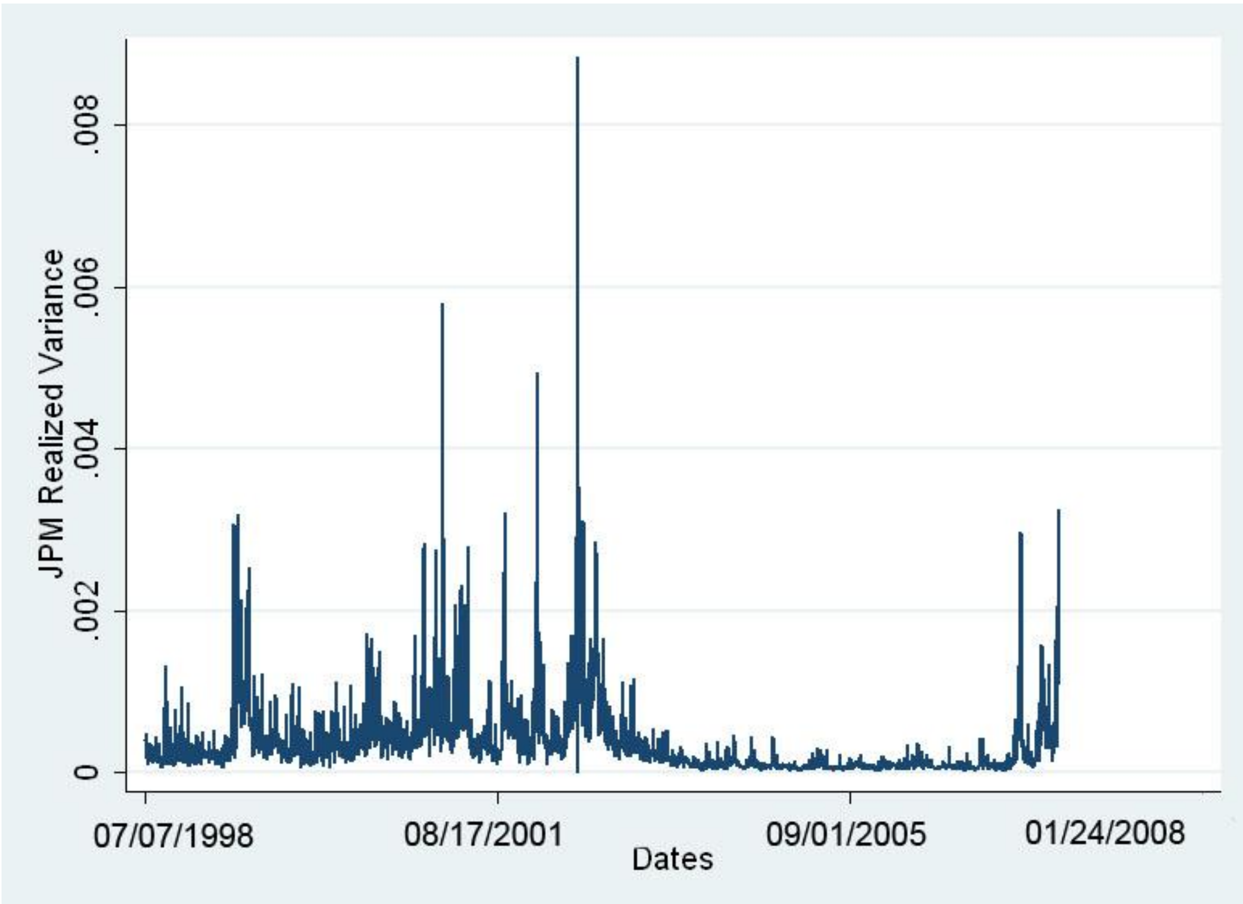
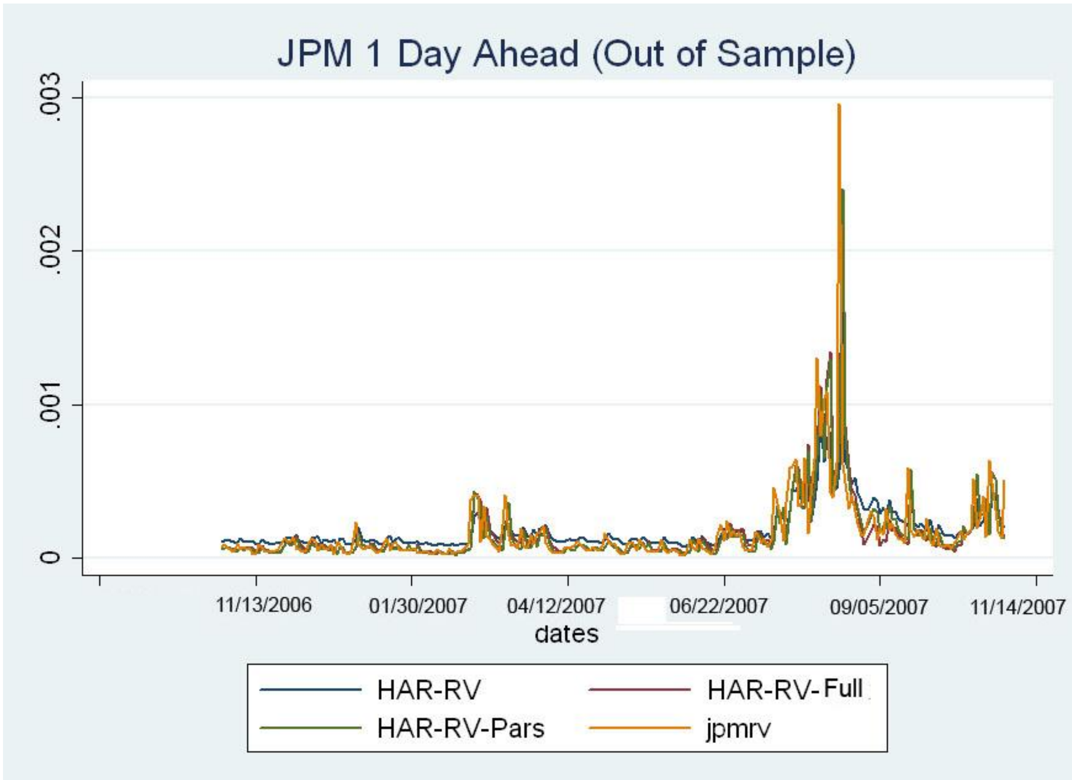


Figure 2: JPM Day Ahead Out-of-sample Forecasts and Residuals



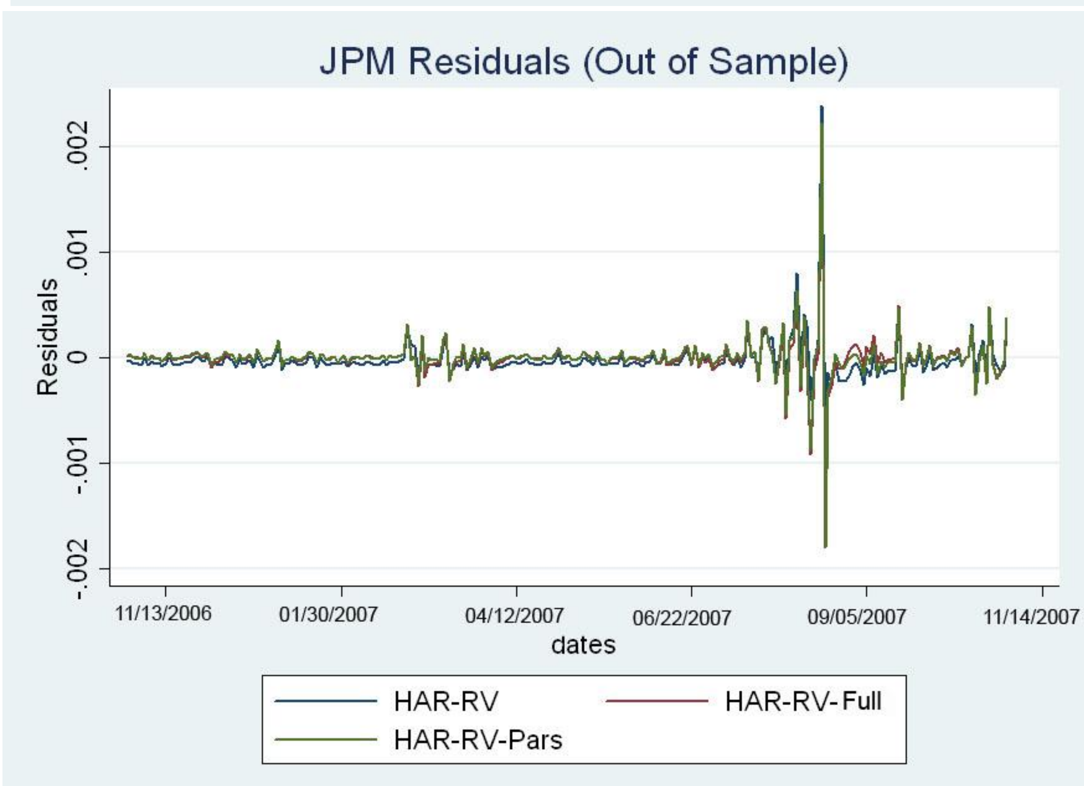


Figure 3: JPM 5 Days Ahead Out-of-sample Forecasts and Residuals

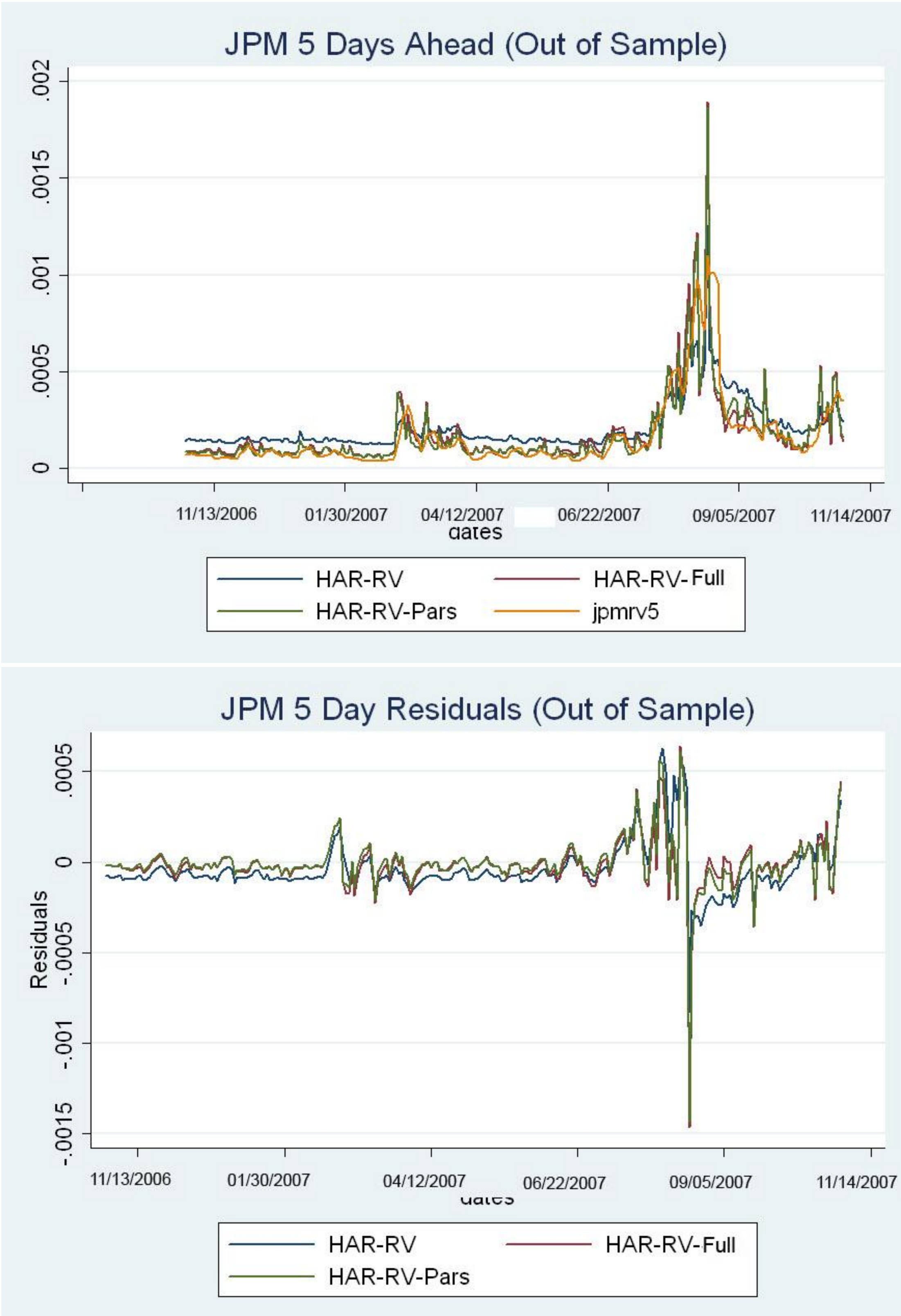
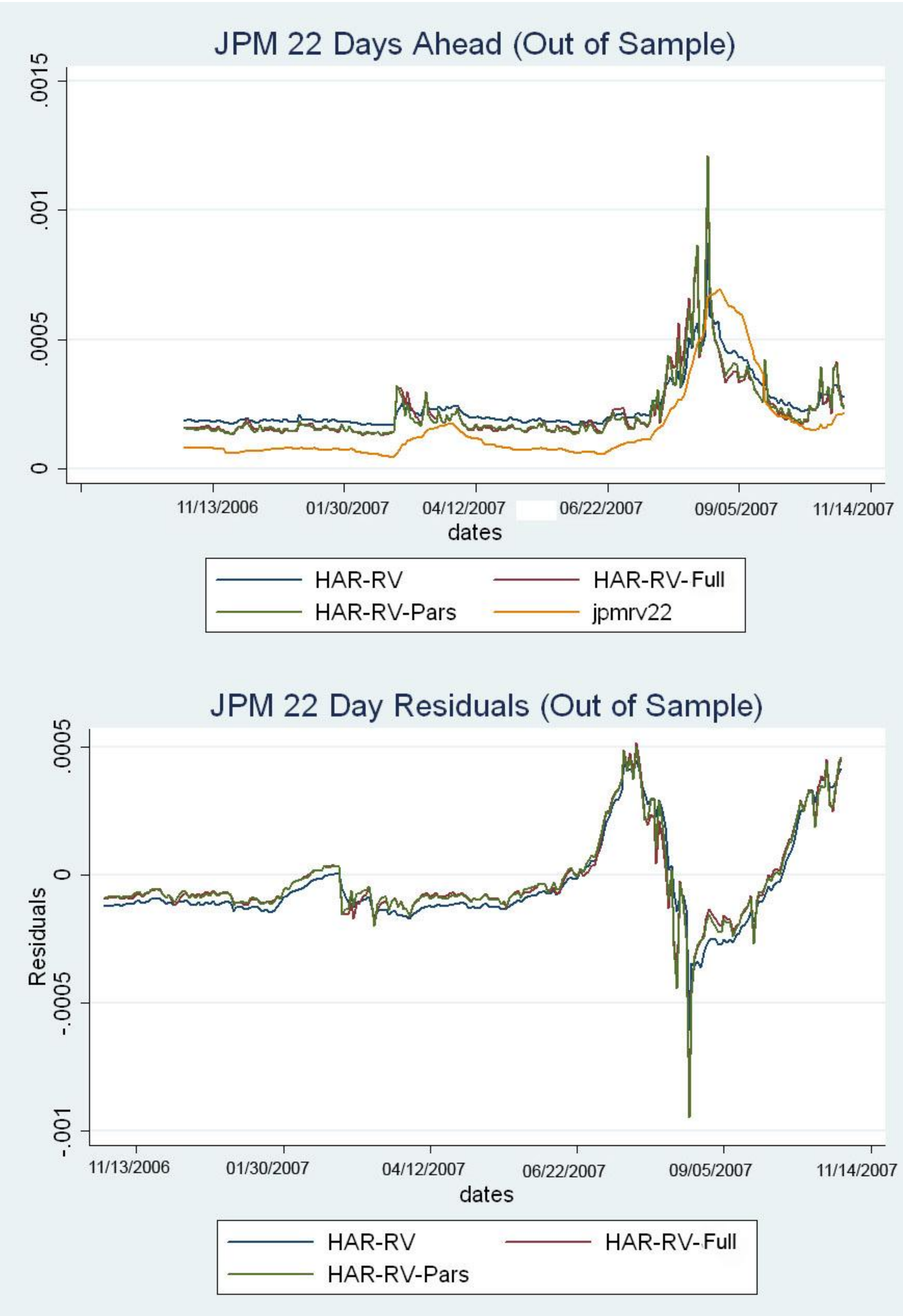


Figure 4: JPM 22 Days Ahead Out-of-sample Forecasts and Residuals



References

Andersen, T.G. and T. Bollerslev (1998). "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review,* 39, 885-905.

Andersen, T. G., Bollerslev T., Diebold F. X., and P. Labys (2003). "Modeling and forecasting realized volatility." *Econometrica* 71, 579-625.

Andersen, T. G., Bollerslev T., F. X. Diebold (2006). "Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility."

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics*, 31, 307-327.

Brandorff-Nielsen, O. and N. Shephard. (2004). "Power and Bipower Variation with Stochastic Volatility and Jumps." *Journal of Financial Econometrics* 2, 1-37.

Corsi, F. (2003). "A simple long memory model of realized volatility". Unpublished manuscript, University of Southern Switzerland.

Fama, E. F., and K. R. French (2003). "The Capital Asset Pricing Model: Theory and Evidence." *Tuck Business School Working Paper,* 03-26.

Fradkin, A. (2007). "The Informational Content of Implied Volatility in Individual Stocks and the Market". Unpublished manuscript, Duke University.

Ghysels, E., A. Harvey and E. Renault (1996). "Stochastic Volatility." in G.S. Maddala (ed.), *Handbook of Statistics Vol. 14, Statistical Methods in Finance*, 119-191.

Granger, C.W.J., and P. Newbold (1974). "Spurious regressions in econometrics." *Journal of Econometrics,* 2, 111-120, 1974.

Harvey, A. C. (1980). "On Comparing Regression Models in Levels and First Differences." *International Econometrics Review* 21, 707-720.

Huang, X., and G. Tauchen. (2005). "The Relative Contribution of Jumps to Total Price Variation." *Journal of Financial Econometrics* 3, 456-499.

Merton, R. C. (1976). "Option Pricing When Underlying Stock Returns are Discontinuous." *Journal of Economic Theory* 3, 125-144.

Petersen, M. A. (2008). "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches." *Review of Financial Studies*, forthcoming.