Forecasting Stock Market Volatilities Using MIDAS Regressions: An Application to the Emerging Markets

C. Emre Alper and Salih Fendoglu and Burak Saltoglu

Bogazici University, Department of Economics, Bogazici University, Center for Economics and Econometrics, University of Maryland, College Park, Department of Economics

March 2008

Forecasting Stock Market Volatilities Using MIDAS Regressions: An Application to the Emerging Markets

C. Emre Alper, Salih Fendoglu, Burak Saltoglu *

March 5, 2008

Abstract

We explore the relative weekly stock market volatility forecasting performance of the linear univariate MIDAS regression model based on squared daily returns *vis-á-vis* the benchmark model of GARCH(1,1) for a set of four developed and ten emerging market economies. We first estimate the two models for the 2002-2007 period and compare their in-sample properties. Next we estimate the two models using the data on 2002-2005 period and then compare their out-of-sample forecasting performance for the 2006-2007 period, based on the corresponding mean squared prediction errors following the testing procedure suggested by West (2006). Our findings show that the MIDAS squared daily return regression model outperforms the GARCH model significantly in four of the emerging markets. Moreover, the GARCH model fails to outperform the MIDAS regression model in any of the emerging markets significantly. The results are slightly less conclusive for the developed economies. These results may imply superior performance of MIDAS in relatively more volatile environments.

Keywords. Mixed Data Sampling regression model; Conditional volatility forecasting; Emerging Markets.
JEL No. C22; C53; G12.

*Corresponding author. Tel: +90 (212) 359-7648. Fax: + 90 (212) 287-2453. E-mail: burak.saltoglu@boun.edu.tr.

1 Introduction

There has been a voluminous literature on volatility analysis since the seminal papers by Engle (1982) and Bollerslev (1986). The different directions into which this research developed are investigated and portrayed in Granger and Poon (2003). In addition, the seminal paper by Andersen *et al.* (2003) on realized volatility has opened a new research avenue in the field of financial econometrics. This new approach provides a framework for integration of high-frequency intraday data into forecasting of daily and lower frequency return volatilities and return distributions. This method proved to be a more satisfactory method compared to more conventional GARCH type methods for modeling high frequency return data.

In a similar context, the Mixed Data Sampling (MIDAS) regression models are introduced by the seminal papers of Ghysels *et al.* (2004, 2005, 2006a,b). MIDAS regressions allow one to study parsimoniously parameterized regressions using data sampled at different frequencies. The major appealing feature of the MIDAS method is that it offers a more general analytical framework for not only daily data but also weekly, monthly or even quarterly financial and macroeconomic data. Consequently, the MIDAS methodology has welcomed considerable attention in recent years. For instance, Chen *et al.* (2007) extend the MIDAS setting to a multi-horizon semi-parametric framework while Ghysels *et al.* (2007a) analyze the U.S. commercial real estate market within the MIDAS context. Clements and Galvao (2006) study forecasts of the U.S. output growth and inflation in this context. Hogrefe (2007) employs a study on data revisions of GDP within a mixed frequency sampling approach. Finally, Kotze (2007) uses MIDAS regressions for inflation forecasting with high frequency asset price data.

Many of the previous empirical studies based on MIDAS regressions and realized volatility used U.S. equity return data and there seems very few studies applied on emerging market equity return data. Given the global integration of international financial markets, and different nature of the emerging markets, it is a natural question to ask how these recent models fare in these countries and under different frequencies.

In this paper, we explore the relative forecasting performance of MIDAS regressions based on squared daily returns *vis-á-vis* the benchmark model of weekly GARCH(1,1) using equity return data for a set of four developed and ten emerging markets. In essence, using MIDAS regressions, we explore whether individual daily return volatilities contribute significantly to predicting the following week's return volatility and whether this prediction is better than the GARCH(1,1) model which uses the current week's volatility to forecast the following week's volatility. We conduct out-of-sample forecasting with a recursive scheme for both models, and evaluate them by comparing the corresponding Mean Squared Prediction Errors (MSPEs) following the procedure proposed by West (2006). Our motivation is that given the heterogeneity among and within these markets, such an analysis might provide valuable insights in assessing the MIDAS methodology. To our best knowledge, this paper is a first attempt in exploring the relative volatility forecasting performance of MIDAS regression model using stock return data from both the emerging and developed markets. We utilize weekly data for two main reasons. First, intradaily stock data for emerging markets was simply unavailable. Second, as we focus on weekly return series, we provide additional evidence on how MIDAS regression model fares under relatively less frequent samples. Our findings suggest for the emerging market economies that MIDAS model by and large produce more precise forecasting performance than that of GARCH(1,1) benchmark. For developed economies we obtain a less decisive picture.

The rest of the paper proceeds as follows. Section 2 describes the methodology. Section 3 provides the data diagnostics and the empirical results, and Section 4 concludes.

2 Methodology

The univariate MIDAS regression model can be represented with the following econometric specification:

$$Y_t = \alpha_0 + \alpha_1 \sum_{k=0}^{k^{max}} B(k,\theta) X^{(m)}_{t-k/m} + \varepsilon_t \tag{1}$$

where $B(k,\theta)$ is a polynomial weighting function depending on both the elapsed time k and the parameter vector θ, and $X^{(m)}_t$ is sampled m times faster than Y_t.

The main appealing aspect of using MIDAS regression model lies in the above specification. In particular, MIDAS not only parameterizes the polynomial $B(k,\theta)$ in a parsimonious and flexible manner but also uses data sampled at different frequencies and hence offers a gain in efficiency by exploiting information hidden in the higher frequency data.[1]

In our methodology, we set $m = 5$, $k^{max} = 5$, and t denoting weekly sampling, which implies a projection of weekly Y_t series on daily $X^{(m)}_t$ data. We primarily focus on forecasting one-week-ahead realized volatility using previous week's individual squared daily returns. We define our returns as

$r^{(m)}_{t,t-1/m} = [\log(P^{(m)}_t) - \log(P^{(m)}_{t-1/m})] \times 100$

[1] Ghysels *et al.* (2004, 2005, 2006a,b) discuss extensions of the given univariate MIDAS model to a non-linear and/or a multivariate setting, the asymptotic properties of MIDAS models in general, their advantages over distributed lag models.

where $P_t^{(m)}$ refers to daily closing value of the stock market and $P_{t-j/m}^{(m)}$ denotes the closing value of the stock market for the $5-j^{th}$ day of the week.

We use the following MIDAS specification:

$$V_{t+1,t} = \alpha_0 + \alpha_1 \sum_{k=0}^{k^{max}} B(k,\theta)[r_{t-k/m}^{(m)}]^2 + \varepsilon_t \tag{2}$$

where t denotes weekly sampling, k^{max}, θ and m are defined as above.[2] In equation (2), $r_{t-k/m}^{(m)}$ is the lag of daily stock returns and $V_{t+1,t}$ is the conditional volatility that we wish to predict. Accordingly, equation (2) specifies how the previous 5 individual daily squared returns should to be weighted in predicting the following week's realized volatility (which is constructed based on non-overlapping 5 days). There are various alternatives for the polynomial specification but throughout this paper, by following Ghysels *et al.* (2006b), we use Beta lag polynomial.[3] In particular, the Beta polynomial can be specified as the following:

$$B(k,\theta) = \frac{f(k/k^{max};\theta_1,\theta_2)}{\sum_{k=1}^{k^{max}} f(k/k^{max};\theta_1,\theta_2)} \tag{3}$$

with

$$f(x,\theta_1,\theta_2) = \frac{x^{\theta_1-1}(1-x)^{\theta_2-1}\Gamma(\theta_1+\theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)}$$

where $\Gamma(.)$ is the conventional Gamma function. Moreover, we let weights be normalized to add up to one so that α_1 in equation (2) can be identified and estimate the MIDAS parameters through non-linear least squares, among other procedures. We leave further features of MIDAS regression models to Ghysels *et al.* (2004, 2005, 2006a,b) and proceed to how we evaluate the MIDAS and the GARCH models in our out-of-sample experiment.

2.1 Forecast Evaluation

A simple criterion to compare out-of-sample forecast accuracies of competing models is to choose the model that provides a smaller mean, mean-absolute or mean-squared prediction error, among other measures. Improving upon this naive criterion, Diebold and Mariano (1995), West (1996), and

[2] We could have used absolute returns or daily range as well, among many other volatility measures. However, since our benchmark is the GARCH(1,1) model, it would be more convenient to use the squared returns for comparison purposes.

[3] Essentially, compared to other polynomial specifications, the Beta lag polynomial appears to be well-performing in predicting following week's realized volatility in our in-sample forecast experiment.

McCracken (2000) question the adequacy of such a procedure, and provide a formal test of equal forecast accuracy between (non-nested) models for a wide variety of loss measures. Readers may refer to McCracken (2004) and West (2006) for surveys on this literature. In this paper, we follow West (2006) who suggests an approach for comparing forecast accuracies of non-nested models. We start by introducing some notation.

Suppose we have two competing models. Let $e_{1,t}$ and $e_{2,t}$ denote the population forecast errors under the first and the second model, $L(.)$ is the loss function used to evaluate the forecast accuracy, and $f_t = L(e_{1,t}) - L(e_{2,t})$. Assume for simplicity that $L(.)$ denotes MSPE such that $L(e_t) = e_t^2$. Moreover, let $\sigma_i^2 = E[e_{i,t}^2]$ denote the population MSPEs for model $i = 1$ or 2. Then denoting the sample counterpart of the variables with a "$\widehat{\ }$", we have $\widehat{f}_t = \widehat{e}_{1,t}^2 - \widehat{e}_{2,t}^2$.

Diebold and Mariano (1995) propose a simple Wald test for the hypothesis $E[f_t] \equiv \sigma_1^2 - \sigma_2^2 \leq 0$, that is, regressing $\widehat{f}_t$ on a constant and comparing the resulting t-statistic with standard normal critical values. We apply the method by West (2006) which suggests using heteroskedasticity and autocorrelation consistent (HAC) t-statistics to solve the inconsistency problem posed in Diebold and Mariano test. Henceforth, we let the first model be the GARCH and the second model be the MIDAS. We use MSPE as the loss function $L(e_t) = e_t^2$ for both models, and test the hypothesis of equal MSPEs, i.e. $\sigma_1^2 - \sigma_2^2 = 0$ by following West (2006). We present our empirical findings in the next section.

3 Data and Empirical Results

Our data set consists of weekly stock returns of four developed and ten emerging market economies. In particular, we study S&P500 (the U.S.), FTSE (the U.K.), DAX (Germany), and NIKKEI (Japan) among developed economies; and BSE30 (India), HSI (Hong Kong), IBOVESPA (Brazil), IPC (Mexico), JKSE (Indonesia), KLSE (Malaysia), KS11 (South Korea), MERVAL (Argentina), STI (Singapore), and TWII (Taiwan) among emerging market economies. The market indices for each stock market are daily closing values for the period between January 7, 2002 (Monday) and December 21, 2007 (Friday) with a number of week-days totaling 1555 except for MERVAL which begins by January 21, 2002 (Monday) and has 1545 daily observations. We choose that specific sample period in order to remove the effects of September 11 and some other financial crises occurred in some of the countries. The stock market indices are taken from Bloomberg. The 'missing' observations due to fixed or moving holidays are replaced by the most recent available observation to achieve uninterrupted series of observations. For our out-of-sample forecast experiment, we first using the initial 1005 daily or equivalently 201 weekly observations, roughly the 2002-2005 period, and predict the remaining 110 weekly realized volatilities, roughly 2006-2007, on a one-step-ahead basis.[4]

The diagnostics for daily/weekly return data are provided in Tables 1 and 3. We also provide average values of the descriptive statistics for each group of countries in Tables 2 and 4. We observe that emerging stock markets provide higher returns as well as higher volatility and higher negative skewness. Daily returns for both groups of stock markets exhibit similar leptokurtic behavior and persistence at 5 lags. For daily squared returns, we observe a much higher serial correlation at 5 lags for developed economies. The same can also be drawn for weekly (squared) returns. Weekly EM stock returns also exhibit higher returns, higher volatility and less return correlation.

We present in-sample MIDAS regression diagnostics in Table 5. It can be seen that MIDAS regression coefficients $\hat{\alpha}_0$ and $\hat{\alpha}_1$ appear to be positive and significant for all countries. This implies that, as expected, daily squared returns contribute positively to the following week's realized volatility. Moreover, $\hat{\theta}_1$ being close to 1 and $\hat{\theta}_2 > 1$ for most of the countries imply that the weights are, in general, decaying gradually (see also Table 5, columns 5 through 8). Residuals obtained from most emerging market countries exhibit no serial correlation implying a relatively satisfactory model specification. However, MIDAS residuals obtained from developed countries, on average, exhibit more serial correlation.[5] In order to further investigate these series we will conduct an out-of-sample forecasting exercise next.

Table 6 presents our essential empirical results. Using Mean Squared Prediction Error (MSPE) as a forecast accuracy criterion, we find that the MIDAS model outperforms the GARCH specification for all emerging markets but one (MERVAL). For the developed countries, MIDAS still outperforms but in a less decisive way. It performs better than the GARCH(1,1) model for S&P500, FTSE, and NIKKEI; and worse for DAX.

The last column of Table 6 presents the test statistics obtained for West(2006) forecast accuracy test. A positive and large test statistics imply that the MIDAS regression model forecasts outperform the benchmark GARCH(1,1) model. As can be seen, for four out of ten emerging market stocks, the forecasts obtained from the MIDAS method proved to be statistically more accurate than that of the GARCH(1,1) method. One can also observe that the emerging stock markets where the MIDAS fared better, receive considerable more weighting for the most recent week. This implies that, given (k^{max}), higher frequency data indeed embed valuable information about the lower frequency data over some future horizon which is one of the main properties of the MIDAS methodology as opposed to GARCH(1,1). One can also notice from Table 6 that the forecasting performance of GARCH(1,1)

[4]For MERVAL, we predict 108 weekly realized volatilities.

[5]We note that the squared MIDAS residuals which resemble the fourth moment of conventional disturbances shows no serial correlation for two out of four developed economies, and three out of ten emerging markets.

method benchmark never outperformed the MIDAS forecasts significantly for the emerging markets. This may further imply that the MIDAS volatility forecasting framework may have more appealing features for the weekly emerging market countries stock data. Either the more volatile structure or rapidly changing volatility dynamics may lead us to obtain these relatively improved forecasting results.

However, we obtained relatively less decisive results for the forecasting performance of developed countries stock data. As can be seen from the Table 6, on the basis of MSPE statistics, except for DAX, MIDAS has a relatively better performance than the benchmark. If we compare these two models on the basis of West (2006), MIDAS the model has no clear clear advantage over GARCH model. While for NIKKEI, MIDAS produce statistically better forecast accuracy, for DAX, GARCH produce a better forecasting precision. For the other two countries, S&P500 and FTSE, we do not observe any clear winners in terms of forecast performance.

As a general conclusion, the forecasting accuracy difference between emerging markets and developed economies may be attributable to the relatively more volatile nature of the former.

4 Conclusions

In this paper, we evaluate the forecasting performance of a linear univariate MIDAS regression model based on squared daily returns compared to the benchmark model of GARCH (1,1) for equity return volatilities of ten emerging markets and four developed economies. More concretely, for both sets of stock markets, we investigate to what extent individual daily volatilities of the very recent week convey significant information about next week's realized volatility. Given the heterogeneity among and within these stock markets, we aim to unravel some novel features of the MIDAS regressions empirically and question how MIDAS performs under relatively less frequent samples.

We conclude that for the emerging stock markets, which are relatively more volatile markets, the MIDAS model appears to be a better forecasting model whereas for the less volatile developed economies' stock markets we do not have clear-cut results. One explanation why MIDAS works better is that it optimally weights the recent return uncertainty. Therefore, making use of the higher frequency data helps predicting future volatility structure under more volatile economic environments. Studying other MIDAS specifications with different sampling frequencies are left for future research.

Acknowledgements

The authors would like to thank Eric Gyhsels for introducing the topic and Oya Ardic, Tae Hwy Lee and Arthur Sinko for useful discussions. Alper acknowledges financial support from TUBA-GEBIP (Turkish Academy of Sciences - Young Scientists Scholarship Program). The usual disclaimer applies.

References

[1] Andersen, T., Bollerslev, T., Diebold, F.X. and Labys, P. (2003), Modeling and Forecasting Realized Volatility, Econometrica, 71, 529-626.

[2] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31:307-327.

[3] Chen, X., and Ghysels, E. (2007). News -Good or Bad- and Its Impact on Multiple Horizons. UNC-Chapel Hill, Working Paper.

[4] Clements, M. P., and Galvao, A. B. (2006). Macroeconomic Forecasting with Mixed Frequency Data: Forecasting US output growth and inflation. Warwick Economic Research Paper No. 773, University of Warwick.

[5] Diebold, F. and R. Mariano (1995). Comparing Predictive Accuracy, *Journal of Business and Economic Statistics* 13(3): 253-263.

[6] Ghysels, E., Santa-Clara, P., and Valkanov, R. (2004). The MIDAS Touch: Mixed Data Sampling Regression Models. UNC and UCLA Discussion Paper.

[7] Ghysels, E., Santa-Clara, P., and Valkanov, R. (2005). There is a risk-return tradeoff after all. *Journal of Financial Economics* 76: 509-548.

[8] Ghysels, E., Santa-Clara, P., and Valkanov, R. (2006a). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics* 131, 59-95.

[9] Ghysels, E., Sinko, A, and Valkonov, R. (2006b). MIDAS Regressions: Further Results and New Directions. Mimeo, University of North Carolina, Chapel Hill, NC.

[10] Ghysels, E., Plazzi, A., and Valkanov, R. (2007). Valuation in the US Commercial Real Estate. UNC-Chapel Hill, Working Paper.

[11] Granger, C. W. J, and Poon, S.-H. (2003). Forecasting Volatility in Financial Markets: A Review. Journal of Economic Literature 41(2): 478-539.

[12] Hogrefe, J. (2007). Forecasting data revisions of GDP : a mixed frequency approach. Christian-Albrechts-Universitaet Kiel, Working Paper.

[13] Kotze, G.L. (2007). Forecasting Inflation with High Frequency Asset Price Data. University of Stellenboch, Working Paper.

[14] McCracken, M. W. (2000). Robust out-of-sample inference. *Journal of Econometrics* 99, 195–223.

[15] McCracken, M. W. (2004). Parameter estimation and tests of equal forecast accuracy between non-nested models. *International Journal of Forecasting* 20: 503-514.

[16] West, K. (1996). Asymptotic Inference about Predictive Ability, *Econometrica*, 64(5): 1067-1084.

[17] West, K. (2006). Forecast Evaluation. In *Handbook of Economic Forecasting*, V. 1 (Eds. Elliot, G., Granger, C. W. J., and Timmermann, A.), pp. 99-134.

Table 1: Descriptive Statistics for the daily return data (January 7 (Monday), 2002 - December 21 (Friday), 2007)

	S&P500		FTSE100		DAX		NIKKEI225	
	daily return	daily squared return	daily return	daily squared return	daily return	daily squared return	daily return	daily squared return
Obs.	1555	1555	1555	1555	1555	1555	1555	1555
Mean	0.015	0.993	0.011	1.145	0.026	2.239	0.020	1.531
Variance	0.992	4.973	1.145	8.486	2.238	28.473	1.530	7.717
Skewness	0.125	5.764	-0.207	5.768	-0.019	5.072	-0.236	4.503
Kurtosis	6.037	52.342	7.477	45.162	6.676	35.347	4.305	34.967
Q(5)	11.734	581.980	42.344	758.250	14.777	749.240	3.446	68.948
	BSE30		**HSI**		**IBOVESPA**		**IPC**	
	daily return	daily squared return	daily return	daily squared return	daily return	daily squared return	daily return	daily squared return
Obs.	1555	1555	1555	1555	1555	1555	1555	1555
Mean	0.112	1.834	0.053	1.259	0.095	2.809	0.096	1.407
Variance	1.822	27.447	1.257	9.193	2.802	22.052	1.397	8.509
Skewness	-0.628	14.472	-0.246	10.410	-0.290	3.633	-0.156	5.919
Kurtosis	9.446	329.648	6.856	199.267	3.864	22.270	5.377	56.568
Q(5)	21.322	406.630	18.415	178.590	4.707	63.042	10.583	160.130
	JKSE		**KLSE**		**KS11**		**MERVAL**	
	daily return	daily squared return	daily return	daily squared return	daily return	daily squared return	daily return	daily squared return
Obs.	1555	1555	1555	1555	1555	1555	1545	1545
Mean	0.124	1.732	0.045	0.528	0.058	2.135	0.107	3.679
Variance	1.716	24.020	0.526	1.661	2.131	18.836	3.671	86.546
Skewness	-0.798	12.524	-0.254	7.813	-0.378	5.514	-0.253	7.622
Kurtosis	9.416	245.735	7.051	96.553	5.196	49.961	7.471	89.727
Q(5)	22.400	123.870	46.224	349.510	9.844	133.910	3.580	135.260
	STI		**TWII**					
	daily return	daily squared return	daily return	daily squared return				
Obs.	1555	1555	1555	1555				
Mean	0.044	0.995	0.019	1.690				
Variance	0.994	4.601	1.690	13.675				
Skewness	-0.107	6.072	-0.274	4.991				
Kurtosis	5.660	64.130	5.799	38.205				
Q(5)	8.001	121.400	10.194	138.350				

Notes: MERVAL (Argentina) is dated from January 21 (Monday), 2002. The missing market indices (due to existence of fixed/moving holidays) are substituted by the very last available index. Q(5) denotes the corresponding Ljung-Box (1979) Q-statistic for five lags.

Table 2: Group-wise Descriptive Statistics of Stock Market -Daily Data-

	Daily Return		Daily Squared Return	
	Developed M.	Emerging M.	Developed M.	Emerging M.
Mean	0.018	0.075	1.477	1.918
Variance	1.476	1.801	12.412	14.482[a]
Skewness	-0.084	-0.338	5.277	6.557[b]
Kurtosis	6.124	6.614	41.955	58.881[c]
Q(5)	18.075	15.527	539.605	134.880[d]

Notes: Developed countries' equity markets consist of S&P500 (the U.S.), FTSE (the U.K.), DAX (Germany), and NIKKEI (Japan). Emerging equity markets consist of BSE30 (India), HSI (Hong Kong), IBOVESPA (Brazil), IPC (Mexico), JKSE (Indonesia), KLSE (Malaysia), KS11 (South Korea), MERVAL (Argentina), STI (Singapore), and TWII (Taiwan). The table values are calculated simply by averaging the corresponding statistics for each group of countries. Q(5) denotes the corresponding Ljung-Box (1979) Q-statistic for five lags.

[a] averaging all but the outlier MERVAL. Including the outlier yields an average variance of 23.489.
[b] averaging all but the outlier BSE30. Including the outlier yields an average skewness of 7.546.
[c] averaging all but the outliers BSE30, HSI and JKSE. Including the outliers yields an average kurtosis of 110.275.
[d] averaging all but the outliers BSE30 and KLSE. Including the outliers yields an average Q(5) of 195.677.

Table 3: Descriptive Statistics for the weekly return data (January 7 (Monday), 2002 - December 21 (Friday), 2007)

	S&P500		FTSE100		DAX		NIKKEI225	
	weekly return	weekly squared return	weekly return	weekly squared return	weekly return	weekly squared return	weekly return	weekly squared return
Obs.	311	311	311	311	311	311	311	311
Mean	1.924	4.964	1.973	5.725	2.776	11.199	2.524	7.657
Variance	1.265	46.145	1.838	91.968	3.504	299.843	1.288	48.094
Skewness	1.687	3.469	2.025	4.051	1.765	3.168	0.736	2.020
Kurtosis	6.573	18.353	8.066	23.933	6.170	14.949	3.635	8.486
Q(1)	155.360	177.230	188.260	153.560	207.810	187.600	70.991	55.909
	BSE30		HSI		IBOVESPA		IPC	
	weekly return	weekly squared return	weekly return	weekly squared return	weekly return	weekly squared return	weekly return	weekly squared return
Obs.	311	311	311	311	311	311	311	311
Mean	2.615	9.173	2.239	6.297	3.484	14.046	2.375	7.034
Variance	2.338	254.402	1.288	66.536	1.913	142.277	1.394	66.324
Skewness	2.766	7.901	1.997	4.935	0.983	2.329	1.566	3.269
Kurtosis	16.962	89.768	9.793	38.404	4.566	10.172	6.338	17.278
Q(1)	79.041	51.686	90.170	114.980	40.576	35.015	69.241	60.711
	JKSE		KLSE		KS11		MERVAL	
	weekly return	weekly squared return	weekly return	weekly squared return	weekly return	weekly squared return	weekly return	weekly squared return
Obs.	311	311	311	311	311	311	309	309
Mean	2.559	8.659	1.414	2.640	2.932	10.675	3.654	18.280
Variance	2.114	175.721	0.640	14.946	2.079	132.342	4.928	724.686
Skewness	2.254	5.690	2.094	4.945	1.224	2.820	1.814	4.330
Kurtosis	12.064	46.175	10.126	35.377	5.107	13.251	8.005	28.714
Q(1)	32.414	26.885	61.691	75.605	75.685	50.540	49.364	30.920
	STI		TWII					
	weekly return	weekly squared return	weekly return	weekly squared return				
Obs.	311	311	311	311				
Mean	1.977	4.978	2.549	8.454				
Variance	1.071	31.990	1.960	97.911				
Skewness	1.200	3.284	1.193	2.600				
Kurtosis	5.132	21.765	4.554	11.533				
Q(1)	72.342	57.608	58.840	44.925				

Notes: MERVAL (Argentina) is dated from January 21 (Monday), 2002. The missing market indices (due to existence of fixed/moving holidays) are substituted by the most recent available index value. Q(1) denotes the corresponding Ljung-Box (1979) Q-statistic for one lag.

Table 4: Group-wise Descriptive Statistics of Stock Market -Weekly Data-

	Weekly Return		Weekly Squared Return	
	Developed M.	Emerging M.	Developed M.	Emerging M.
Mean	2.299	2.580	7.386	9.032
Variance	1.974	1.973	121.513	165.049[a]
Skewness	1.553	1.709	3.177	3.800[b]
Kurtosis	6.111	8.264	16.430	24.741[c]
Q(1)	155.605	62.936	143.574	48.211[d]

Notes: Developed countries' equity markets consist of S&P500 (the U.S.), FTSE (the U.K.), DAX (Germany), and NIKKEI (Japan); and Emerging equity markets consist of BSE30 (India), HSI (Hong Kong), IBOVESPA (Brazil), IPC (Mexico), JKSE (Indonesia), KLSE (Malaysia), KS11 (South Korea), MERVAL (Argentina), STI (Singapore), and TWII (Taiwan). The table values are calculated simply by averaging the corresponding statistics for each group of countries. Q(1) denotes the corresponding Ljung-Box (1979) Q-statistic for one lag.

[a] averaging all but the outliers BSE30 and MERVAL. Including the outliers yields an average variance of 170.713.
[b] averaging all but the outlier BSE30. Including the outlier yields an average skewness of 4.210.
[c] averaging all but the outliers BSE30. Including the outlier yields an average kurtosis of 31.244.
[d] averaging all but the outliers in HSI. Including the outliers yields an average Q(1) of 54.888.

Table 5: In-sample MIDAS Regression Diagnostics

	α_0	α_1	θ_1	θ_2	Week 1	Week 2	Week 3	Week >3	$Q(10)$	$Q^2(10)$	MSPE
S&P500	0.785	4.340	0.991	13.054	0.771	0.181	0.040	0.008	0.036	0.000	16.94
	(2.587)	(22.287)	(137.547)	(4.904)							
FTSE	1.077	4.193	0.982	7.425	0.598	0.237	0.105	0.059	0.491	0.700	43.27
	(2.232)	(17.434)	(105.816)	(4.924)							
DAX	1.011	4.625	0.987	5.996	0.507	0.253	0.134	0.106	0.002	0.000	105.88
	(1.373)	(23.453)	(131.994)	(5.972)							
NIKKEI	2.920	3.018	0.999	5.654	0.453	0.267	0.149	0.129	0.023	0.940	34.11
	(4.435)	(8.098)	(61.576)	(2.762)							
BSE30	4.328	2.546	0.000	5.007	1.000	-	-	-	0.374	0.046	153.24
	(5.473)	(14.401)	(0.000)	(0.379)							
HSI	1.382	4.166	0.989	8.387	0.620	0.241	0.095	0.045	0.000	0.000	40.09
	(2.732)	(13.979)	(94.732)	(3.793)							
IBOVESPA	6.524	2.679	1.008	8.543	0.579	0.268	0.105	0.048	0.917	0.403	113.68
	(5.382)	(7.207)	(69.504)	(3.029)							
IPC	2.398	3.645	0.964	6.438	0.618	0.207	0.102	0.072	0.634	0.001	50.61
	(3.588)	(8.605)	(78.042)	(3.083)							
JKSE	4.784	2.453	1.047	37.797	0.967	0.032	-	-	0.947	0.958	152.87
	(5.165)	(6.123)	(48.371)	(4.538)							
KLSE	1.083	3.147	1.007	21.786	0.893	0.099	0.008	-	0.661	0.000	11.20
	(4.337)	(9.377)	(83.967)	(4.741)							
KS11	3.675	3.282	1.008	5.769	0.440	0.277	0.153	0.130	0.518	0.372	97.59
	(3.851)	(8.967)	(45.158)	(3.753)							
MERVAL	5.987	3.199	0.990	5.080	0.443	0.254	0.151	0.152	0.871	0.000	322.99
	(3.572)	(7.898)	(48.691)	(3.235)							
STI	1.977	3.174	1.031	22.933	0.882	0.110	0.007	-	0.000	0.000	21.91
	(4.900)	(9.534)	(60.356)	(2.724)							
TWII	2.789	3.431	0.973	4.477	0.459	0.226	0.144	0.170	0.626	0.000	74.08
	(3.325)	(8.393)	(64.452)	(2.186)							

Notes: The table values are based on the MIDAS regression model (see equation 2) with unrestricted beta polynomial and with k^{max}=50. The corresponding t-statistics are provided in parentheses below the parameter estimates. The weighting schemes are illustrated by the columns "Week 1" to "Week >3" subsequently, where "Week 1" represents the total weight of the most recent week (or 5 days), "Week 2" represents that of the second last week (or the days 6-10), and so forth. The columns $Q(10)$ and $Q^2(10)$ represent the p-values for the Ljung-Box (1979) Q-statistic for ten lags of the MIDAS residuals and of the MIDAS squared residuals, respectively. MSPE denotes the mean squared prediction error.

Table 6: The forecasting performances of MIDAS and GARCH models

	$MSPE_{GARCH}$	$MSPE_{MIDAS}$	t-stat.
S&P500	13.020	11.189	0.510
FTSE	31.527	23.853	0.842
DAX	12.978	16.589	$-1.808^{\dagger}$
NIKKEI	42.365	30.780	1.888^{**}
BSE30	263.549	170.146	1.726^{**}
HSI	129.717	95.746	1.549^{*}
IBOVESPA	111.100	99.805	0.875
IPC	80.502	80.156	0.031
JKSE	181.784	109.872	1.293^{*}
KLSE	28.699	13.742	1.625^{*}
KS11	72.765	69.422	0.371
MERVAL	200.742	206.287	-0.178
STI	51.759	29.681	1.110
TWII	53.412	41.913	1.194

Notes: $MSPE$ denotes the mean squared prediction errors. $MSPE$s for each methodology are based on one-step-ahead out-of-sample forecasting. The HAC t-statistics are obtained from the regression of $\widehat{f}_t \equiv e^2_{GARCH,t} - e^2_{MIDAS,t}$ on a constant. The superscripts * and ** imply that the MSPE of GARCH is larger than that of the MIDAS with a significance level of .10 or .05, respectively. The corresponding critical values are 1.282 and 1.645. The superscript † shows that the MSPE of GARCH is significantly lower than that of the MIDAS at the 5% level of significance.