2002s-21

Information Content of Volatility Forecasts at Medium-term Horizons

John W. Galbraith, Turgut Kisinbay

Série Scientifique
Scientific Series

CIRANO
Centre interuniversitaire de recherche en analyse des organisations

Montréal
Février 2002

CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du ministère de la Recherche, de la Science et de la Technologie, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère de la Recherche, de la Science et de la Technologie, and grants and research mandates obtained by its research teams.

Les organisations-partenaires / The Partner Organizations

•École des Hautes Études Commerciales
•École Polytechnique de Montréal
•Université Concordia
•Université de Montréal
•Université du Québec à Montréal
•Université Laval
•Université McGill
•Ministère des Finances du Québec
•MRST
•Alcan inc.
•AXA Canada
•Banque du Canada
•Banque Laurentienne du Canada
•Banque Nationale du Canada
•Banque Royale du Canada
•Bell Canada
•Bombardier
•Bourse de Montréal
•Développement des ressources humaines Canada (DRHC)
•Fédération des caisses Desjardins du Québec
•Hydro-Québec
•Industrie Canada
•Pratt & Whitney Canada Inc.
•Raymond Chabot Grant Thornton
•Ville de Montréal

Les cahiers de la série scientifique (CS) visent à rendre accessibles des résultats de recherche effectuée au CIRANO afin de susciter échanges et commentaires. Ces cahiers sont écrits dans le style des publications scientifiques. Les idées et les opinions émises sont sous l'unique responsabilité des auteurs et ne représentent pas nécessairement les positions du CIRANO ou de ses partenaires.
This paper presents research carried out at CIRANO and aims at encouraging discussion and comment. The observations and viewpoints expressed are the sole responsibility of the authors. They do not necessarily represent positions of CIRANO or its partners.

ISSN 1198-8177

Information Content of Volatility Forecasts at Medium-term Horizons*

John W. Galbraith† *and Turgut Kisinbay*‡

Résumé / *Abstract*

En utilisant la volatilité réalisée pour estimer la volatilité conditionnelle quotidienne des rendements financiers, nous comparons les prévisions de volatilité quotidienne effectuées à partir de modèles GARCH-QVM standard et à partir de projections directes sur les volatilités réalisées. Nous considérons un horizon maximal de trente jours de transaction. Les prévisions sont comparées à la variance non conditionnelle des rendements quotidiens, ce qui nous permet d'estimer l'horizon maximal pour lequel les modèles détiennent un pouvoir de prévision. Nous utilisons des données de l'indice TSE 35 et des taux de change DM/US$ et Yen/US$, et nos résultats montrent qu'il y a un pouvoir de prédiction jusqu'à un horizon de trente jours, et ce, pour chacune des trois séries. Nous montrons aussi que le résultat de Bollerslev et Wright (2001), résultat indiquant que les projections sont supérieures sur l'horizon d'un jour, reste valide dans un horizon s'étendant jusqu'à dix ou quinze jours. Pour des horizons plus longs, les deux types de méthodes de prévision ne se différencient guère.

Using realized volatility to estimate daily conditional volatility of financial returns, we compare forecasts of daily volatility from standard QML-estimated GARCH models, and from projections on past realized volatilities obtained from high-frequency data. We consider horizons extending to thirty trading days. The forecasts are compared with the unconditional sample variance of daily returns treated as a daily volatility forecast, allowing us to estimate the maximum horizon at which the model-based forecasts provide forecasting power, measured by MSE reduction. Using data from a Toronto Stock Exchange equity index and foreign exchange returns (DM/$US and Yen/$US), we find evidence of forecasting power at horizons of up to thirty trading days, on each of the three financial returns series. We also find some evidence that the result of (e.g.) Bollerslev and Wright (2001), that projections on past realized volatility provide better 1-step forecasts than the QML-GARCH forecasts, appears to extend to longer horizons up to around ten to fifteen trading days. At longer horizons, there appears to be little to distinguish the forecast methods.

Mots-clés : GARCH, données à haute fréquence, volatilité intégrée, volatilité réalisée

Keywords: GARCH, high-frequency data, integrated volatility, realized volatility

JEL Classification: C22, C53

---

* We thank the *Fonds pour la formation de chercheurs et l'aide à la recherche (Québec)* and the *Social Sciences and Humanities Research Council of Canada* for financial support. Data were provided by the Toronto Stock Exchange, Market Data Services (TSE 35), and Olsen Associates (foreign exchange).
† Corresponding Author: John W. Galbraith, Department of Economics, McGill University, 888 Sherbrooke Street West, Montréal, Qc, Canada H3A 2T7 Tel.: (514) 985-4008 Fax: (514) 985-4039 email: galbraij@cirano.qc.ca.
‡ McGill University.

1. Introduction

The modelling of conditional variance of economic time series, following Engle (1982) and many subsequent contributions, has permitted characterization and forecasting of volatility in data such as time series of asset returns. While returns themselves are usually approximately unforecastable, forecasting of the *volatility* of returns has been quite successful. Of course, evaluating the degree of this success, as with the forecasting itself, has been complicated by the unobservability of the realized conditional variance.

Recent work with high-frequency data, in particular by Andersen and Bollerslev (1997, 1998), has provided an avenue for relaxing this constraint. Using direct estimates of daily conditional variance obtained from high-frequency (intra-day) squared returns may permit better evaluation of model-based daily volatility forecasts (Andersen and Bollerslev 1998), alternative methods of estimating parametric models of conditional variance (Bollen and Inder 1998, Maheu and McCurdy 2000, Meddahi and Renault 2000, Galbraith and Zinde-Walsh 2001), and characterization of daily conditional volatilities without recourse to parametric models (Andersen, Bollerslev, Diebold and Labys 2001a/b, Bollerslev and Wright 2001).

The present paper uses intra-day asset returns to examine model-based volatility forecasts at various horizons. As Christoffersen and Diebold (2000) note, risk depends on the horizon considered, and different horizons are relevant for different problems or, in the case of financial risk management, different classes of market participant. Nonetheless, little is known about the forecastability of volatility at horizons beyond the very short, such as one day. We investigate forecasts, both from the standard QML-estimated GARCH model and from projecting directly on past values of the realized volatility, at horizons of up to thirty trading days.

These forecasts are compared with the forecast implicit in the unconditional sample variance. We are then able to consider the maximum horizon at which the QML-GARCH or realized-volatility forecasts contain useful information about deviations of daily conditional variance from the unconditional variance. In doing so we use the definitions of forecast content of Galbraith (2001), and compute the forecast content function to the thirty-day horizon.

The next section of the paper describes the methods that will be applied in evaluating the information content of volatility forecasts. The third section describes the high-frequency data and forecasting models, and the fourth estimates QML-GARCH and realized-volatility forecast content for the daily conditional volatilities of the returns, and interprets the results.

2. Volatility forecast and content horizons

The conditional variance of a time series is not directly observable. For this reason, both estimation and evaluation of forecasting models is difficult relative to the conditional mean case. However, estimates of volatility can be obtained from the quadratic variation or realized volatility. For example, Andersen and Bollerslev (1998) address the question of the forecast performance of GARCH models using estimates of daily conditional volatility obtained from intra-day returns, and compare an R^2 -type measure of predictive power at a 1-day horizon. The exercise that we will perform here instead uses a measure based on relative MSE, to examine performance of models in forecasting daily volatility at a range of horizons.

The model of asset returns used by Andersen and Bollerslev, which we adopt, assumes that returns follow a continuous-time diffusion process such that instantaneous returns obey the relation $dp_t = \sigma_t dW_{p,t}$, where p_t is the logarithm of the price, σ_t the instantaneous standard deviation, and $W_{p,t}$ a Wiener process. Consider the sequence of $(Tm+1)$ discretely-sampled logarithmic prices $\{p_j\}, \quad j = 0, \frac{1}{m}, \frac{2}{m}, \ldots, T$, where integer values of the index j represent end-of-day closing prices and non-integer values represent intra-day observations. Transform the sequence to obtain Tm returns, $r_{[\ell,\ell+1]}$, $\ell = 0, \frac{1}{m}, \frac{2}{m}, \ldots, (T - \frac{1}{m})$. Defining $r_{[a,b]}$ as $p_b - p_a$, the daily return is $r_{[t-1,t]} = p_t - p_{t-1}$ and the return on the last of m intra-day periods of equal length between $t-1$ and t is $r_{[t-\frac{1}{m},t]} = p_t - p_{t-\frac{1}{m}}$.[1]

With these definitions, the daily integrated volatility is $\int_0^1 \sigma_{t+\tau}^2 d\tau$, and using the result that (Andersen and Bollerslev, 1998)

$$plim_{m\to\infty}\left(\int_0^1 \sigma_{t+\tau}^2 d\tau - \sum_{i=1}^{m} r^2_{[t+\frac{(i-1)}{m},t+\frac{i}{m}]}\right) = 0, \tag{1}$$

it follows that the summation of intra-day squared returns, or daily realized volatility, provides an estimate of the daily conditional volatility. It is the performance of daily volatility models in forecasting this quantity of interest which is evaluated both by Andersen and Bollerslev and in the present paper. Note that the result (1) requires that the discretely-sampled squared returns $r^2_{[t+\frac{(i-1)}{m},t+\frac{i}{m}]}$ be serially uncorrelated and that the price process be

[1]This is a slight modification of Andersen and Bollerslev's notation for the returns.

continuous. In practice this summation will provide only an approximation to $\int_0^1 \sigma^2_{t+\tau} d\tau$, the quality of which will depend upon the adequacy of these assumptions as well as on the intra-day sampling frequency m (that is, the asymptotics). Meddahi (2001) provides a detailed examination of the discrepany between integrated and realized volatility.

The daily realized volatility therefore gives us the target for the forecasting technique, analogous to the realized outcome in forecasting an observable series. With this information we proceed to the evaluation of forecasts at different horizons, making use of the forecast content function as defined in Galbraith (2001), modified slightly here to reflect the fact that we are evaluating second-moment forecasts. There, the forecast content function was defined as the proportionate reduction in MSE obtainable relative to the unconditional mean forecast, in forecasting the level of a quantity observable *ex post*. Adapting that definition for a conditional volatility forecast, we set

$$C(s) = 1 - \frac{MSE_{(\tilde{\sigma^2}(s))}}{MSE_{(\hat{\sigma^2}(s))}}, \qquad s = 1, \ldots, S, \tag{2}$$

where $\tilde{\sigma}(s)$ and $\hat{\sigma}(s)$ are respectively the model-based s-step-ahead forecast of the daily conditional volatility, and the estimated unconditional variance. As in the case of conditional mean forecasts, $C(s) \to 0$ as $T, s \to \infty$ for a forecast $\tilde{\sigma^2}_{T+s|T}$ based on a correctly-specified forecasting model; $C(s)$ may be less than zero where model parameter uncertainty dominates the contribution of these parameters to prediction; $C(s)$ near zero indicates model-based forecasts no better as predictors than the unconditional standard deviation.

For general autoregressive processes, the content function can be expressed as a function of the autoregressive parameters and sample size available for estimation, taking account of parameter estimation uncertainty; expressions for $C(s)$ are given in Galbraith (2001). For a set of forecasts from an arbitrary forecasting model, $C(s)$ can be estimated from a sequence of outcomes and forecasts using the sample mean squared errors and associated asymptotic inference. This is the method used below for the forecasts evaluated in Section 4.

3. Data and Models

We evaluate the forecast content functions of two types of asset: an equity price index and two currencies, priced relative to the U.S. dollar. In each case we examine the daily logarithmic returns. High-frequency intra-day data (bid, ask and index value at last trades) are available on the equity index at 15-second intervals,and on the foreign exchange prices (bid and ask) at five-minute intervals.

A point relevant to all of the data series, noted in Andersen and Bollerslev (1998) and explored further in Andersen, Bollerslev, Diebold and Labys (2001a/b), is that the highest possible frequency of observation may not be optimal from the point of view of daily volatility estimation, because of market micro-structure or other effects. As in ABDL (2001a/b), we consider various possible sampling frequencies for the estimation of daily volatility.

3.1 (i) TSE 35 Index

The equity index that we use is the Toronto Stock Exchange (TSE) 35 index of large-capitalization stocks traded on the TSE. This index was chosen because of the availability of a reasonably long daily time series for standard QML estimation of GARCH models (dating from creation of the index in 1987) combined with a recent sequence of high-frequency data allowing us to compute realized volatility measures for forecast evaluation.[2]

In addition to daily data from 1987 through 1998 inclusive, we use intra-day data on the TSE 35 index value from calendar year 1998, to evaluate the forecasts of volatility. These data are available at intervals of fifteen seconds throughout the 9:30 a.m. to 4:00 p.m. trading day, for a total of approximately 1560 observations a day. Bid and ask are also available, beginning several hours before the trading day.

Despite the simplicity of the Andersen-Bollerslev result (1) for estimating the daily volatility, there are several difficulties in computing estimates on these data, and in general in actual realizations of asset prices.

Here, a first problem arises from the fact that trading does not take place throughout the 24-hour day, as with some foreign exchange markets. Consider an example in which an asset trades solely within 9:30 - 4:30 period. If the opening value at date t is equal to the closing value at date $t-1$, then all volatility is captured by the intra-day returns in the 9:30 - 4:30 period, and (1) may be applied directly, with the first squared return of the day

[2]The Toronto 35 was introduced as a notional portfolio of shares of the companies, chosen to track the broader TSE 300 index fairly closely. The primary determinant of a stock's weighting in the portfolio is its 'float quoted market value.' This term is defined by the TSE as the trade-weighted average price for the year, multiplied by the year-end share float. (The set of policies governing inclusion, weighting and maintenance is recorded at www.tse.com.) In February of 2000 unitholders of the TSE 35 (and also TSE 100) Index Participation Fund approved a merger of the fund with the S & P / TSE 60 Index Participation Fund (i60).

captured from (e.g.) 9:30:00 to 9:30:15. By contrast a change in the index between close and open on the next day will not be captured by this computation.

In actual equity price data, including these index data, there is often a change between the 4:30 close and 9:30 open on the following day, that represents a contribution to volatility which, if ignored, will lead to underestimation of $\hat{\sigma}_t^2$. One possibility for incorporating this component of volatility is to treat the squared return from close to subsequent opening as an additional element of $\hat{\sigma}_t^2$, added to other squared returns. If we think of the index as having a notional value throughout the non-trading hours, the overnight squared return is an unbiased, but noisy estimate of overnight volatility (analogous to the use of squared daily return, r_t^2, as a proxy for the daily volatility, in the pre-integrated-volatility literature). Because this change is generally small, however, the amount of noise added should be correspondingly small.

A related problem arises here in that the first few minutes of the trading day typically show the index value outside the bid/ask range; within the first two minutes of trading, the index value is usually again within the range. One interpretation of this is that the bid/ask better reflect the notional overall value of the index to investors, but the actual calculated index value, arising from the aggregation of individual security prices, takes some time to reflect this value through arbitrage. The opening index values may therefore be somewhat unrealistic (too closely tied to the previous day's close). One option for handling this, which we explore, is to use the midpoint between bid and ask for the first two minutes of the trading day, by which point the two measures are almost invariably compatible.[3]

Each of three sets (for different sampling frequencies: 1-, 5- and 20-minute) of daily realized volatility estimates, $\{\hat{\sigma}_t^2\}_{t=1}^{252}$, is computed on each of the 252 trading days in 1998 (with the exception, noted above, of the error-laden observations for 23 January), and is used in the forecast evaluation below. The daily squared return as a volatility measure is also included for comparison.

3.1 (ii) Foreign exchange data

The foreign exchange data are Deutschemark-U.S. dollar and Yen-U.S. dollar spot exchange rates, as described by Andersen and Bollerslev (1998), taken from the Olsen & Associates HFDF-2000 data set. The original sources of the raw data that are used to

[3]An exception is 23 January 1998, for which the index values and bid/ask are grossly different throughout much of the trading day, evidently a data recording problem. We replace this day's estimated volatility by its squared return in the computations below.

construct HFDF are the DM/$US and Yen/$US bid-ask quotes displayed on the Reuters FXFX screen. These raw data are subject to various microstructure frictions, outliers and other anomalies, and have to be filtered; see Müller et al. (1990) and Dacorogna et al. (1993) for the filtering procedures and the construction of the returns. Because these data have been fairly widely used, our description will be brief.

Each data point in the HFDF-2000 data set contains a mid-price, a bid-ask spread over a 5-minute interval of spot rate quote and a time stamp, resulting in a total of 288 5-minute returns for a given day. The 5-minute returns, expressed in basis points (i.e. multiplied by 10,000), are computed as the mid-quote price difference. The mid- price at a given regular time point is estimated through a linear interpolation between the previous and following mid-price of the irregularly spaced tick-by-tick data. The bid-ask spread is the average over the last 5-minute interval, again expressed in basis points. If there is no quote during this interval, the mean bid-ask spread is zero. The sample spans the period from January 2, 1987 to December 31, 1998, providing us with a sample of 1,262,016 5-minute returns, expressed in USD terms. However, it is well known that trading activity in the foreign exchange market slows markedly over the weekends and certain holidays. Following Andersen et al. (2001a), among others, we remove such low-activity days from our data set. Whenever a day t is excluded from the data set, we cut from 21.05 GMT of the previous calendar day $t-1$ to 21:00 GMT on the calendar day t. This definition of a "day" is standard in the literature and follows Bollerslev and Domowitz (1993).

In addition to the low-activity period from Friday 21:05 GMT to Sunday 21.00 GMT, we remove the following fixed holidays from our data set: Christmas (December 26-26), New Year's (December 31-January 2), and the Fourth of July. We also remove the moving holidays of Good Friday, Easter Monday, Memorial Day, the Fourth of July, Labour Day, Thanksgiving (US) and the day after Thanksgiving. Finally, we cut the days for which the indicator variable (the bid-ask spread) had 144 or more zeros, corresponding with the technical "holes" in the recorded data. After these adjustments, we are left with 2,968 trading days of DM/$US data corresponding with $2,968 \times 288 = 854,784$ five-minute return observations, and 2970 days of Yen/$US data corresponding with $2,970 \times 288 = 855,360$ five-minute return observations. A final adjustment is made to the Yen/$US data set: because of the extremely large volatility observations occurring at 7-8 October 1998 and surrounding dates,[4] we terminate this data set at the end of May 1998, leaving 2830 days

[4]These days correspond with the period around the collapse of Long Term Capital Manage-

of high-frequency data. We comment below on the effect of this trimming on the relative performance of the two forecast types.

As with the TSE index data, we report results for three different aggregations of the intra-day data as well as for the daily squared return, for comparison. These data sets differ from the TSE data in that the highest frequency available is 5-minute returns, and so the three aggregations that we choose are correspondingly less fine (5-, 10- and 30-minute returns). However, the sequence of daily high-frequency returns corresponds to over ten times as many days as for the TSE data. The TSE data also differ in the existence of a prior sample of daily data from which to construct QML GARCH estimates. Here, we instead use an initial sample of daily returns computed from the high-frequency data for QML estimation.

3.2 Forecasting models

We use two classes of forecasting model to evaluate the information content of volatility forecasts at horizons extended to thirty days. The first is the standard GARCH model estimated by quasi-Maximum Likelihood methods on samples of daily returns data; these forecasts do not use the information present in the high-frequency data. The second class is the forecast of conditional volatility obtained from projection of realized volatility on past values of realized volatility: that is, autoregressive models of the realized volatility, as examined by ABDL (2001a/b) and Bollerslev and Wright (2001).

In the QML case we restrict ourselves to the GARCH(1,1) model which Bollerslev and Wright, for example, found to produce the best one-step volatility forecasts by a substantial margin, on a set of DM/$US data very similar to that used here. The latter method is not applied to the TSE equity index data because of the relatively small number of days (252) for which high-frequency data were available.

Parameters of the GARCH(1,1) model

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2, \tag{3a}$$

or in terms of squared returns,

$$(1 - \alpha L - \beta L)\varepsilon_t^2 = \omega + (1 - \beta L)(\varepsilon_t^2 - \sigma_t^2), \tag{3b}$$

are estimated for the TSE 35 index on the initial sample of 2772 observations from January 1987 through 31 December 1997, and updated with values from 1998 as the data from

---

ment.

which forecasts are made advances. From each forecast date, beginning with 2 January 1998, out-of-sample forecasts are produced for s days in the future, $s = 1, 2 \ldots 30$. For the two foreign exchange series, an initial sample of daily data is taken from the first two years, 1987-88, for inital estimation of the GARCH(1,1) parameters. Thereafter these estimates are again updated at each day for recursive estimation and out-of-sample forecasting.

Forecasts by autoregressive projection on past realized volatilities are also computed recursively beginning with the same initial samples. Note that these forecasts use the additional information present in the high-frequency data, and are therefore not based on the same information set as the QML GARCH forecasts.

For each data set, we then also estimate the unconditional variance recursively using the same initial sample, in order to evaluate the performance of the forecasting models in predicting deviations of the conditional variance from this quantity; that is, we ask whether the model-based forecast of volatility on a given day is superior to simply using the estimated unconditional variance to predict volatility.

Estimates of the volatility forecast content function are obtained as follows. Let $\tau = 1, 2, \ldots, T$ index the full sample of trading days for which high-frequency information is available, on any one of the three data sets (that is, T=252, 2968 and 2970 for TSE, DM and Yen respectively). Let $s = 1, 2, \ldots, S$ index the forecast horizon; we consider a maximum horizon of $S = 30$ days. Using an initial sample of size t_0, which can be set to 0 for the TSE data because of the existence of the daily pre-sample, construct the $(T - t_0 + 1) \times S$ matrix E having typical element $\epsilon_{\tau,s}$, defined as the forecast error from the model-based forecast for trading day τ given a forecast made s days in the past. Construct the corresponding $(T - t_0 + 1) \times 1$ vector U having typical element u_τ, defined as the difference between the estimated unconditional variance of the returns and the realized volatility for trading day τ. The latter quantities do not depend on s because the unconditional variance does not, of course, depend on s. For the entries in both E and U, we take the sample mean squared error obtained in comparing the forecast with the realized volatility, at the chosen level of aggregation (e.g. 5 minutes, 10 minutes). These sample quantities are substituted into (2), for each s.[5]

[5]Pointwise standard errors for the content function are computed using the approximate expression

$$var\left(\frac{X}{Y}\right) \simeq \left(\frac{\mu_X}{\mu_Y}\right)^2 \left[\frac{var(X)}{\mu_X^2} + \frac{var(Y)}{\mu_Y^2} - \frac{2cov(X,Y)}{\mu_X \mu_Y}\right].$$

4.Empirical results

Since the TSE high-frequency data set used here is both relatively short and less well known than the others, we begin with some plots of relevant quantities which will serve to illustrate some typical relationships. We do not plot the realized volatilities for the foreign exchange data.

Figure 1(a-c) shows estimated daily volatilities $\{\sigma_t^2\}_{t=1}^{252}$ for a variety of values of a time-aggregation parameter k which indicates the number of 15-second returns which are summed to obtain a single intra-day squared return: $k = 4$ therefore uses 1-minute returns, $k = 20$ uses 5-minute returns, and $k = 40$ 10-minute returns.[6] For comparison, Figure 1d plots the squared returns on each of the trading days of 1998 together with a representative set of 1-step-ahead forecasts, from a GARCH(1,1) model estimated by QML on the pre-sample of daily data. The high equity-market volatility of August 1998, and its tracking by the GARCH estimates, is readily observable. The GARCH forecasts are of course much smoother than the squared returns, being forecasts of the conditional expectations of the squared returns; each of these features implies smoothing.[7] In comparing the different values, note the difference in vertical scales; while the overall form is similar among the different values of k, higher values produce greater extremes of variation (approaching the daily squared return as k increases to 1560). For any moderate value of k, the daily volatilities are a much smoother sequence than the squared returns of which they are (estimates of) the conditional expectations.

The next figures record the estimated forecast content functions for the three data sets and the two (except in the case of the TSE) forecast methods. Each figure shows the forecast content function for three different time aggregations, together with the function estimated on squared returns, for comparison. The latter is of interest primarily to underline, in accordance with Andersen and Bollerslev (1998), the value of realized volatility in exploring the performance of conditional variance forecasts; using the very noisy squared return series,

---

Here X is the estimated MSE of model-based forecasts for each s, and Y is the estimated MSE of the forecast implicit in the unconditional variance. We replace each of the population quantities on the RHS of this expression with their sample counterparts for asymptotic inference.

[6] ABDL (2001b) suggest a method for choosing a time aggregate based on 'volatility signature plots'.

[7] The parameters of the estimated GARCH(1,1) process (log returns scaled by 100) are $(\hat{\omega}, \hat{\beta}, \hat{\alpha}) = (0.03, 0.85, 0.12)$.

which was used for forecast evaluation in literature pre-dating Andersen and Bollerslev (1998), yields no significant indication of forecasting power on any data set. Each content function C(s) is shown with ± 2 standard error bands.

Figure 2(a-d) plots estimated forecast content functions for the TSE 35 index, forecast with the QML-estimated GARCH(1,1) model, on the three realized volatility measures and the untransformed daily squared returns. Figures 3(a-d) and 4(a-d) do so for the DM/$US and Yen/$US data respectively, with aggregations of 5, 10 and 30 minutes used in estimating realized volatility, as well as with the squared daily return for comparison. The TSE data show much wider confidence intervals, reflecting the fact that the data set contains fewer than one tenth the number of days of high-frequency data available for foreign exchange. Nonetheless the forecast content on 5-minute returns is significantly positive to about sixteen days forward, and the point estimate is positive to the thirty-day maximum horizon considered here. Results on the 10-minute aggregation are similar, although the 5-minute results are the strongest. Results on the 1-minute aggregation are very weak, suggesting substantial noise from microstructure effects. These weak results also suggest that the absence of higher-frequency returns in the foreign exchange data does not represent a substantial loss of information for the present purpose.

On the foreign exchange data, all estimates on realized volatility measures (parts a-c) show positive point estimates of forecast content remaining at the thirty-day horizon: use of a formal volatility forecasting model for daily volatility shows positive information content 30 days into the future, in the sense that the forecasts improve on the unconditional variance of the process as an indicator of the realized volatility that will emerge. In each data set, the five-minute aggregation provides the highest estimate of forecast content.

The results for volatility forecasts made via autoregressions on realized volatility appear in Figures 5 and 6. The pattern is very similar to that appearing in Figures 3 and 4 respectively; however, the point estimates of forecast content are in general higher for the realized volatility forecasts for approximately 10-15 trading days. Toward the end of the 30-day maximum horizon that we consider, the two types of forecasts appear to be very similar, although point estimates of forecast content become higher for the GARCH forecasts at the longer horizons considered. As one would expect, there is virtually no evidence of significant content in applying this method directly to squared returns (Fig. 5d, 6d). Again, the strongest results on both data sets appear in the 5-minute returns.

It is important to recall that even the best realized volatility measure contains measurement noise relative to the daily integrated volatility $\int_{t-1}^{t} \sigma_{t+\tau}^2 d\tau$. Such measurement noise tends to lower all measures of the value of forecasts, relative to that which would be measured with knowledge of the true daily volatility, by obscuring the match between forecast and outcome. For this reason, we interpret the highest forecast content across measures as evidence in favour of the superiority of that measure. The evidence in these data sets therefore favours the five-minute aggregation.

A numerical summary of some of these results appears in Table 1, for the 5-minute aggregation.[8]

Table 1
Numerical summary of forecast content
Realized volatility of 5-minute returns

Forecast method	TSE 35	DM/$US	Yen/$US
QML-GARCH			
C(1) (se)	0.30 (.12)	0.26 (.02)	0.25 (.03)
C(5) (se)	0.20 (.08)	0.15 (.02)	0.11 (.02)
C(20) (se)	0.08 (.06)	0.07 (.01)	0.04 (.01)
(30) (se)	0.01 (.05)	0.05 (.01)	0.02 (.01)
AR-rv			
C(1) (se)	—	0.39 (.03)	0.32 (.03)
C(5) (se)	—	0.20 (.02)	0.15 (.02)
C(20) (se)	—	0.05 (.01)	0.02 (.01)
C(30) (se)	—	0.001 (.02)	-0.004 (.01)

[8]Note however that in trimming the Yen sample to remove the apparent outliers, we have given an advantage to the AR-realized volatility forecasts, which are more sensitive to this effect than the GARCH forecasts; the latter are based on estimated daily volatilities which fluctuate much less extremely than the actualy realized volatilities. There is some indication that limiting the degree of correlation in very large realized volatilities may be less than in more moderate values.

These results tend to confirm, for moderate horizons, the results of Bollerslev and Wright (2001) concerning the superiority of AR forecasts over QML-GARCH. At longer horizons, however, the GARCH forecasts tend to have slightly higher point estimates of information content. More precise evidence on this point comes from a sequence of Diebold-Mariano (1995) tests for the null of equal forecast accuracy, on the two sequences of forecasts, for each of the two currencies.

Table 2
Diebold-Mariano tests of equal forecast loss: p-values
GARCH(1,1) vs AR-realized volatility forecasts
Realized volatility of 5-minute returns

Horizon (s)	DM/$US	Yen/$US
1	<0.01	0.14
2	0.01	0.04
3	0.01	0.28
4	0.07	0.21
5	0.03	0.09
10	0.09	0.13
15	0.17	0.48
20	0.56	0.72
30	0.67	0.36

Clearly the results are stronger on the DM$US data; there is some weak evidence on the Yen/$US data that the AR-realized volatility forecast may be superior at fairly short horizons. In either case, the point estimates of the content of the two types of forecasts tend to converge to zero around the thirty-day horizon. Although, as we have noted, the GARCH forecasts do have higher point estimates of forecast content for the longer horizons, there do not appear to be statistically significant differences.

5.Concluding remarks

To measure forecast performance, we need to estimate the outcome (volatility) which was forecast. The use of realized volatility, as in Andersen and Bollerslev (1998) and various subsequent contributions, has clearly improved our ability to do so, and we take advantage of this measurement here to examine the usefulness of forecasts at longer horizons than have commonly been investigated. Nonetheless, as in ABDL (2001a/b), the choice of frequency of measurement has non-trivial affects on the measured realized volatility. These different estimates affect our measures of forecast content; we find, however, that the five-minute realized volatilities provide the strongest results on all data sets, including the TSE data for which higher-frequency measurements, as frequent as 15 seconds, are available.

The results have a number of other features which appear common to the different data sets. First, forecasts of conditional volatility appear to have information content to a horizon of approximately thirty trading days, although the precise maximum horizon of course depends on the data set and forecast method. It is possible that a measure of daily volatility with even less noise than the five-minute realized volatilities would provide yet stronger evidence of the value of volatility forecasts. Second, forecasts based on autoregressive projection on past realized volatilities do appear to provide better results at short horizons, as was found by Bollerslev and Wright (2001) for the 1-day horizon (also on DM/$US data); evidence on the Yen/$US data on the significance of this effect is weak, however. At longer horizons, as the content of each type of forecasts nears zero, the methods become difficult to distinguish on statistical grounds. Either method, however, provides some forecasting power to approximately thirty trading days.

References

Andersen, T.G. and T. Bollerslev (1997) Intraday Periodicity and Volatility Persistence in Financial Markets. *Journal of Empirical Finance* 4, 115-158.

Andersen, T.G. and T. Bollerslev (1998) Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. *International Economic Review* 39(4), 885-905.

Andersen, T.G., T. Bollerslev, F.X.Diebold and P.Labys (2001a) The Distribution of Realized Exchange Rate Volatility. *Journal of the American Statistical Association* 96, 42-55.

Andersen, T.G., T. Bollerslev, F.X.Diebold and P.Labys (2001b) Modeling and Forecasting Realized Volatility. Working paper.

Barndorff-Neilsen, O.E. and N. Shephard (2000) Econometric analysis of realised volatility and its use in estimating Lévy-based non-Gaussian OU type stochastic volatility models. Working paper, Nuffield College, Oxford.

Bollen, B. and B. Inder (1998) A General Volatility Framework and the Generalised Historical Volatility Estimator. Working paper, Monash University.

Bollerslev, T. and I. Domowitz (1993) Trading Patterns and Prices in the Interbank Foreign Exchange Market. The Journal of Finance 48, 1421-43.

Bollerslev, T. and J.H. Wright (2001) High-frequency Data, Frequency Domain Inference, and Volatility Forecasting. *Review of Economics and Statistics* 83, 596-602.

Christoffersen, P. and F.X. Diebold (2000) How Relevant is Volatility Forecasting for Financial Risk Management? *Review of Economics and Statistics* 82, 12-22.

Dacorogna, M.M., U.A. Müller, R.J. Nagler, R.B. Olsen, and O. V. Pictet (1993) A geographical Model for the Daily and Weekly Seasonal Volatility in the Foreign Exchange Market. *Journal of International Money and Finance* 12, 413-438.

Diebold, F.X. and R. Mariano (1995) Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13, 253-59.

Engle, R.F. (1982) Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica* 50, 987-1008

Galbraith, J.W. (2001) Content Horizons for Univariate Time Series Forecasts. Forthcoming, *International Journal of Forecasting* .

Galbraith, J.W. and V. Zinde-Walsh (2001) GARCH Model Estimation Using Noisy Realized Volatility. Working paper, McGill University and CIRANO.

Maheu, J.M. and T.H. McCurdy (2000) Nonlinear Features of Realized FX Volatility. Working paper, University of Alberta/University of Toronto.

Meddahi, N. (2001) A Theoretical Comparison Between Integrated and Realized Volatilities. Working Paper 2001s-71, CIRANO.

Meddahi, N. and E. Renault (2000) Temporal Aggregation of Volatility Models. Working paper 2000s-22, CIRANO.

Müller, U.A., M.M. Dacorogna, R.B. Olsen, O.V. Pictet, M. Schwarz and C. Morgenegg (1990) Statistical Study of Foreign Exchange Rates, Empirical Evidence of a Price Scaling Law, and Intraday Analysis. *Journal of Banking and Finance* 14, 1189-1208.

Figure 1a: TSE 35
Realized volatility, 1-minute (k=4)

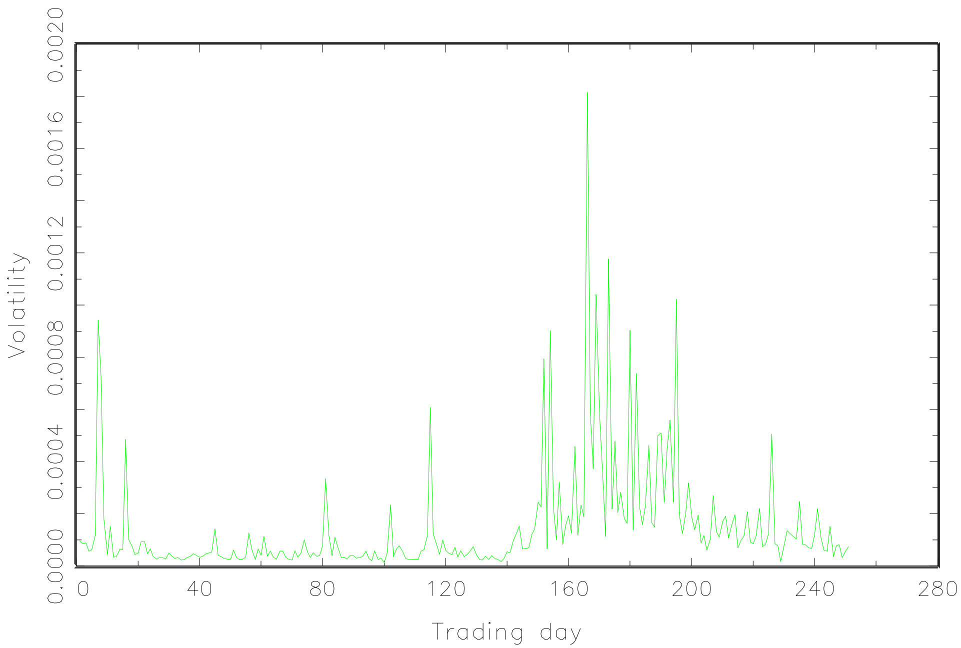


Figure 1b: TSE 35
Realized volatility, 5-minute (k=20)

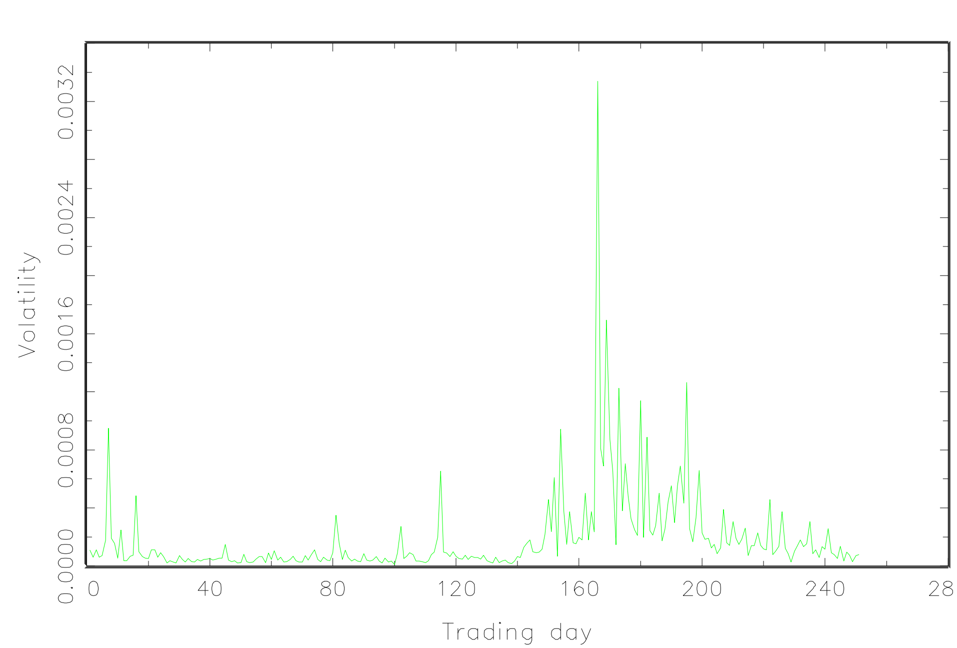


Figure 1c: TSE 35
Realized volatility, 10-minute (k=40)

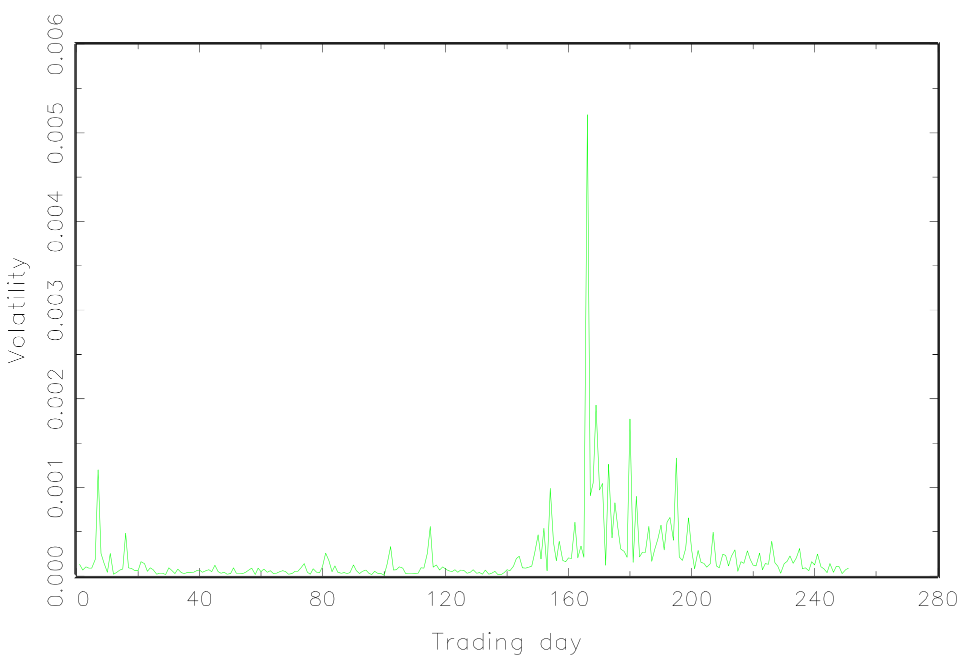


Figure 1d: TSE 35
GARCH(1,1) 1-step forecasts and squared returns (k=1560)

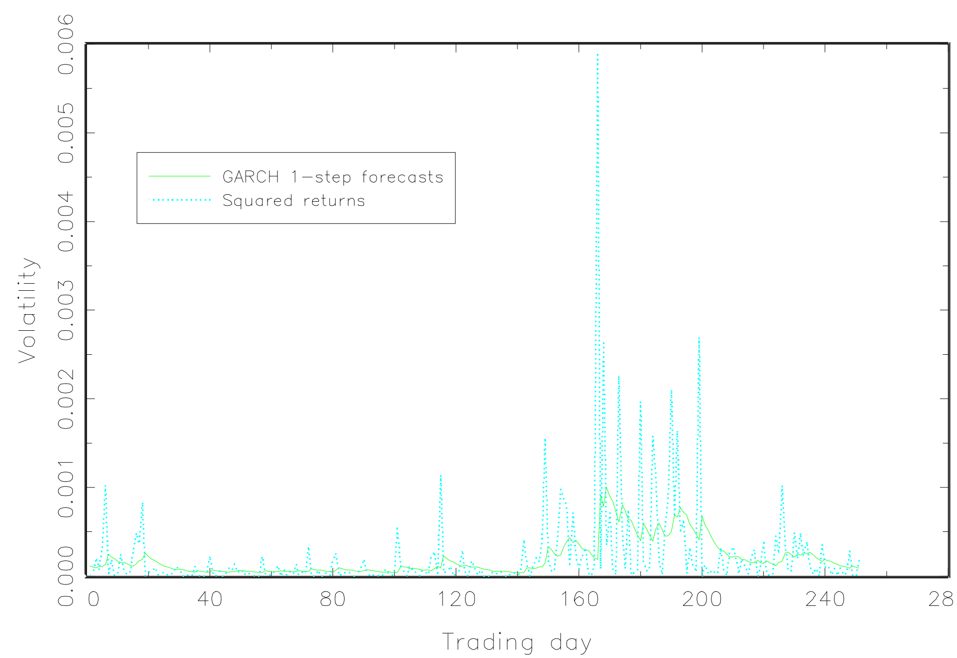


Figure 2a: TSE 35
Forecast content function, GARCH(1,1) forecasts,
realized volatility of 1−minute returns

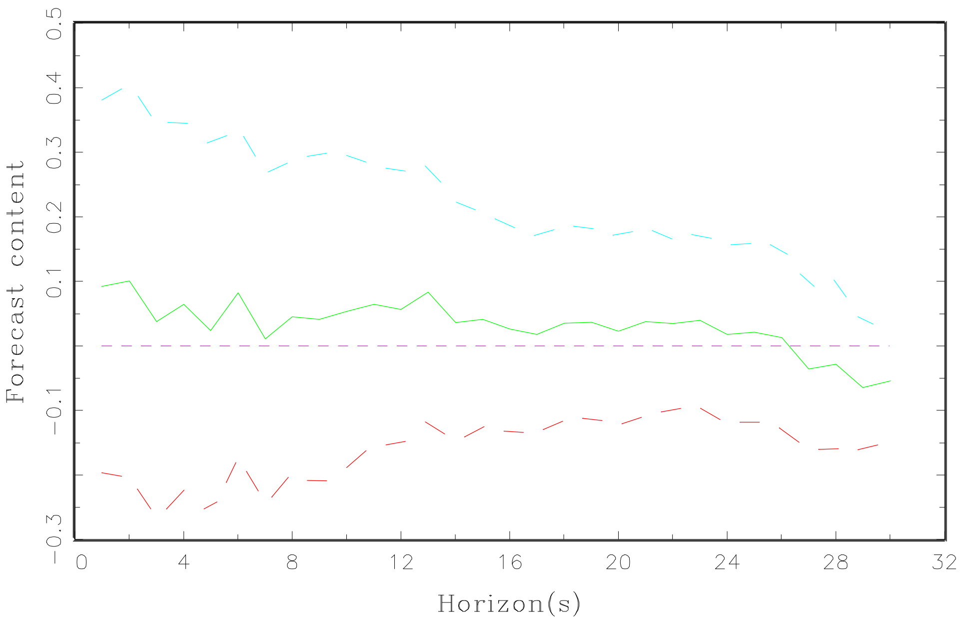


Figure 2b: TSE 35
Forecast content function, GARCH(1,1) forecasts,
realized volatility of 5−minute returns

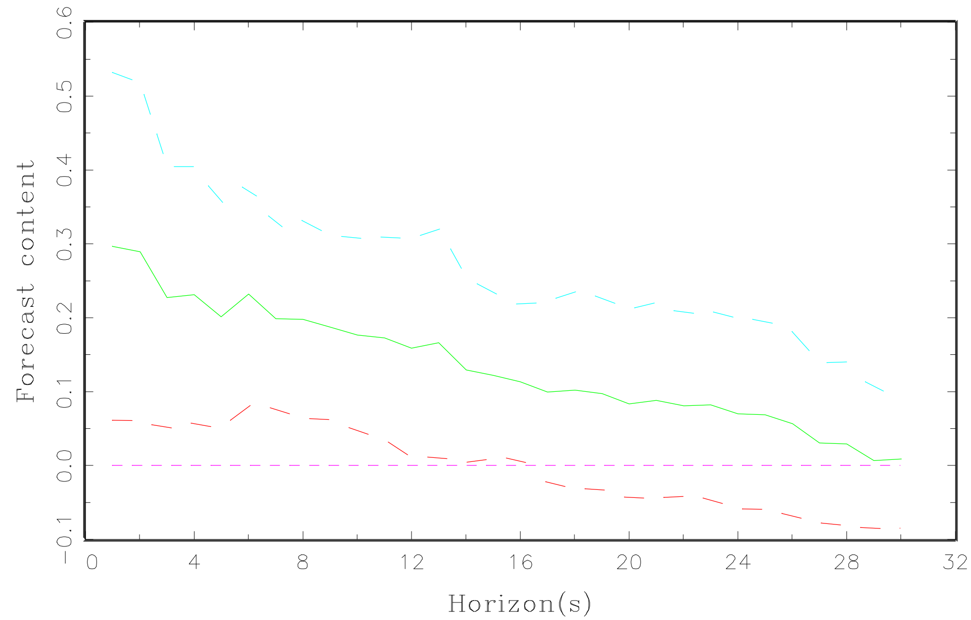


Figure 2c: TSE 35
Forecast content function, GARCH(1,1) forecasts,
realized volatility of 10−minute returns

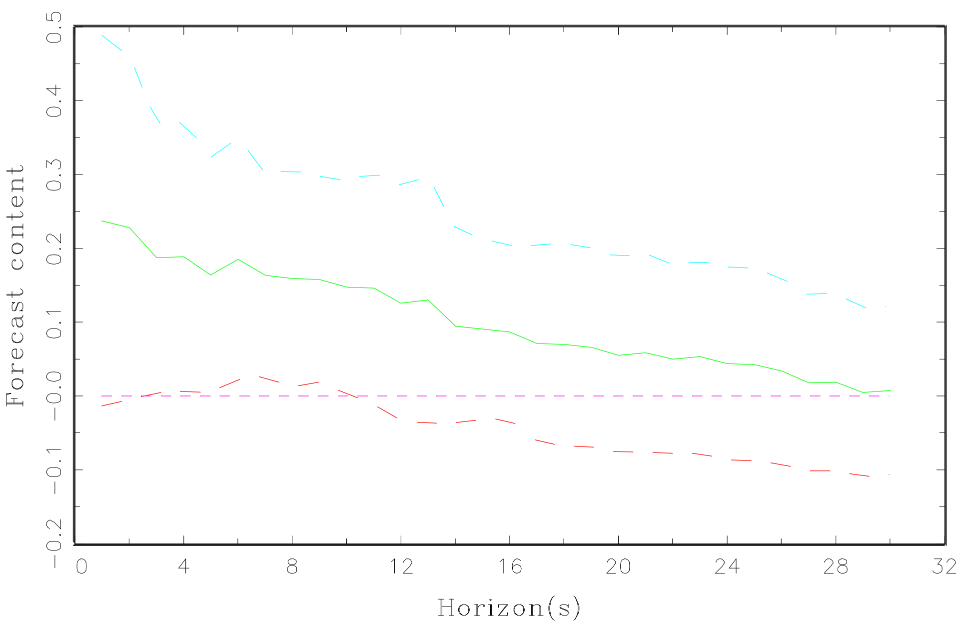


Figure 2d: TSE 35
Forecast content function, GARCH(1,1) forecasts,
squared daily returns

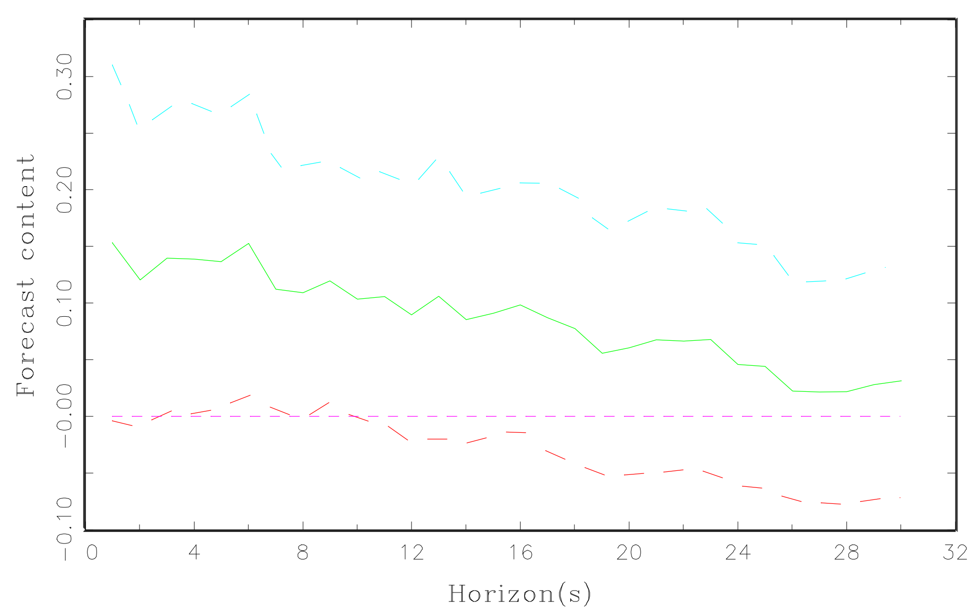


Figure 3a: DM/USD
Forecast content function, Garch (1,1) forecasts,
realized volatility of 5-minute returns

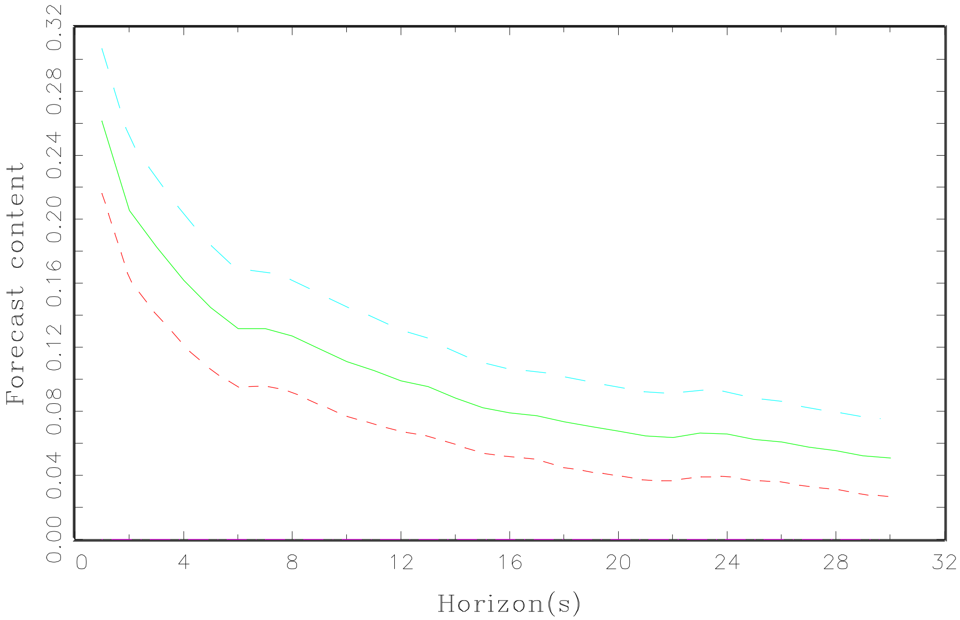


Figure 3b: DM/USD
Forecast content function, Garch (1,1) forecasts,
realized volatility of 10-minute returns

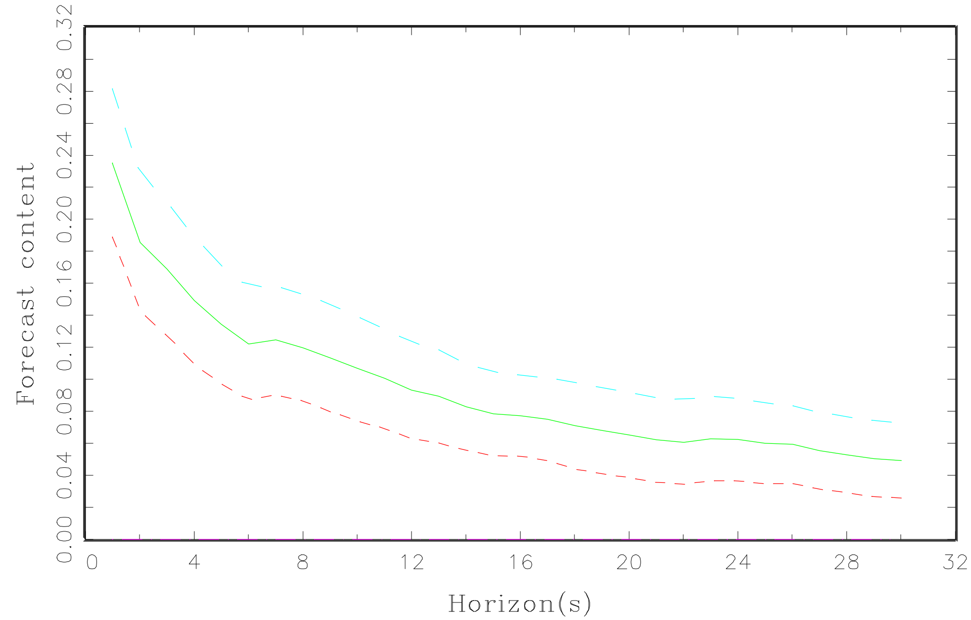


Figure 3c: DM/USD
Forecast content function, Garch (1,1) forecasts,
realized volatility of 30-minute returns

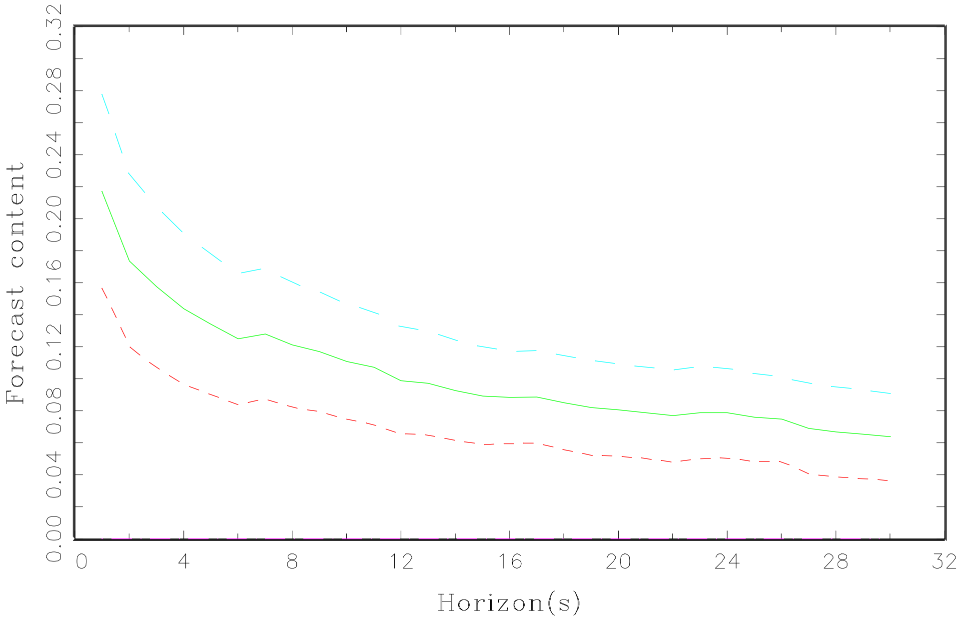


Figure 3d: DM/USD
Forecast content function, Garch (1,1) forecasts,
squared daily returns

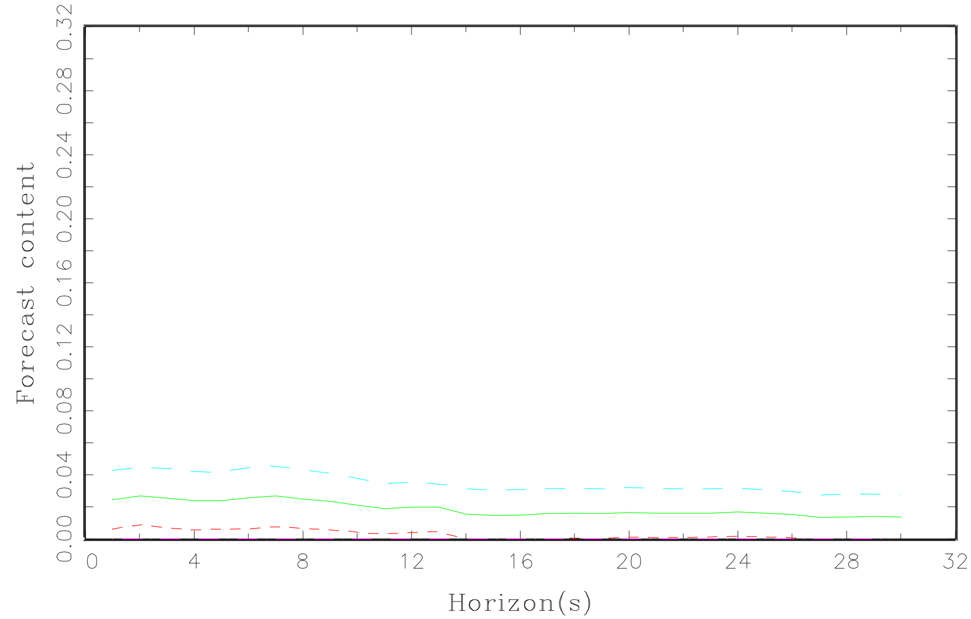


Figure 4a: Yen/USD
Forecast content function, Garch (1,1) forecasts,
realized volatility of 5-minute returns

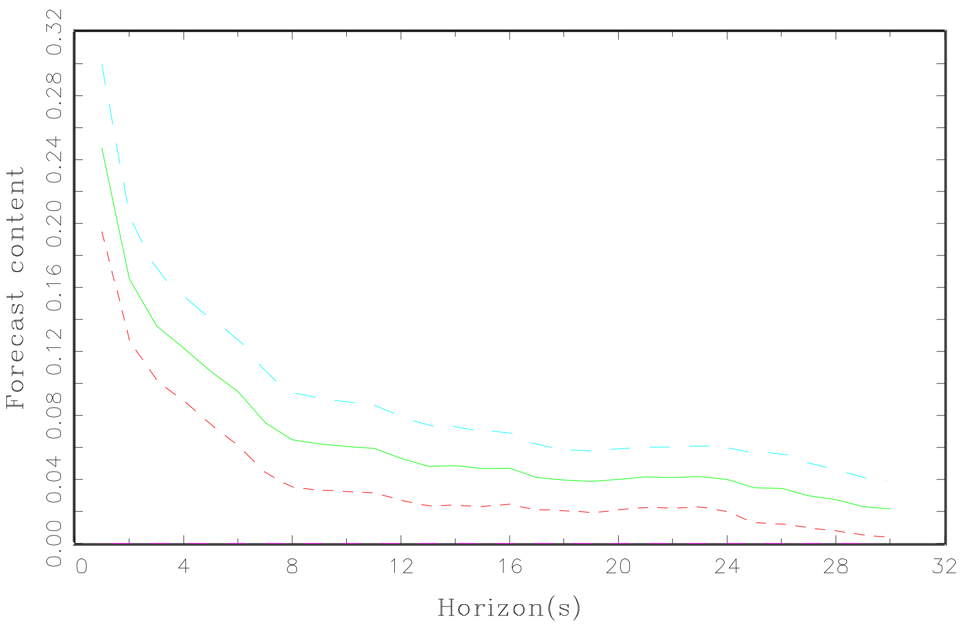


Figure 4b: Yen/USD
Forecast content function, Garch (1,1) forecasts,
realized volatility of 10-minute returns

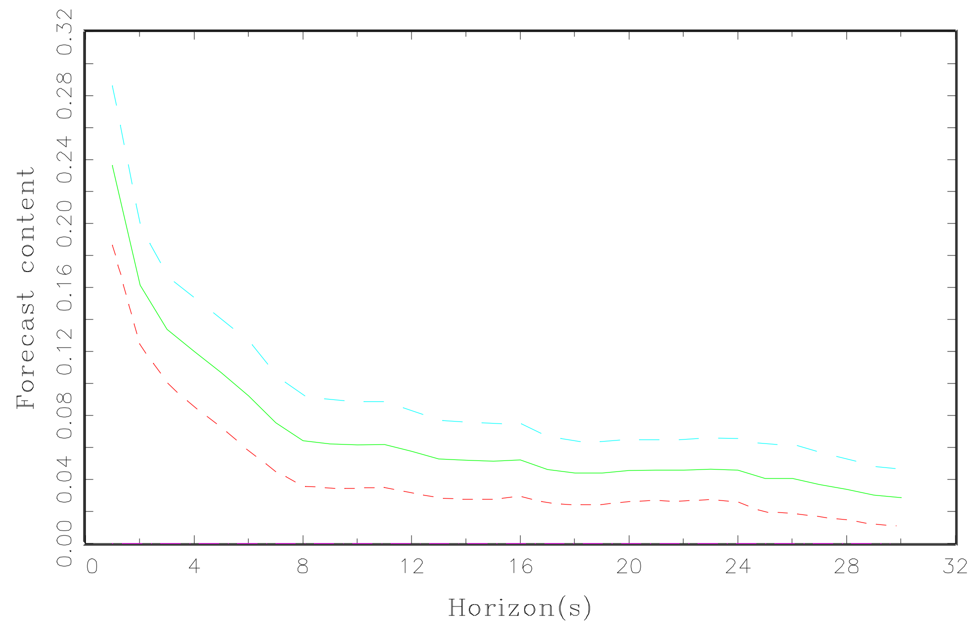


Figure 4c: Yen/USD
Forecast content function, Garch (1,1) forecasts,
realized volatility of 30-minute returns

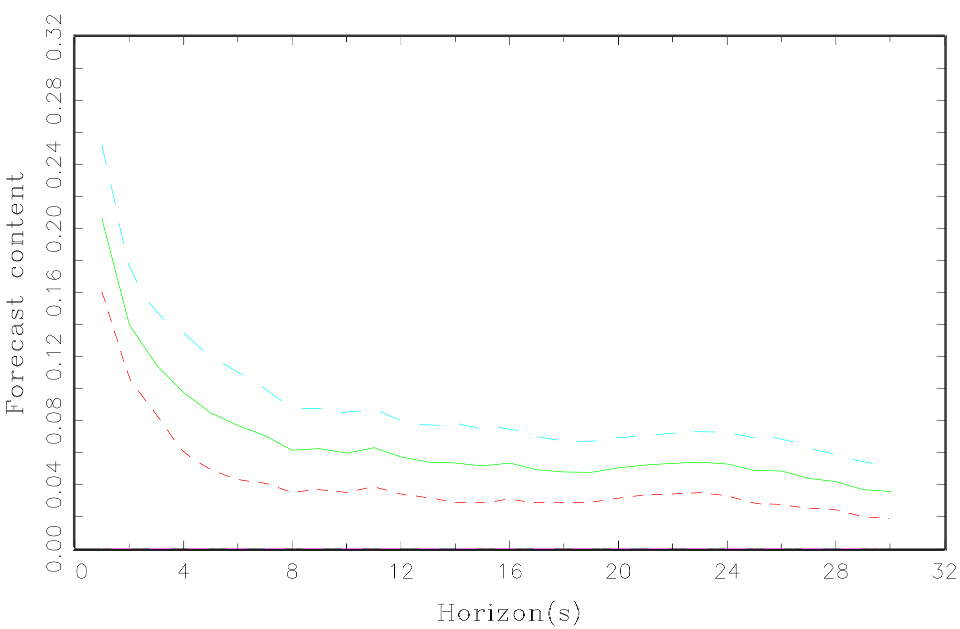


Figure 4d: Yen/USD
Forecast content function, Garch (1,1) forecasts,
squared daily returns

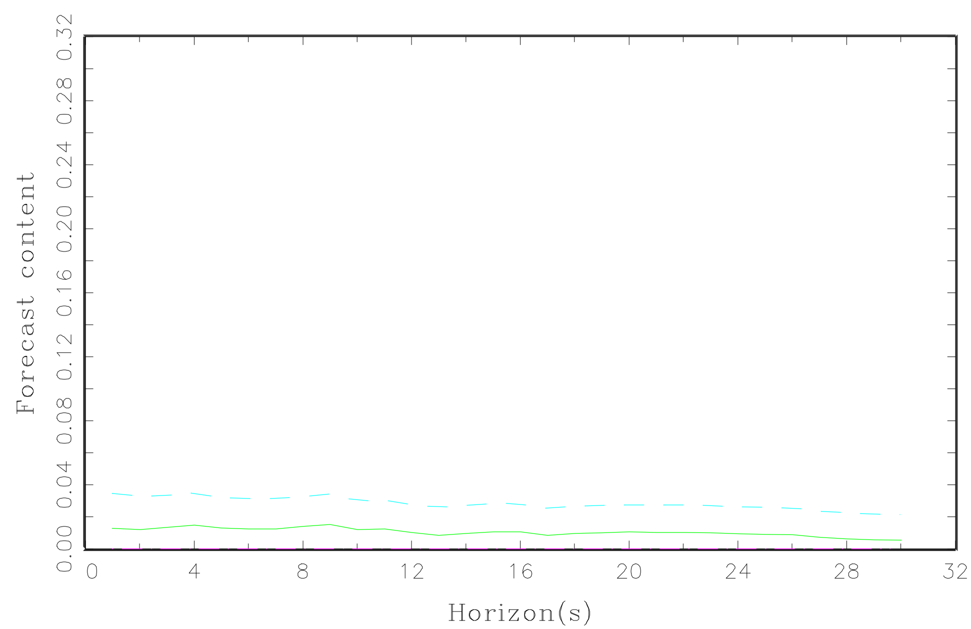


Figure 5a: DM/USD
Forecast content function, AR(12) forecasts,
realized volatility of 5-minute returns

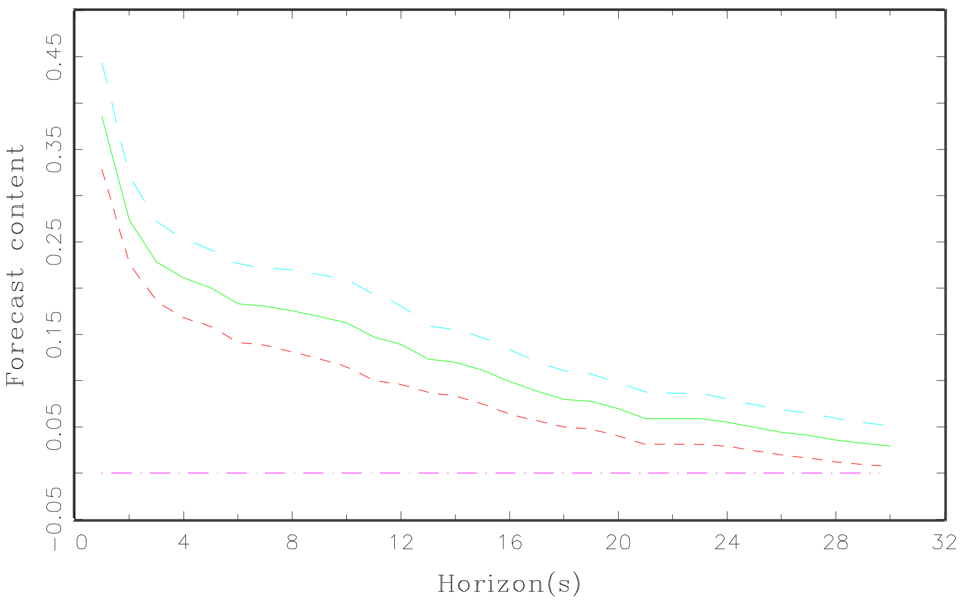


Figure 5b: DM/USD
Forecast content function, AR(12) forecasts,
realized volatility of 10-minute returns

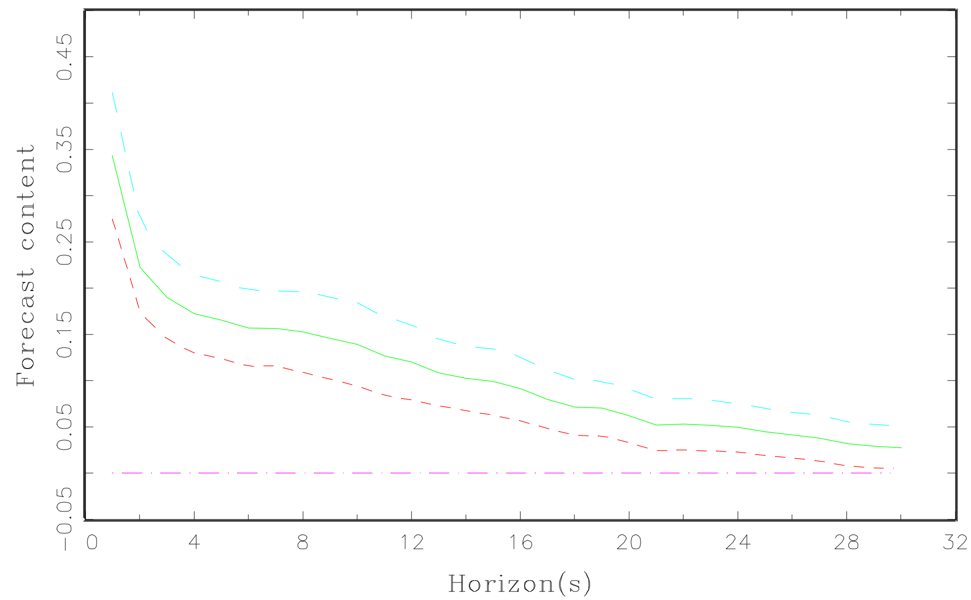


Figure 5c: DM/USD
Forecast content function, AR(12) forecasts,
realized volatility of 30-minute returns

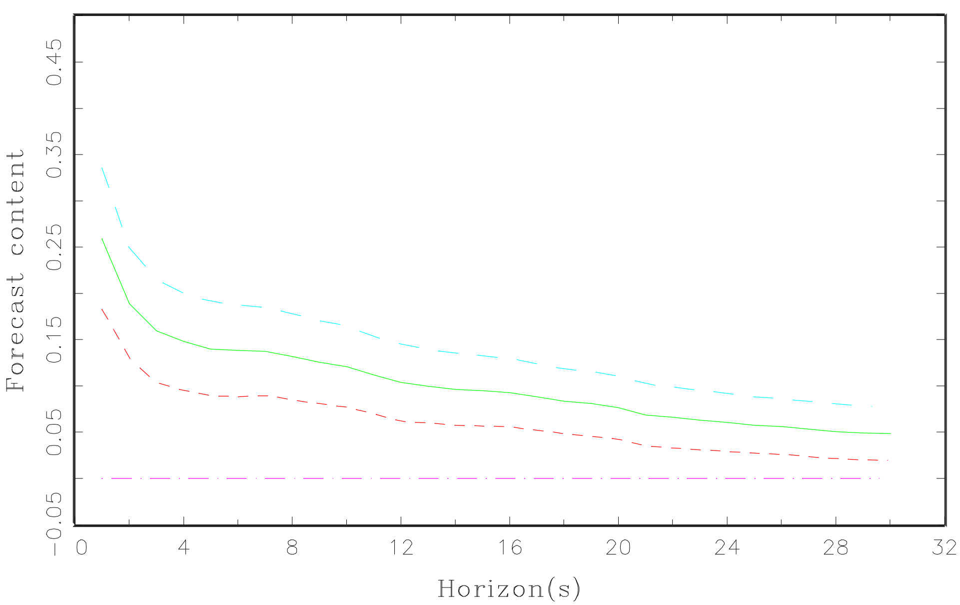


Figure 5d: DM/USD
Forecast content function, AR(12) forecasts,
squared daily returns

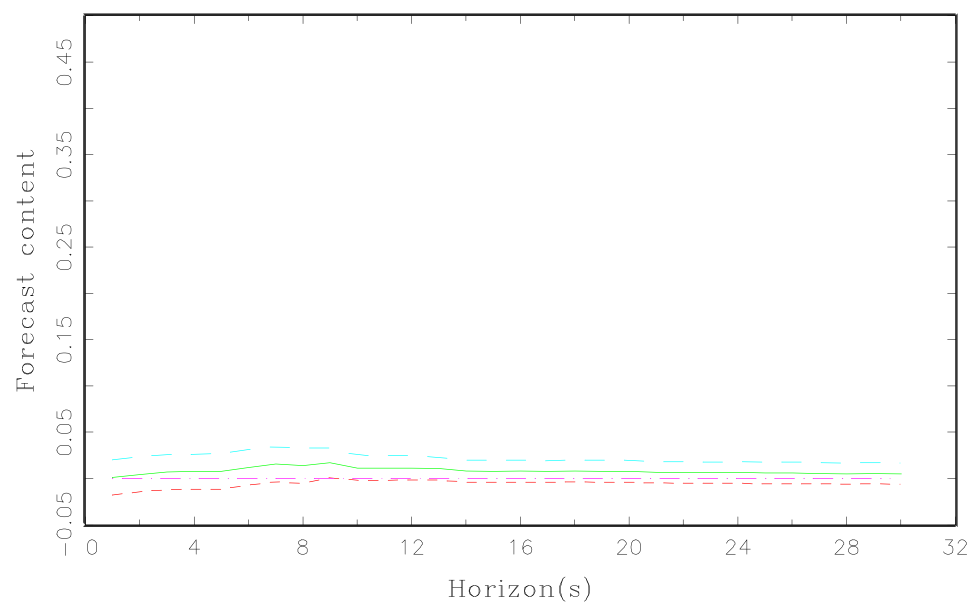


Figure 6a: Yen/USD
Forecast content function, AR(12) forecasts,
realized volatility of 5−minute returns

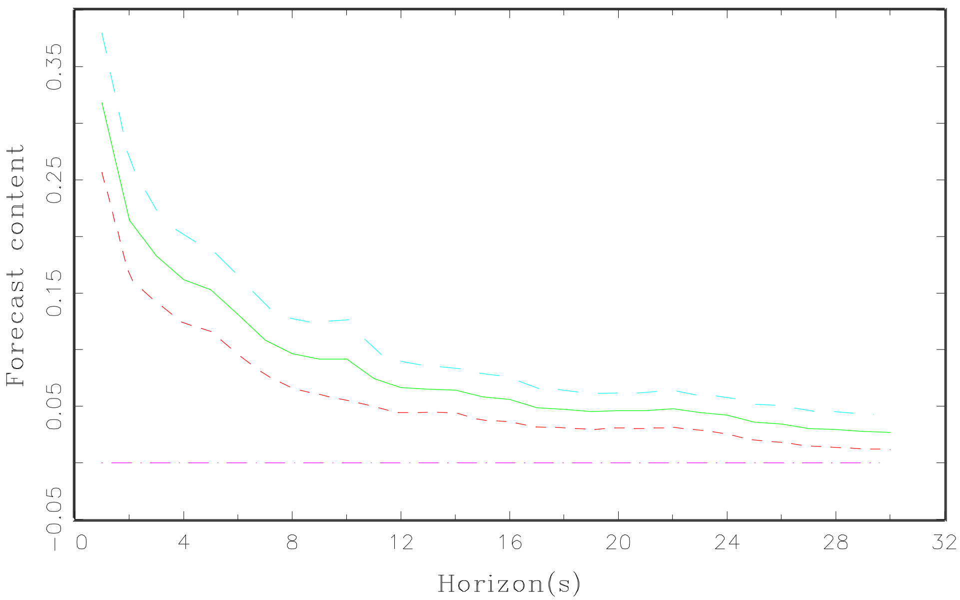


Figure 6b: Yen/USD
Forecast content function, AR(12) forecasts,
realized volatility of 10−minute returns

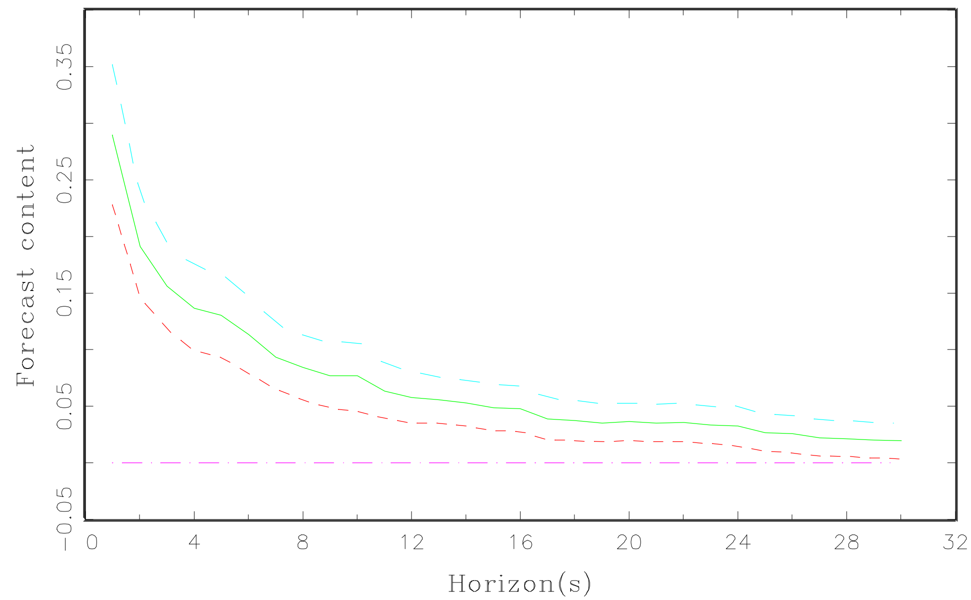


Figure 6c: Yen/USD
Forecast content function, AR(12) forecasts,
realized volatility of 30−minute returns

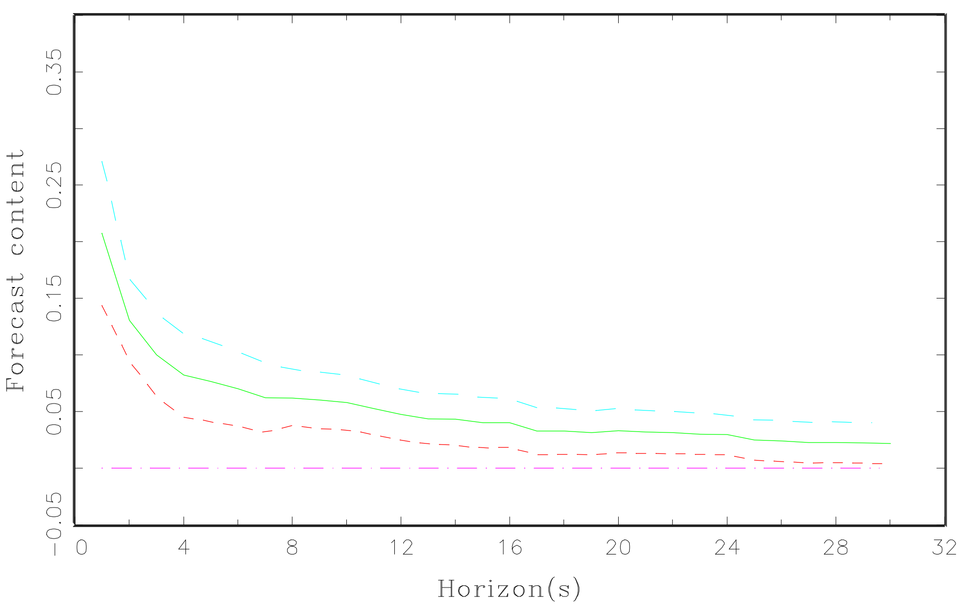
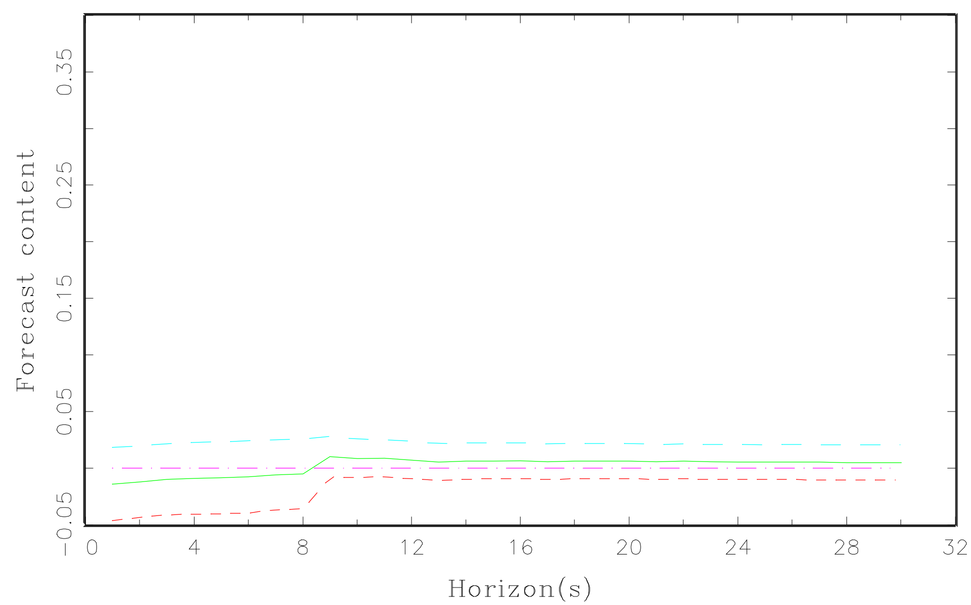


Figure 6d: Yen/USD
Forecast content function, AR(12) forecasts,
squared daily returns



Liste des publications au CIRANO*

Série Scientifique / *Scientific Series* (ISSN 1198-8177)

---

* Consultez la liste complète des publications du CIRANO et les publications elles-mêmes sur notre site Internet :

http://www.cirano.qc.ca/publication/documents.html

2002s-01 An Empirical Analysis of Water Supply Contracts / Serge Garcia et Alban Thomas
2001s-71 A Theoretical Comparison Between Integrated and Realized Volatilities Modeling / Nour Meddahi
2001s-70 An Eigenfunction Approach for Volatility Modeling / Nour Meddahi
2001s-69 Dynamic Prevention in Short Term Insurance Contracts / M. Martin Boyer et Karine Gobert
2001s-68 Serial Cost Sharing in Multidimensional Contexts / Cyril Téjédo et Michel Truchon
2001s-67 Learning from Strike / Fabienne Tournadre et Marie-Claire Villeval
2001s-66 Incentives in Common Agency / Bernard Sinclair-Desgagné
2001s-65 Detecting Mutiple Breaks in Financial Market Volatility Dynamics / Elena Andreou et Eric Ghysels
2001s-64 Real Options, Preemption, and the Dynamics of Industry Investments / Marcel Boyer, Pierre Lasserre, Thomas Mariotti et Michel Moreaux
2001s-63 Dropout, School Performance and Working while in School: An Econometric Model with Heterogeneous Groups / Marcel Dagenais, Claude Montmarquette et Nathalie Viennot-Briot
2001s-62 Derivatives Do Affect Mutual Funds Returns : How and When? / Charles Cao, Eric Ghysels et Frank Hatheway
2001s-61 Conditional Quantiles of Volatility in Equity Index and Foreign Exchange Data / John W. Galbraith, Serguei Zernov and Victoria Zinde-Walsh
2001s-60 The Public-Private Sector Risk-Sharing in the French Insurance "Cat. Nat. System" / Nathalie de Marcellis-Warin et Erwann Michel-Kerjan
2001s-59 Compensation and Auditing with Correlated Information / M. Martin Boyer et Patrick González
2001s-58 Resistance is Futile: An Essay in Crime and Commitment / M. Martin Boyer
2001s-57 The Unreliability of Output Gap Estimates in Real Time / Athanasios Orphanides et Simon van Norden
2001s-56 Exact Nonparametric Two-Sample Homogeneity Tests for Possibly Discrete Distributions / Jean-Marie Dufour et Abdeljelil Farhat
2001s-55 Les coûts de la réglementation : une revue de la littérature / Robert Gagné, Paul Lanoie, Pierre-Carl Micheud et Michel Patry
2001s-54 Testing for structural Change in the Presence of Auxiliary Models / Eric Ghysels et Alain Guay
2001s-53 Environmental Regulation and Productivity: New Findings on the Porter Hypothesis / Paul Lanoie, Michel Patry et Richard Lajeunesse
2001s-52 The Aftermarket Performance of Initial Public Offerings in Canada / Maher Kooli et Jean-Marc Suret
2001s-51 Capital Structure and Risk Management / Karine Gobert
2001s-50 The Underpricing of Initial Public Offerings: Futher Canadian Evidence / Maher Kooli et Jean-Marc Suret
2001s-49 How Innovative Are Canadian Firms Compared to Some European Firms? A Comparative Look at Innovation Surveys / Pierre Mohnen et Pierre Therrien