ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Fernandes, Marcelo; Medeiros, Marcelo C.; Scharth, Marcel

Working Paper

Modeling and predicting the CBOE market volatility index

Texto para discussão, No. 548

Provided in Cooperation with:
Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro

Suggested Citation: Fernandes, Marcelo; Medeiros, Marcelo C.; Scharth, Marcel (2007) : Modeling and predicting the CBOE market volatility index, Texto para discussão, No. 548, Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Departamento de Economia, Rio de Janeiro

This Version is available at:
https://hdl.handle.net/10419/176031

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

TEXTO PARA DISCUSSÃO

No. 548

Modeling and predicting the CBOE market volatility index

Marcelo Fernandes
Marcelo C. Medeiros
Marcel Scharth

DEPARTAMENTO DE ECONOMIA
www.econ.puc-rio.br

Modeling and predicting the CBOE market volatility index

Marcelo Fernandes
Queen Mary University of London
m.fernandes@qmul.ac.uk

Marcelo C. Medeiros
Pontifical Catholic University of Rio de Janeiro
mcm@econ.puc-rio.br

Marcel Scharth
Tinbergen Institute, University of Amsterdam
marcelsfp@gmail.com

Abstract: This paper performs a thorough statistical examination of the time-series properties of the daily market volatility index (VIX) from the Chicago Board Options Exchange (CBOE). The motivation lies on the widespread consensus that the VIX is a barometer to the overall market sentiment as to what concerns investors' risk appetite. Our preliminary analysis suggests that the VIX index displays long-range dependence. This is well line with the strong empirical evidence in the literature supporting long memory in both options-implied and realized variances. We thus resort to both parametric and semiparametric heterogeneous autoregressive (HAR) processes for modeling and forecasting purposes. Our main findings are as follows. First, we confirm the evidence in the literature that there is a strong negative relationship between the VIX index and the S&P 500 index return as well as a positive contemporaneous link with the volume of the S&P 500 index. Second, we find that the VIX index tends to decline as the long-run oil price increases. This is not entirely surprising given the high demand from oil in the last years as well as the recent trend of shorting energy prices in the hedge fund industry. Third, the term spread has no long-run impact in the VIX index despite of the positive contemporaneous link. Fourth, there is some weak evidence that increases in the value of the US dollar tend to move down options-implied market volatility. Finally, we cannot reject the linearity of the above relationships, neither in sample nor out of sample. As for the latter, we actually show that it is pretty hard to beat the pure HAR process because of the very persistent nature of the VIX index. It is not impossible, though. We set out a semiparametric HAR-type model that performs very well across different forecasting horizons by using the above explanatory variables in a quite efficient manner.

JEL classification numbers: G12, C22, C53, E44

Keywords: heterogeneous autoregression, implied volatility, neural network, VIX.

Acknowledgments: The authors are very grateful to Richard Baillie (Editor) as well as to two anonymous referees for valuable comments. Fernandes and Medeiros thank the financial support from the ESRC under the grant RES-062-23-0311 and from CNPq-Brazil, respectively. The usual disclaimer applies, though Scharth is ready to assume the responsibility for any remaining error.

1 Introduction

The Chicago Board Options Exchange (CBOE) computes since 1993 the volatility index VIX to measure market expectations of the near-term volatility implied by stock index option prices. Actually, as from September 2003, the CBOE reports two market volatility indices. The VXO represents the implied volatility of a hypothetical 30-calendar-day at-the-money S&P 100 index option, whereas the VIX hinges on the prices of a portfolio of 30-calendar-day S&P 500 calls and puts with weights being inversely proportional to the squared strike price. The latter thus gauges the expected market volatility by pooling the information from option prices over the whole volatility skew, not just at-the-money strikes as in the VXO index. Moreover, the VIX considers a model-free estimator of the implied volatility, and so it does not depend on any particular option pricing framework. The motivation for using options on the S&P 500 index rather than the S&P 100 lies on the fact that the S&P 500 is the main stock market benchmark in US not only for derivative markets, but also for the hedge fund industry.

This means that the VIX essentially offers a market-determined, forward-looking estimate of one-month stock market volatility (Hentschel, 2003). Most studies in the literature that tackle the information content of implied market volatility employ either VIX or VXO time series. See, among others, Canina and Figlewski (1993), Christensen and Prabhala (1998), Fleming (1998), Blair, Poon and Taylor (2001ab)), Martens and Zein (2004) Koopman, Jungbacker and Hol (2005), and Bandi and Perron (2006). All in all, options-implied volatility is typically more informative than time-series volatility models based on stock market index returns for forecasting purposes, though the latter may sometimes carry incremental information. Jorion (1995), Xu and Taylor (1995), Taylor and Xu (1997), and Martens and Zein (2004) provide similar evidence for foreign exchange markets.

This paper departs from this literature in that we do not attempt to compare the information content of VIX relative to volatility models based on the S&P 500 index returns. In some sense, we restrict attention to a much more basic task, namely, to understand the statistical behavior of the new VIX time series, though we carry out such exercise in a multivariate setting that controls for macroeconomic and financial market conditions. Our motivation lies on the widespread notion that the VIX stands for a barometer to the overall market conditions (Whaley, 2000). High VIX levels typically reflect pessimism, causing equity prices to overshoot on the downside and thus leading to subsequent rallies. In turn, low VIX levels would mirror complacency among market participants, setting up the market for disappointment and raising the likelihood of a market correction.

Our analysis complements well the evidence put forth by Fleming, Ostdiek and Whaley (1995) in their examination of the statistical behavior of the VXO index. They conclude that the daily changes in the VXO index display a slightly first-order positive autocorrelation, whereas weekly changes exhibit significant mean reversion, even if there is no sign of either intraday or intraweek seasonality. In addition, Fleming et al. (1995) also evince a strong negative and asymmetric association with contemporaneous stock market returns. Our findings suggest that the VIX index behaves in a somewhat different manner. First, the contemporaneous relationship between the VIX index and the S&P 500 index returns does not seem to feature any sort of nonlinearity or asymmetric effect. Second, we uncover a strong long-range dependence in the data in line with the long memory that typically characterizes both options-implied and realized volatility measures (Corsi, 2004; Koopman et al., 2005; Bandi and Perron, 2006).

To capture the long-memory in the VIX index, we resort to the family of heterogeneous autoregressive (HAR) processes (Muller, Dacorogna, Dav, Olsen, Pictet and von Weizsacker, 1997; Corsi, 2004; Hillebrand and Medeiros, 2007). Apart from the pure HAR model, we also set out parametric and semiparametric HAR-type processes with additional explanatory variables so as to account for the (contemporaneous and predictive) relationships between the VIX index and key financial and macroeconomic variables. We cover a bit more than the usual suspects. Apart from the changes in the S&P 500 index and volume, we include multiperiod returns on the one-month oil futures contract, the change in the foreign value of the US dollar, the term spread, the credit spread, and the difference between the effective and target Federal Fund rates. These are all linked to different dimensions of the overall market conditions in US. Both oil prices and term spread convey information about the present and future real economic activity, whereas the credit spread relates to the amount of liquidity in the market. The strength of the dollar and the deviation in the Fed rates both reflect to some extent the macroeconomic conditions in US.

The results we obtain are very robust in that average partial effect of the macro-finance variables do not vary much across specifications. Even though accounting for nonlinear dependence matters very little from January 1992 to December 2008, sophistication turns out to pay off in the out of sample analysis. Bayesian regularization avoids overfitting by automatically shrinking insignificant partial effects in the semiparametric HAR model to zero, thereof enabling a remarkable performance across different forecasting horizons. A careful analysis of the average partial effects within the full sample unveils some interesting relationships. As expected, we find a strong negative relationship with both contemporaneous and lagged S&P 500 index returns as well as a positive link with the

contemporaneous S&P 500 volume. It is however a bit surprising that we establish an inverse relationship between the VIX index and long-run movements in the oil price. In addition, the term spread seems to have no long-run effect in the VIX index despite of a contemporaneous positive effect. Finally, the VIX index does not seem to depend either on the deviation in the Fed rates or on the credit spread, whereas there is only weak evidence of a link with the foreign value of the US dollar.

The remainder of this paper is organized as follows. Section 2 provides some background and describes how the CBOE computes the market volatility index. Section 3 discusses the main features of the VIX data so as to shed some light on the specification of the econometric model. Section 4 then evaluates both the in-sample and out-of-sample performances of several heterogeneous autoregressive models. Section 5 offers some concluding remarks.

2 Background for the VIX index

The idea of constructing a volatility index from option prices emerges soon after the introduction of exchange-traded index options in 1973. Gastineau (1977) proposes a volatility index that averages the volatilities implied by at-the-money call options of 14 stocks, whereas Cox and Rubinstein (1985, Appendix 8A) ameliorate the procedure by employing multiple call options on each stock and by weighting the volatilities in such a fashion that the index is at money with a constant time to expiration. The CBOE volatility indices capture the spirit of these earlier efforts, extending the notion in two important directions. First, the VIX hinges on index options rather than stock options. Second, it depends on the implied volatilities of both call and put options. This not only increases the amount of information that the index pools, but also mitigates any eventual bias due to staleness in the observed index level and due to mismeasurement in the riskless rate.

In a nutshell, the VIX index measures the market expectations of the near-term volatility implied by stock index option prices. It features three main differences with respect to the VXO index. First, the VIX index relies on S&P 500 index options with a wide array of strike prices rather than restricting attention to at-the-money strike prices as in VXO. Second, the VIX does not assume the Black-Scholes-Merton option pricing framework, employing a model-free estimator of the implied volatility (Britten-Jones and Neuberger, 2000; Jiang and Tian, 2005). Third, the VIX calculation consider options on the S&P 500 index rather than on the S&P 100 index. This seems much more natural for the S&P 500 is the primary stock market benchmark for both the hedge fund industry and derivative markets.

The model-free estimator of the implied volatility that CBOE employs to calculate the VIX index reads

$$\sigma^2_{\text{CBOE}} = \frac{2}{T} \sum_i \frac{\Delta K_i}{K_i^2} e^{r\,T} Q(K_i) - \frac{1}{T} \left(\frac{F}{K_0} - 1 \right)^2, \tag{1}$$

where T is time to expiration, F is the forward index level derived from the index options prices, K_i is the strike price of the ith out-of-the-money option (either a call if $K_i > F$ or a put if $K_i < F$), $\Delta K_i = (K_{i+1} - K_{i-1})/2$ (for the lowest/highest strike price is the difference between the lowest/highest strike and the next higher/lower strike, respectively), K_0 is the first strike below the forward index level, r is the risk-free interest rate to expiration, and $Q(K_I)$ is the mid-quote for the option with strike of K_i. The VIX index then equals 100 times the options-implied volatility given by σ_{CBOE} in (1). See discussion in Demeterfi, Derman, Kamal and Zou (1999).

The CBOE computes the VIX using primarily the put and call options in the two nearest-term expiration months so as to bracket a 30-day calendar period. At eight days to expiration, the VIX rolls to the second and third contract months to alleviate any sort of pricing anomaly that may occur due to the expiration proximity. For the sake of precision, the CBOE fixes the risk-free rate at $r = 1.162\%$ and measures the time to expiration T in minutes rather than days: $T = T_{SC}/T_Y$, where T_{SD} is the total number of minutes remaining until 8:30 on the settlement day and T_Y refers to the number of minutes in a year.

As for the forward index level, the CBOE assumes that $F = K_* + e^{r\,T_*}\,(C_* - P_*)$, where C_* and P_* are respectively the prices of the 'at-the-money' call and put options with a time to maturity of T_* and a strike price of K_* that minimizes the distance between the call and put prices. Finally, one determines the threshold strike K_0 as the strike price immediately below the forward index level F. The algorithm then sorts all options in ascending order by strike price so as to select only the call/put options with nonzero bid quote and strike price either at or below/above K_0, respectively. To avoid double counting, one must average the mid-quote prices of the call and put options at K_0. The CBOE executes the above calculations for both the near and next term options, resulting in a forward index level and a threshold strike for each term. This ultimately means that the algorithm will end up with estimates of the implied volatility in (1) for the near term options and for the next term options. The single VIX index then stems from a linear interpolation of these two estimates that ensures a constant maturity of 30 days to expiration.

3 Daily behavior of the market volatility index

We examine the daily VIX index for the period running from January 2, 1992 to December 10, 2008. The sample include altogether 4,269 daily observations. We use the full sample for the in-sample analyses, namely, descriptive statistics and contemporaneous modeling, whereas we employ a rolling window of 1,000 observations for the estimation of all predictive regressions. This means that the sample size for the out-of-sample performance evaluation amounts to over 3,200 observations after controlling for starting values.

Figure 4 illustrates the time evolution of the VIX index in the full-sample period. The VIX seems to oscillate in long swings between a quite volatile regime with high index values and a more stable regime with low index values. High volatility characterizes the periods ranging from January to December 1990, from July 1997 to April 2003, and from August 2007 onwards. In contrast, low volatility seems dominant from January 1991 to June 1997 and from April 2004 to July 2007. This is consistent with Whaley's (2000) claim that one may interpret the VIX index as the investors' fear gauge. There are a series of financial crises in the periods featuring a high VIX index, e.g., Asian crisis in 1997, Russian crisis in 1998, Brazilian crisis in 1999, the internet bubble burst in 2000, the 9/11 terrorist attack in 2001, the corporate scandals in 2002, the quantitative long/short equity hedge funds meltdown in the first week of August 2007, and the subsequent credit crunch and global financial crisis.

3.1 Statistical properties of the VIX time series

In this section, we attempt to characterize some of the statistical properties of the daily VIX index. Table 1 documents the results of our preliminary descriptive analyses. In particular, it reports the sample mean, standard deviation, minimum, first quartile, median, third quartile, maximum, and skewness for the VIX index time series as well as the p-value of the Jarque-Bera test for normality. These descriptive statistics do not seem to change much according to the sample despite the seemingly different regimes in Figure 4. The only exception is the skewness coefficient, which substantially increases in the second half of the sample. As expected, the VIX time series is very skewed to the right, leptokurtic, and far from Gaussian. Further analyses show that applying a logarithmic transformation to the VIX index solves the excessive kurtosis, though a good deal of skewness (and hence nonnormality) remains.

Table 1 also evaluates the persistence of the VIX index through a battery of testing procedures. It reports the p-values of the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests for

unit root as well as the values of the KPSS test statistics for the null hypothesis of stationarity. We select the number of lags in the ADF test using the Bayesian information criterion, whereas we run the PP and KPSS tests using the quadratic spectral kernel with Andrews's (1991) bandwidth choice. Finally, Table 1 also displays the rescaled range (R/S) and rescaled variance (V/S) tests for long memory by Lo (1991) and Giraitis, Kokoszka, Leipus and Teyssière (2003), respectively.

We strongly reject the null hypothesis of a unit root for the VIX index with the ADF and PP tests in the first half of the sample as well as in the full sample. In contrast, there is some evidence of nonstationary behavior in the second half of the sample according to the unit root tests. This reflects the rapid increase in the VIX index due to the subprime crisis and the subsequent credit crunch after a period of consistent decline running from April 2003 to April 2006. In contrast, the KPSS test cannot reject the null of stationarity for both subsamples as well as for the full sample. Such a set of mixed results is typical of time series exhibiting long memory. This interpretation is consistent with the results we observe for Lo's (1991) R/S and Giraitis et al.'s (2003) V/S tests given that they both easily reject the null of short memory for the VIX index. The sample autocorrelation and partial autocorrelation functions in Figure 4 corroborates our story in that the VIX series displays a highly persistent nature. The values of the sample autocorrelation function remain highly significant up to lag 500, though the partial correlation function seems to die out very fast. This explains the long swings in Figure 4 as well as the mixed results of the unit-root tests.

3.2 Modeling the volatility index

Corsi (2004) argues that HAR specifications are particularly suitable to modeling and forecasting both realized and implied volatilities because they are able to capture the long-range dependence that arises from the asymmetric propagation of volatility between long and short horizons.[1] The HAR model implicitly assumes an additive cascade of different partial volatilities generated by the actions of distinct types of market participants (Müller, Dacorogna, Dav, Olsen, Pictet and Ward, 1993). At each level of the cascade (or time scale), the corresponding unobserved partial volatility process is a function not only of its past value, but also of the expected values of the other partial volatilities. Corsi shows by straightforward recursive substitutions of the partial volatilities that this additive structure for the volatility cascade leads to a simple restricted linear autoregressive model featuring volatilities realized over different time horizons. The heterogeneous

[1] We indeed find that ARFIMA models perform very poorly for the VIX index both in-sample and out-of-sample. The problem is that ARFIMA models impose a linear form of long memory that depends exclusively on a single parameter, i.e., the fractional integration order (see, e.g., Abadir and Talmain, 2002; Bhardwaj and Swanson, 2006).

nature of the model derives from the fact that, at each time scale, the partial volatility relies on different autoregressive structures.

Let $\bar{y}_t^{(h)} = \frac{1}{h}\sum_{s=1}^{h} y_{t-s+1}$ and define $\mathbf{x}_t = \left(1, \bar{y}_t^{(\iota_1)}, \ldots, \bar{y}_t^{(\iota_p)}\right)' \in \mathbb{R}^{p+1}$ for some vector of indexes $\boldsymbol{\iota} = (\iota_1, \ldots, \iota_p)' \in \mathbb{Z}_+^p$. The time series $\{y_t,\ 1 \leq t \leq T\}$ then follows a HAR model if $y_t = \boldsymbol{\beta}'\mathbf{x}_{t-1} + \varepsilon_t$, where ε_t denotes a generic (weak) white noise. A typical choice in the literature for the index vector is $\boldsymbol{\iota} = (1, 5, 22)'$ so as to mirror the daily, weekly, and monthly components of the volatility process. In this paper, we augment the index vector by also including a biweekly and a quarterly component, so that $\boldsymbol{\iota} = (1, 5, 10, 22, 66)'$.

We consider three variations of the HAR specification. The first includes a set of additional regressors $\mathbf{z}_t$ such that

$$y_t = \boldsymbol{\beta}'\mathbf{x}_{t-1} + \boldsymbol{\gamma}'\mathbf{z}_t + \varepsilon_t, \tag{2}$$

where $\mathbf{z}_t = (z_{1t}, \ldots, z_{kt})$ is a k-dimensional vector of explanatory variables. Among the latter, we include the following macro-finance variables (both contemporaneously and with one lag): the m-day continuously compounded return on the S&P500 index for $m = 1, 5, 10, 22, 66$ (S&P 500 m-day return); the first difference of the logarithm of the volume of the S&P500 index (S&P 500 volume change); the m-day continuously compounded return on the one-month crude oil futures contract (oil m-day return); the first difference of the logarithm of the trade-weighted average of the foreign exchange value of the US dollar index against the Australian dollar, Canadian dollar, Swiss franc, euro, British sterling pound, Japanese yen, and Swedish kroner (USD change); the excess yield of the Moody's seasoned Baa corporate bond over the Moody's seasoned Aaa corporate bond (credit spread); the difference between the 10-Year and 3-month treasury constant maturity rates (term spread); and the difference between the effective and target Federal Fund rates (FF deviation). We refer to (2) as the HARX specification.

The motivation for using S&P 500 returns is to take into account possible leverage effects and asymmetries. Figure 4 evinces not only that the VIX index acts as a good proxy for the S&P 500 volatility, but also that there seems to exist a strong negative link between the VIX index and the S&P 500 index return. This is consistent with the evidence in the literature (see, among others, Fleming et al., 1995; Giot, 2005). We also include multiperiod returns so as to comply with the HAR nature of the model. Given the well-documented positive relationship between volume and volatility (Lamoureux and Lastrapes, 1990), we add the S&P 500 volume change to the set of explanatory variables. The remaining regressors are all linked to different dimensions of the overall market conditions in US. Both oil prices and term spread contain information about the future real

economic activity (Estrella and Hardouvelis, 1991) as well as about future investment opportunities (Petkova, 2006). The credit spread gauges to some extent the amount of liquidity in the market, whereas USD change and FF deviation are both related to the macroeconomic conditions in US.

The second variant is an HAR-type specification that controls for explanatory variables with asymmetric effects. In particular, the AHARX model is given by

$$y_t = \boldsymbol{\beta}'\mathbf{x}_{t-1} + \boldsymbol{\gamma}_0'\mathbf{z}_t^{(-)} + \boldsymbol{\gamma}_1'\mathbf{z}_t^{(+)} + \varepsilon_t, \tag{3}$$

where $\mathbf{z}_t^{(+)} = \{\max(z_{1t}, \bar{z}_1), \ldots, \max(z_{kt}, \bar{z}_k)\}$ and $\mathbf{z}_t^{(-)} = \{\min(z_{1t}, \bar{z}_1), \ldots, \min(z_{kt}, \bar{z}_k)\}$ and $\bar{z}_i$ is the sample average of z_{it} for $1 \leq i \leq k$. Finally, we also consider a semiparametric specification that captures more general forms of nonlinear dependence through a neural network approximation. The motivation rests on the typical success that neural networks experience in the context of volatility modeling and forecasting (Donaldson and Kamstra, 1997; Hu and Tsoukalas, 1999; Hamid and Iqbal, 2004). As in Hillebrand and Medeiros (2007), we specify our neural network HARX (NNHARX) model as

$$y_t = \boldsymbol{\beta}_0'\mathbf{x}_{t-1} + \boldsymbol{\gamma}_0'\mathbf{z}_t + \sum_{m=1}^{M} \frac{\lambda_m}{1 + e^{-\boldsymbol{\beta}_m'\mathbf{x}_{t-1} - \boldsymbol{\gamma}_m'\mathbf{z}_t}} + \varepsilon_t. \tag{4}$$

We estimate the semiparametric NNHARX model using Bayesian regularization with m set either to 3 or 10. The results are very robust to changes in m and hence we report only those corresponding to the more parsimonious model with $m = 3$.[2]

Table 2 reports least-square estimates of the HARX coefficients and their heteroskedasticity-consistent standard errors. In addition, it also documents the corresponding average partial effects and their 95% bootstrap-based confidence intervals within the semiparametric NNHARX model. We omit the AHARX estimates for we find no statistical evidence favoring asymmetric effects in that we cannot reject the equality of $\boldsymbol{\gamma}_0$ and $\boldsymbol{\gamma}_1$. The results are both qualitatively and quantitatively very similar across specifications.[3] In particular, there is a strong negative link between the contemporaneous and lagged S&P 500 index 1-day return and the VIX index, whereas we find no significant influence from the S&P 500 index multiperiod returns. In addition, the positive volume effect is exclusively contemporaneous.

The only oil return that seems to matter is the 66-day continuously compounded return on the one-month crude oil futures contract. Both specifications predict a small negative oil effect in the

[2] This is not surprising given that we find little evidence of nonlinear dependence in our empirical analysis.

[3] The p-values of the Lagrange multiplier tests for autocorrelation up to lag m (with $m = 1, 5, 10, 22$) in Table 2 suggest no evidence of residual autocorrelation regardless of the specification we use. This is reassuring because it ensures the consistency of the coefficient estimates. Further analysis show that, as expected, there is some strong evidence of conditional heteroskedasticity. That is why we employ heteroskedasticity-consistent standard errors.

long run, as measured by the sum of the contemporaneous and lagged coefficients. This is somewhat surprising for we would expect a positive impact given that oil price serves in principle as a proxy for uncertainty in the real economy. However, there is a significative increase in the demand for oil within our sample period, mostly due to China and India thirsty for oil. Most of the recent ups in the oil price are thus demand-driven rather than supply-driven, and so oil return becomes more of a proxy to world economic activity than to uncertainty. In addition, there is also some evidence that the recent decrease in oil price is due to the unwinding of hedge-fund positions in oil futures contract after the banking industry meltdown. Further fueling the collapse in oil prices, hedge funds also start shorting oil exchange-traded funds (ETFs) as a mean to obtain leverage given that they could not short banking ETFs (or any other financial stock) due to the halt of short selling imposed by the US Securities and Exchange Commission. This naturally coincides with a period of extremely high volatility, so that a negative link between oil prices and the VIX index arises.

The linear HARX specification also uncovers a negative contemporaneous relation with the change in the USD index. This is not surprising that a strong dollar typically serves well the US stock market and hence market uncertainty should decline as the value of the US dollar increases. Although of similar magnitude, the NNHARX average partial effect is not significant at the usual significance levels. The term spread seems to affect the VIX index in a positive manner contemporaneously, though the long-run impact is close to zero in both specifications. Finally, both credit spread and the difference between the effective and target rates of the Federal Fund have no impact in the VIX.

In the next section, we turn our attention to predictive regressions that exclude all contemporaneous terms for forecasting purposes. We employ a rolling window of 1,000 observations to estimate the regression coefficients of HARX, HARX, AHARX and NNHARX specifications and then assess their out-of-sample performance in the remaining of the sample by looking at m-day ahead forecasts, with $m \in \{1, 5, 10, 22\}$.

3.3 Forecasting the volatility index

Table 3 displays some descriptive results of the out-of-sample evaluation for forecasts 1, 5, 10, and 22 days ahead, respectively. In particular, we report the mean and standard deviation of the forecast errors as well as the corresponding coefficient of determination (R^2), mean absolute error (MAE), mean squared error (MSE), and the p-value of Hansen's (2005) test of superior predictive ability (SPA) according to the MSE criterion. Apart from the HAR-type models, we also include the results for a random walk with drift as a benchmark.

The random walk with drift on consistently entails the smallest bias regardless of the forecasting horizon. However, the mean forecast error is very close to zero even for the worst specification and hence contributions to the MSE are only very marginal. As for the standard deviation of the forecast errors, the pure HAR model seems to perform very well, confirming persistence as the most prominent feature of the VIX index. It consistently beats the random walk model as well as the HARX and AHARX specifications across the different forecasting horizons, whereas it compares well with the semiparametric NNHARX model for all but the 22-day ahead forecast.

The latter is not surprising. The VIX index measures the market expectations about the risk-neutral volatility 30 calendar days ahead, so that the overlapping implied by the daily frequency contributes to the strong persistence in the data. After 22 trading days (about 30 calendar days), the overlapping effect disappears, reducing persistence and increasing the relative contribution of the macro-finance factors. Accordingly, we also observe a drop of about 16% in the coefficient of determination once we move from 10-day to 22-day forecasts. This is a decline of dramatic proportion if one compares to the reduction of only about 7.5% from 1-day to 5-day forecasts and from 5-day to 10-day forecasts. As persistence subsides, the coefficient of determination is bound to decrease.

The MSE and MAE criteria tell exactly the same story. The NNHARX entails the best 1-day forecast results, even if only marginally better than the pure HAR model, whereas the opposite applies for the 5-day horizon. Their performances are again very similar for the 10-day-ahead forecasts, and so the only striking difference resides in the 22-day forecasts. The results of the SPA test corroborate by a long chalk this evidence, individuating the NNHARX model as the sole model to entail superior predictive ability at the 22-day horizon. In contrast, the HARX and AHARX models perform relatively more poorly in every horizon, reflecting the lack of precision that arises from the relatively much larger number of parameters that we have to estimate.

Although the NNHARX specification seemingly have many more parameters to estimate than the HARX and AHARX models, the Bayesian regularization automatically shrinks the average partial effect of the insignificant coefficients to zero, thereby controlling the precision of the overall estimation/forecasting exercise. Figure 4 illustrates well this point. The average partial effects of the NNHARX model are typically within the confidence bands of the HARX coefficient estimates. This suggests that the NNHARX entails better forecasting ability not because it captures nonlinear effects but because of the regularization procedure that automatically selects which subset of explanatory variables to rely upon. The average partial effect of the change in the USD index indeed

is the only that differs markedly from the one implied by the HARX specification. Another striking feature in Figure 4 relates to the relative instability of the average partial effects over time, even if not very significant in statistical terms.

We complement the above results by running Harvey, Leybourne and Newbold's (1997) variant of Diebold and Mariano's (1995) test for the mean squared forecast error.[4] Table 4 reports the p-values for testing the the null hypothesis that the column and row models perform equally well in terms of mean absolute forecast error. The NNHARX model performs significantly better than any other model at the 1-day and 22-day horizons, though we cannot reject that the random walk and pure HAR models forecast the VIX index 5 and 10 days ahead equally well.

4 Conclusion

This paper examines the time-series properties of the CBOE's market volatility index (VIX) at the daily frequency. The motivation lies on the widespread consensus that the VIX index is a barometer to the overall market sentiment as to what concerns risk appetite. As expected, preliminary analysis unearths strong evidence that the VIX time series displays long-range dependence and so we we employ HAR-type processes for modeling and forecasting purposes. In particular, we employ a pure HAR specification as well as both parametric and semiparametric HAR-type models that also use the information coming from several macro-finance variables. Among the latter, we include multiperiod returns on the S&P 500 index and on the one-month oil futures contract as well as the change in the volume of the S&P 500 index, the credit and term spreads, the change in the foreign value of the US dollar, and the difference between the effective and target Federal Fund rates.

The VIX index does not seem to depend either on the deviation in the Fed rates or on the credit spread. It however holds a very strong negative relationship with both contemporaneous and lagged S&P 500 index returns as well as a positive link with the contemporaneous S&P 500 volume. Market uncertainty is also a decreasing function of oil futures returns in the last quarter, reflecting the fact that oil prices are mostly demand driven in the last years. Although the term spread typically contains information about the future real economic activity and investment opportunities, we find no long-run impact in the VIX index. The HAR models also uncover that the value of the US dollar significantly affects the VIX index in a linear positive fashion, though this effect does not remain significant if one controls for nonlinear dependence of unknown form. Interestingly, this is

[4] We find no qualitative change in the results if we consider Giacomini and White's (2006) conditional predictive ability test using the information set spanned by the lag values of the explanatory variables as well as of the loss-function difference. We employ Newey-West standard errors in both tests so as to account for possible heteroskedasticity in the VIX index.

the only link for which accounting for nonlinearity actually matters. All of the other relationships hold with similar magnitudes regardless of whether we take a semiparametric route.

As per the forecasting results, it turns out that the pure HAR process is a tough cookie to beat because of the very persistent nature of the VIX index. This is partly due to the daily sampling frequency. Given that the VIX index reflects the market expectations about the stock market volatility 30 calendar days ahead, looking at daily figures implies a certain degree of overlapping that exacerbates data persistence. As a consequence, persistence becomes almost the only feature that matters for forecasting purposes at short horizons. In particular, this explains why exploiting the macro-finance information becomes relatively more valuable as the forecasting horizon approaches the 30 calendar days ahead threshold. We nevertheless find that our semiparametric HAR model performs very well across all forecasting horizons, mainly because it automatically selects which macro-finance effect is influential at any given day.

References

Abadir, K. M., Talmain, G., 2002, Aggregation, persistence and volatility in a macro model, Review of Economic Studies 69, 749–779.

Andrews, D. W. K., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, Econometrica 59, 817–858.

Bandi, F. M., Perron, B., 2006, Long memory and the relation between implied and realized volatility, Journal of Financial Econometrics 4, 636–670.

Bhardwaj, G., Swanson, N. R., 2006, An empirical investigation of the usefulness of ARFIMA models for predicting macroeconomic and financial time series, Journal of Econometrics 131, 539–578.

Blair, B. J., Poon, S.-H., Taylor, S. J., 2001a, Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high frequency index returns, Journal of Econometrics 105, 5–26.

Blair, B. J., Poon, S.-H., Taylor, S. J., 2001b, Modelling S&P 100 volatility: The information content of stock returns, Journal of Banking and Finance 25, 1665–1679.

Britten-Jones, M., Neuberger, A., 2000, Option Prices, implied price processes, and stochastic volatility, Journal of Finance 55, 839–866.

Canina, L., Figlewski, S., 1993, The informational content of implied volatility, Review of Financial Studies 6, 659–681.

Christensen, B. J., Prabhala, N. R., 1998, The relation between implied and realized volatility, Journal of Financial Economics 50, 125–150.

Corsi, F., 2004, A simple long memory model of realized volatility, University of Lugano.

Cox, J. C., Rubinstein, M., 1985, Options Markets, Prentice Hall, New Jersey.

Demeterfi, K., Derman, E., Kamal, M., Zou, J., 1999, More than you ever wanted to know about volatility swaps, Journal of Derivatives 6, 9–32.

Diebold, F. X., Mariano, R. S., 1995, Comparing predictive accuracy, Journal of Business and Economic Statistics 13, 253–263.

Donaldson, R. G., Kamstra, M., 1997, An artificial neural network-GARCH model for international stock return volatility, Journal of Empirical Finance 4, 17–46.

Estrella, A., Hardouvelis, G., 1991, The term structure as a predictor of real economic activity, Journal of Finance 46, 555–576.

Fleming, J., 1998, The quality of market volatility forecasts implied by S&P 100 index option prices, Journal of Empirical Finance 5, 317–345.

Fleming, J., Ostdiek, B., Whaley, R. E., 1995, Predicting stock market volatility: A new measure, Journal of Futures Markets 15, 265–302.

Gastineau, G. L., 1977, An index of listed option premiums, Financial Analysts Journal 30, 70–75.

Giacomini, R., White, H., 2006, Tests of conditional predictive ability, Econometrica 74, 1545–1578.

Giot, P., 2005, Relationships between implied volatility indices and stock index returns, Journal of Portfolio Management 31, 92–100.

Giraitis, L., Kokoszka, P., Leipus, R., Teyssière, G., 2003, Rescaled variance and related tests for long memory in volatility and levels, Journal of Econometrics 112, 265–294.

Hamid, S. A., Iqbal, Z., 2004, Using neural networks for forecasting volatility of S&P 500 Index futures prices, Journal of Business Research 57, 1116–1125.

Hansen, P. R., 2005, A test for superior predictive ability, Journal of Business and Economic Statistics 23, 365–380.

Harvey, D., Leybourne, S., Newbold, P., 1997, Testing the equality of prediction mean squared errors, International Journal of Forecasting 13, 281–291.

Hentschel, L., 2003, Errors in implied volatility estimation, Journal of Financial and Quantitative Analysis 38, 779–810.

Hillebrand, E., Medeiros, M. C., 2007, The benefits of bagging for forecast models of realized volatility, forthcoming in Econometric Reviews.

Hu, M. Y., Tsoukalas, C., 1999, Combining conditional volatility forecasts using neural networks: An application to the EMS exchange rates, Journal of International Financial Markets, Institutions and Money 9, 407–422.

Jiang, G., Tian, Y., 2005, Model-free implied volatility and its information content, Review of Financial Studies 18, 1305–1342.

Jorion, P., 1995, Predicting volatility in the foreign exchange market, Journal of Finance 50, 507–528.

Koopman, S. J., Jungbacker, B., Hol, E., 2005, Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements, Journal of Empirical Finance 12, 445–475.

Lamoureux, C. G., Lastrapes, W. D., 1990, Heteroskedasticity in stock return data: Volume versus GARCH effects, Journal of Finance 45, 221–229.

Lo, A. W., 1991, Long-term memory in stock market prices, Econometrica 59, 1279–1313.

Martens, M., Zein, J., 2004, Predicting financial volatility: High-frequency time-series forecasts vis-à-vis implied volatility, Journal of Futures Markets 24, 1005–1028.

Muller, U., Dacorogna, M., Dav, R., Olsen, R., Pictet, O., von Weizsacker, J., 1997, Volatilities of different time resolutions – analysing the dynamics of market components, Journal of Empirical Finance 4, 213–239.

Müller, U., Dacorogna, M., Dav, R., Olsen, R., Pictet, O., Ward, J., 1993, Fractals and intrinsic time – A challenge to econometricians, Proceedings of the XXXIX International AEA Conference on Real Time Econometrics.

Petkova, R., 2006, Do the Fama-French factors proxy for innovations in predictive variables?, Journal of Finance 61, 581–612.

Taylor, S. J., Xu, X., 1997, The incremental volatility information in one million foreign exchange quotations, Journal of Empirical Finance 4, 317–340.

Whaley, R. E., 2000, The investor fear gauge, Journal of Portfolio Management 26, 12–17.

Xu, X., Taylor, S. J., 1995, Conditional volatility and the informational efficiency of the PHLX currency options markets, Journal of Banking and Finance 19, 803–821.

Table 1: **Descriptive statistics for the logarithm of the VIX index**

The sample period runs from January 2, 1990 to December 10, 2008, including altogether 4,269 time-series observations. We report the sample mean, median, minimum, maximum, standard deviation, skewness, and kurtosis for the logarithm of the VIX time series, as well as the p-values of the Jarque-Bera test for normality and of the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests for unit root. In addition, we also report the values of the KPSS test statistics for the null hypothesis of stationarity, whose critical values are 0.347, 0.463, and 0.739 at the 10%, 5%, and 1% significance levels, respectively. We select the number of lags in the ADF test using the Bayesian information criterion, whereas we carry out the PP and KPSS tests using the quadratic spectral kernel with Andrews's (1991) bandwidth choice. Finally, R/S and V/S refer to the rescaled range and rescaled variance tests for long memory by Lo (1991) and Giraitis et al. (2003), respectively. The critical values of the R/S test are 1.747 and 2.098 at the 5% and 0.5% levels, whereas the critical values for the V/S test are 1.36 and 1.63 at the 5% and 1% levels, respectively.

sample statistics	first half	second half	full sample
mean	2.8539	2.9523	2.9031
median	2.8151	2.9538	2.8904
minimum	2.2311	2.2915	2.2311
maximum	3.8230	4.3927	4.3927
standard deviation	0.3097	0.3820	0.3512
skewness	0.4206	0.6142	0.6247
kurtosis	2.4189	3.4931	3.6854
Jarque-Bera	0.0000	0.0000	0.0000
ADF	0.0152	0.1667	0.0084
PP	0.0022	0.1029	0.0006
KPSS	0.6714	0.1943	0.2136
R/S	3.7947	3.5891	3.8032
V/S	6.4433	7.7384	6.9482

Table 2: **Modeling the logarithm of the VIX index**

The sample period runs from January 2, 1990 to December 10, 2008, including altogether 4,269 time-series observations. The first column lists the additional regressors we use apart from the day-of-the-week dummies and the average of the logarithm of the VIX index over the last $k \in \{1, 5, 10, 22, 66\}$ days. S&P500 k-day return is the k-day log-return on the S&P500 index; S&P500 volume change is the first difference of the logarithm of the volume of the S&P500 index; oil k-day return is the k-day log-return on the one-month crude oil futures contract; USD change is the first difference of the logarithm of the foreign exchange value of the US dollar index; credit spread is the excess yield of the Moody's seasoned Baa corporate bond over the Moody's seasoned Aaa corporate bond; term spread is the difference between the 10-Year and 3-month treasury constant maturity rates; and FF deviation is the difference between the effective and target Federal Funds rates. For the parametric specifications, we provide the point estimates for the coefficients as well as their heteroskedasticity-consistent standard errors within parentheses, whereas we report average partial effects with the corresponding 95% bootstrap-based confidence intervals for the semiparametric NNHARX model.

	HARX		NNHARX	
	contemporaneous	past	contemporaneous	past
S&P500 1-day return	−3.7395 (0.1240)	−0.1432 (0.0659)	−3.7255 [−4.9230,−3.2908]	−0.1504 [−0.3593,−0.0072]
S&P500 5-day return	0.0998 (0.0640)	−0.1184 (0.0608)	0.0217 [−0.1853,0.1920]	−0.0231 [−0.2111,0.2099]
S&P500 10-day return	0.0421 (0.0606)	−0.0195 (0.0563)	0.0163 [−0.1258,0.1522]	−0.0163 [−0.1593,0.1234]
S&P500 22-day return	−0.0203 (0.0566)	−0.0220 (0.0546)	−0.0457 [−0.1576,0.0608]	0.0019 [−0.1156,0.1161]
S&P500 66-day return	0.0510 (0.0591)	−0.0404 (0.0584)	0.0140 [−0.0789,0.1072]	0.0020 [−0.0808,0.0958]
S&P500 volume change	0.0190 (0.0033)	0.0003 (0.0033)	0.0197 [0.0110,0.0288]	0.0001 [−0.0065,0.0070]
oil 1-day return	−0.0158 (0.0640)	0.0239 (0.0302)	−0.0167 [−0.1643,0.1506]	0.0335 [−0.0297,0.0972]
oil 5-day return	0.0228 (0.0301)	−0.0212 (0.0294)	0.0171 [−0.0619,0.0810]	−0.0096 [−0.0803,0.0659]
oil 10-day return	0.0035 (0.0296)	−0.0032 (0.0275)	−0.0027 [−0.0769,0.0757]	0.0006 [−0.0708,0.0721]
oil 22-day return	−0.0138 (0.0276)	0.0148 (0.0269)	−0.0136 [−0.0640,0.0344]	0.0126 [−0.0315,0.0648]
oil 66-day return	0.0473 (0.0277)	−0.0505 (0.0274)	0.0452 [−0.0001,0.0911]	−0.0471 [−0.0963,−0.0011]
USD change	−0.3179 (0.1459)	0.0476 (0.1455)	−0.2508 [−0.5780,0.1068]	0.0023 [−0.3396,0.3240]
credit spread	−0.0143 (0.0325)	0.0154 (0.0327)	−0.0014 [−0.0763,0.0868]	0.0010 [−0.0805,0.0766]
term spread	0.0248 (0.0084)	−0.0263 (0.0084)	0.0188 [−0.0174,0.0443]	−0.0202 [−0.0456,0.0154]
FF deviation	0.0023 (0.0036)	−0.0015 (0.0035)	0.0029 [−0.0064,0.0112]	−0.0013 [−0.0107,0.0061]
LM test for autocorrelation up to lag m (p-value)				
$m = 1$	0.9053		0.3645	
$m = 5$	0.2573		0.6333	
$m = 10$	0.3613		0.5826	
$m = 22$	0.1060		0.3939	

Table 3: **Forecasting performance at different horizons**

The sample period runs from January 2, 1990 to December 10, 2008, adding up to 4,269 observations. We use a rolling window of 1,000 time-series observations to estimate the different models and then perform out-of-sample forecasting evaluation in the remaining of the series. We consider the following specifications: random walk with drift (RW), heterogeneous autoregression (HAR), heterogeneous autoregression with exogenous variables (HARX), heterogenous autoregression with exogenous variables and asymmetric effects (AHARX), and the neural-network heterogeneous autoregression with exogenous variables (NNHARX). We gauge forecasting performance by means of the mean forecast error (MFE), the mean squared forecast error (MSE), the mean absolute forecast error (MAE), and the coefficient of determination (R^2). We also report the p-value of Hansen's (2005) test of superior predictive ability (SPA) in terms of the MSE criterion.

	MFE	MSE	SPA	MAE	R^2
ONE DAY AHEAD					
RW	0.0003	0.0035	0.0280	0.0439	0.9715
HAR	0.0019	0.0035	0.4670	0.0436	0.9720
HARX	0.0017	0.0038	0.0000	0.0463	0.9689
AHARX	0.0020	0.0039	0.0000	0.0467	0.9682
NNHARX	0.0018	0.0034	0.7990	0.0428	0.9723
FIVE DAYS AHEAD					
RW	0.0018	0.0127	0.0180	0.0850	0.8969
HAR	0.0073	0.0121	0.8135	0.0837	0.9024
HARX	0.0073	0.0126	0.0000	0.0848	0.8983
AHARX	0.0080	0.0128	0.0355	0.0857	0.8967
NNHARX	0.0079	0.0124	0.4005	0.0839	0.9002
TEN DAYS AHEAD					
RW	0.0037	0.0199	0.1035	0.1073	0.8384
HAR	0.0118	0.0191	0.7970	0.1055	0.8463
HARX	0.0102	0.0197	0.0080	0.1065	0.8410
AHARX	0.0113	0.0197	0.2580	0.1073	0.8409
NNHARX	0.0119	0.0193	0.1052	0.1052	0.8445
TWENTY-TWO DAYS AHEAD					
RW	0.0088	0.0379	0.0020	0.1462	0.6929
HAR	0.0196	0.0356	0.0245	0.1439	0.7144
HARX	0.0163	0.0354	0.0010	0.1394	0.7149
AHARX	0.0164	0.0348	0.0325	0.1386	0.7196
NNHARX	0.0248	0.0329	0.8350	0.1322	0.7381

Table 4: **Diebold-Mariano tests for the mean absolute forecast error**

The sample period runs from January 2, 1990 to December 10, 2008, adding up to 4,269 observations. We use a rolling window of 1,000 time-series observations to estimate the different models and then perform out-of-sample forecasting evaluation in the remaining of the series. We consider the following specifications: random walk with drift (RW), heterogeneous autoregression (HAR), heterogeneous autoregression with exogenous variables (HARX), heterogenous autoregression with exogenous variables and asymmetric effects (AHARX), and the neural-network heterogeneous autoregression with exogenous variables (NNHARX). The p-values in each entry correspond to the modified Diebold-Mariano test for the null hypothesis that the column and row models perform equally well in terms of mean absolute forecast error.

	RW	HAR	HARX	AHARX	NNHARX
ONE DAY AHEAD					
RW		0.1158	0.0000	0.0000	0.0008
HAR	0.1158		0.0000	0.0000	0.0014
HARX	0.0000	0.0000		0.0028	0.0000
AHARX	0.0000	0.0000	0.0028		0.0000
NNHARX	0.0008	0.0014	0.0000	0.0000	
FIVE DAYS AHEAD					
RW		0.1313	0.4477	0.3303	0.2246
HAR	0.1313		0.1506	0.0445	0.4129
HARX	0.4477	0.1506		0.1084	0.0008
AHARX	0.3303	0.0445	0.1084		0.0093
NNHARX	0.2246	0.4129	0.0008	0.0093	
TEN DAYS AHEAD					
RW		0.2087	0.3914	0.4922	0.2106
HAR	0.2087		0.3092	0.2038	0.4398
HARX	0.3914	0.3092		0.2771	0.0068
AHARX	0.4922	0.2038	0.2771		0.0629
NNHARX	0.2106	0.4398	0.0068	0.0629	
TWENTY-TWO DAYS AHEAD					
RW		0.3300	0.1160	0.0979	0.0045
HAR	0.3300		0.1290	0.1148	0.0030
HARX	0.1160	0.1290		0.3776	0.0035
AHARX	0.0979	0.1148	0.3776		0.0200
NNHARX	0.0045	0.0030	0.0035	0.0200	

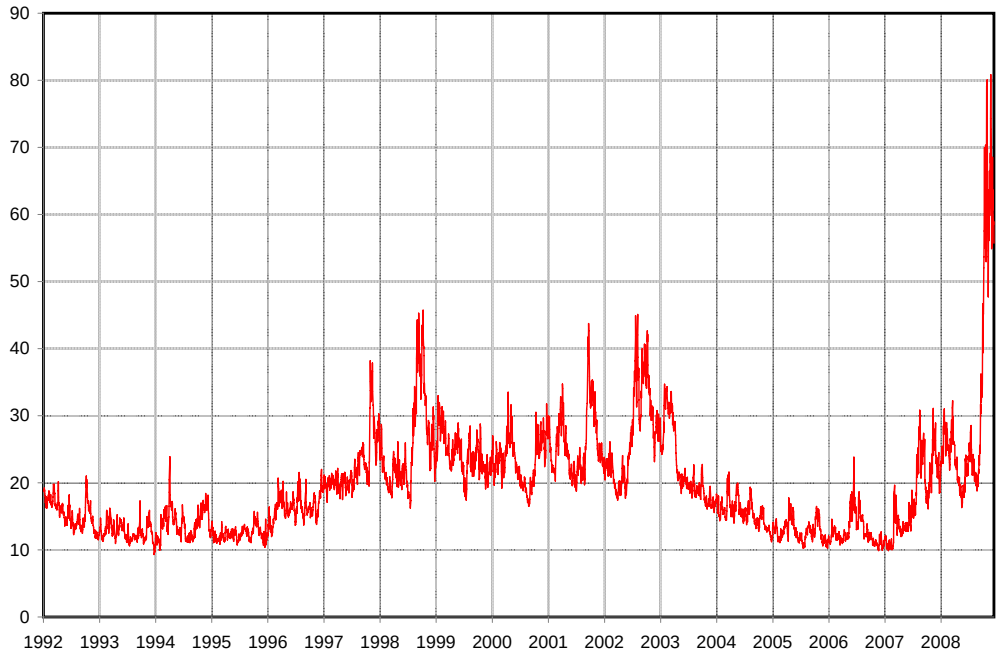

Figure 1: The daily VIX index from January 2, 1992 to December 10, 2008.

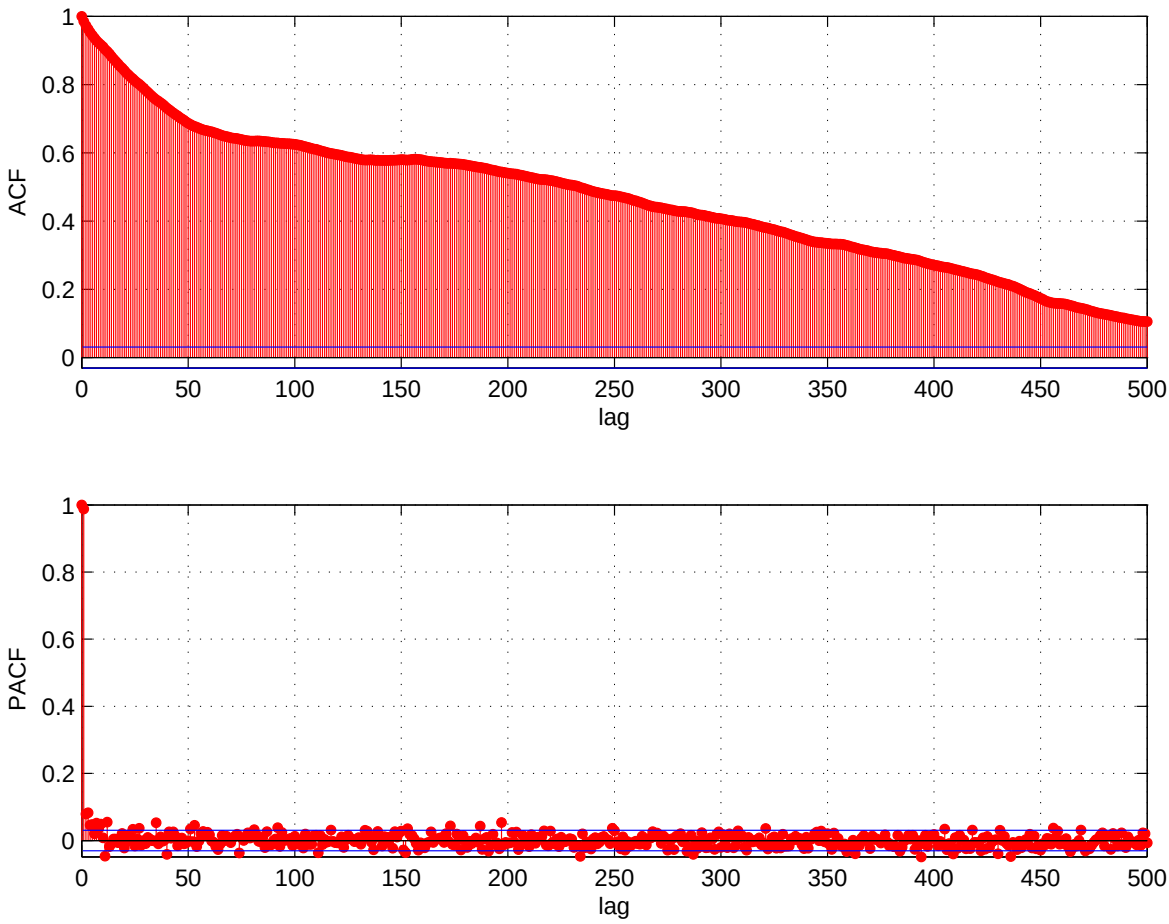

Figure 2: Sample autocorrelation and partial autocorrelation functions of the logarithm of the VIX index from January 2, 1990 to December 10, 2008. The blue line refers to the 95% confidence interval under the null of zero autocorrelation.

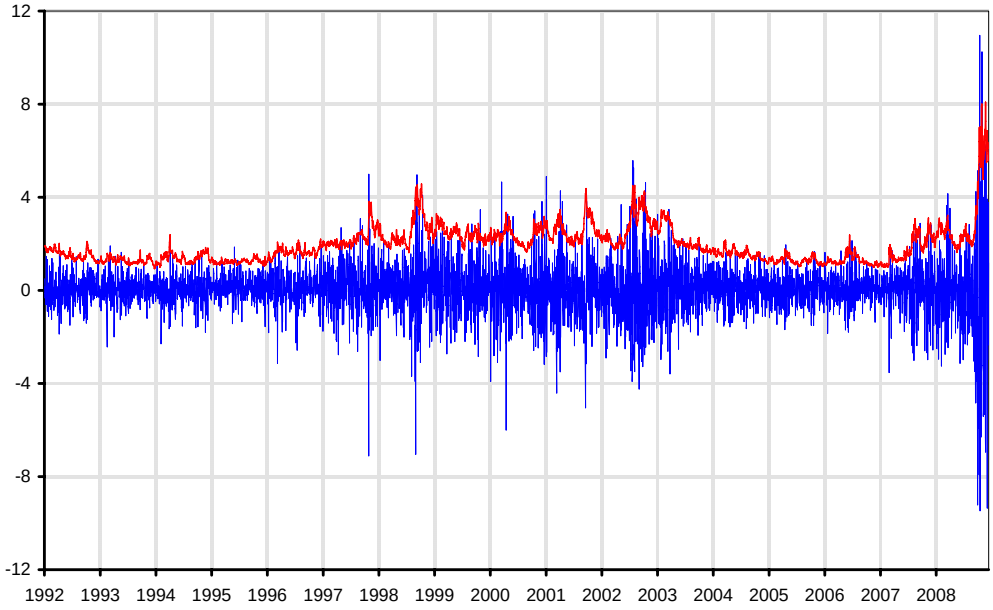

Figure 3: The relationship between the S&P 500 index returns and the VIX index (divided by 10) from January 2, 1990 to December 10, 2008.

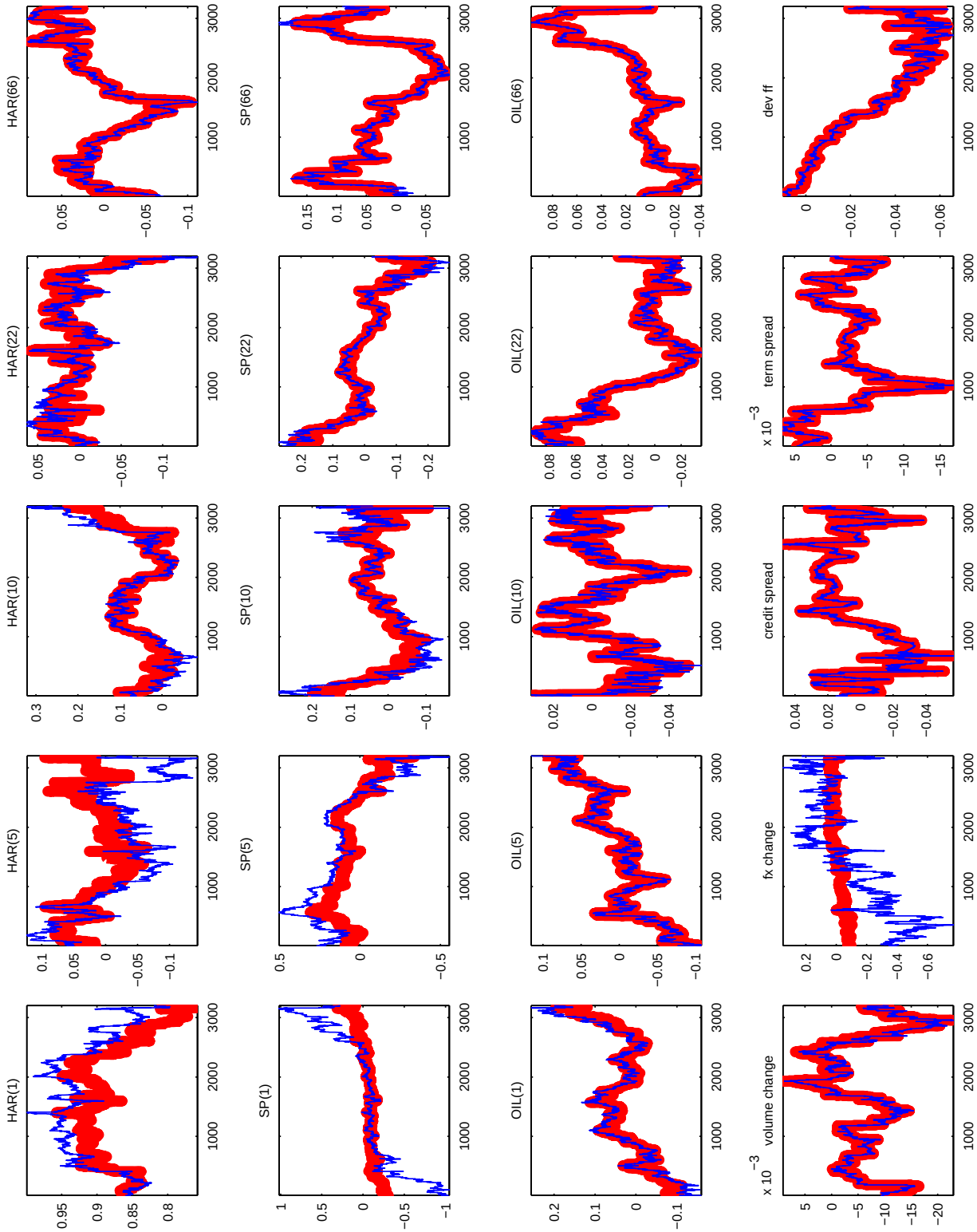

Figure 4: Average partial effects implied by the NNHARX model (in blue) as compared to the corresponding confidence intervals of the HARX coefficient estimates (in red).

Departamento de Economia PUC-Rio
Pontifícia Universidade Católica do Rio de Janeiro
Rua Marques de Sâo Vicente 225 - Rio de Janeiro 22453-900, RJ
Tel.(21) 35271078 Fax (21) 35271084
www.econ.puc-rio.br
flavia@econ.puc-rio.br