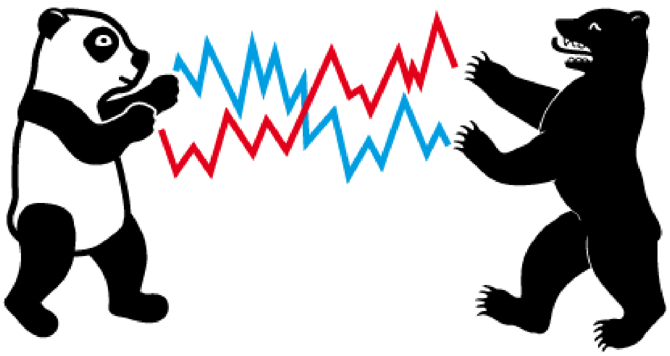
IRTG 1792 Discussion Paper 2018-046

Forecasting the Term Structure of Option Implied Volatility: The Power of an Adaptive Method

Ying Chen *
Qian Han *2
Linlin Niu *2



* National University of Singapore, Singapore
*2 Xiamen University, Republic of China

This research was supported by the Deutsche Forschungsgemeinschaft through the International Research Training Group 1792 "High Dimensional Nonstationary Time Series".

http://irtg1792.hu-berlin.de
ISSN 2568-5619

International Research Training Group 1792

Forecasting the Term Structure of Option Implied Volatility: The Power of an Adaptive Method

Ying Chen[a] Qian Han[b,c] Linlin Niu[b,c,d*]

[a]Department of Statistics & Applied Probability, National University of Singapore

[b]Wang Yanan Institute for Studies in Economics (WISE), Xiamen University

[c]MOE Key Laboratory of Econometrics, Xiamen University

[d]Gregory and Paula Chow Center for Economic Research, Xiamen University

August 2018

Abstract

We model the term structure of implied volatility (TSIV) with an adaptive approach to improve predictability, which treats dynamic time series models of globally time-varying but locally constant parameters and uses a data-driven procedure to find the local optimal interval. We choose two specifications of the adaptive models: a simple local AR (LAR) model for a univariate implied volatility series and an adaptive dynamic Nelson-Siegel (ADNS) model of three factors, each based on an LAR, to model the cross-section of the TSIV simultaneously with parsimony. Both LAR and ADNS models uniformly outperform more than a dozen alternative models with significance across maturities for 1-20 day forecast horizons. Measured by RMSE and MAE, the forecast errors of the random walk model can be reduced by between 20% and 60% for the 5 to 20 days ahead forecast. In terms of prediction accuracy of future directional changes, the adaptive models achieve an accuracy range of 60%-90%, which strictly dominates the range of 30%-59% of the alternative models.

Keywords: Term structure of implied volatility, local parametric models, forecasting

JEL Classification: C32, C53

*Corresponding author. Linlin Niu, Rm A306, Economics Building, Xiamen University, Xiamen, 361005, Fujian, China. Email: llniu@xmu.edu.cn. Phone: +86-592-2182839. Fax: +86-592-2187708.

1 Introduction

Option prices contain rich information about future underlying asset returns as well as volatilities. Canina and Figlewski (1993) and Jiang and Tian (2005) find that the implied volatility of the at-the-money option is a more efficient forecast of realized volatility than volatility measures based on historical data. Breeden and Litzenberger (1978) and voluminous following studies show that, for a given maturity, the entire risk-neutral distribution of the underlying asset returns is foreseen in option prices across the strikes as market investors with heterogeneous expectations about the future price movement trade at different strike prices.

In comparison, the information content of implied volatilities along the maturity dimension, i.e., the term structure of implied volatilities (TSIV), is less explored. Stein (1989) finds that option implied volatilities of long terms overreact to shocks to the short term volatility, which is inconsistent with the expectation hypothesis. Since then, Diz and Finucane (1993), Heynen et al. (1994), Campa and Chang (1995), Byoun et al. (2003), and Mixon (2007) debate whether the expecation hypothesis holds for the implied volatility term structure and do not reach consensus. Some studies conclude that implied volatilities with different durations do behave differently (see, for example, Xu and Taylor, 1994; Christoffersen et al., 2008; Guo et al., 2014). However, it remains unclear how useful the information from the entire term structure is for out-of-sample prediction and how performance varies for different forecast horizons.

We examine this issue using an adaptive approach. The adaptive approach is first proposed in Mercurio and Spokoiny (2004) to forecast volatility where the basic model is simply a varying mean process, and is shown to outperform the more sophisticated GARCH model in the out-of-sample prediction. The approach takes a non-parametric view of the persistence of time series and considers nonstationarity to be generated by time-varying structures in the underlying models. Time series possessing these properties can be observationally equivalent to processes generated by long memory models or models with sophisticated time-varying mechanisms such as clustering properties, abrupt or smooth structural breaks, or regime-switchs, etc., see Diebold and Inoue (2001) and Granger and Hyung (2004) for examples. Chen et al. (2010) extend the approach further with a local AR (LAR) model with time-varying parameters and demonstrate its superiority in out-of-sample forecasts for realized volatility against several popular long memory models and regime switching models. Moreover, Giacomini et al. (2009) estimate portfolio risk with the time-varying copula model. Spokoiny et al. (2013) extend the local approach to a quantile regression to investigate the tail dependence of the Hong Kong stock market and to analyze the distributions of the risk factors of temperature dynamics. Chen and Niu (2014) apply an LAR model with exogenous variables (LARX) to successfully forecast the yield curve. Haerdle et al. (2015) develop a local adaptive multiplicative error model for high-frequency forecasts. In these works, choosing a sophisticated yet simple model as the local model is the key for successful application, especially for forecast performance.

Recognizing that implied volatility, as with many other financial time series, is often highly persistent with nonstationary features, we treat the series as globally time-changing, but locally stationary, processes. We apply a LAR model as a basic specification to forecast the implied volatility of a particular term and strike price, and an adaptive dynamic term structure model built on the LAR as an extended specification for the TSIV of a strike price, one of the Adaptive Dynamic Nelson-Siegel (ADNS) models without exogenous macrovariables used in Chen and Niu (2014) for yield curve modeling. More specifically, we propose two specifications of the adaptive model. The first models the single implied volatility as a univariate time series; the other models the entire term structure given a strike price to explore information across maturity. Mimicking a real-time forecasting senario, at each point in time, we model the time series parametrically with past historic information. We assume that the parameters are approximately constant (homogeneous), but only up to a point beyond which homogeneity will be rejected by a statistical test. Once the test procedure selects the longest possible homonegeous interval, the prediction is made, assuming that the homogeneity will remain within the forecast horizon. This adaptive approach, thus, strikes a balance between information efficiency and stationarity concerns. It not only uses the longest sample possible under homogeneity to increase information efficiency, but also limits the sample to a properly chosen interval to reduce parameter instability. As comparisons, we use thirteen models assessed in Guo et al. (2018).

The results are encouraging. The adaptive models dramatically outperform the alternative models. The ADNS model is more parsimoneous with higher efficiency for the whole TSIV than the LAR model for a single IV. While Guo et al. (2018) find their models may slightly improve predictability against the random walk benchmark model up to the 5-day horizon, we find that our two adaptive models strongly improve predictability against their models up to the 20-day horizon; the longer the horizon, the better the performance. The key driving force is the adaptive process used. The LAR for a single IV and ADNS for the TSIV perform almost equally well, and both, remarkably, outperform the random walk model in all combinations of maturity and horizon. The ADNS parsimonously combines cross-section information in three factor LAR prosesses without modeling each maturity separately.

In a striking comparison to the marginal improvement of other models relative to the random walk, the ADNS significantly reduces forecasting errors for almost all maturities. For example, when it comes to forecast the 20-day ahead implied volatility for 730-day maturity calls, the root mean squared errors (RMSE) and mean absolute errors (MAE) for the ADNS are only 59% and 46% of those of the random walk model, while the best forecast among the other models does not beat the random walk model. Moreover, no single model, except the ADNS, consistently beats the benchmark random walk model across all forecast horizons. Results of out-of-sample R^2 also demonstrate that only the adaptive models can outperform random walk on 5- to 20-days forecast horizons. We further verify the statistical significance of the adaptive models against all models 5- to 20-days forecast horizons by computing the Diebold-Mariano (DM) (Diebold and Mariano, 1995) and the Clark-West (CW) (Clark and West, 2006, 2007) test statistics. In terms of the directional forecast of future changes, while alternative models accurately predict between 30% and 59% of the time, the adaptive models have prediction accuracy ranged between 60% and 90%.

The clear dominance of our adaptive models over other models suggests the power of adaptive forecasting with suitable local models; prediction accuracy increases further for longer maturities and longer horizons.

Our work is closely related to the literature on implied volatility surface (IVS) modeling. Mapping option price quotes in terms of implied volatility against the maturity and strike price dimensions generates the IVS. For practitioners, forecasting the IVS has become more and more important for risk management and in developing trading strategies, and is crucial in the market of volatility derivatives. In the academic literature, the joint dynamics of the IVS are often factorized along both dimensions of strike price and maturity, as shown by Gonclaves and Guidolin (2006) and Neumann and Skiadopoulos (2013), among many others. The strike price dimension, widely known as the option smile, has received much more attention than the maturity dimension. The sources and implications of the implied volatility smile on the underlying asset returns are often emphasized. For example, Bollen and Whaley (2004) proposes net hedging demand as an alternative explanation of the option smile beyond stochastic volatility and jumps in asset prices. In this paper, we focus on the information content embodied in the TSIV in terms of out-of-sample forecasting. Understanding the role of the TSIV in implied volatility forecasting is necessary and helpful for IVS modeling.

Our study is also related to the emerging literature on the VIX and VIX derivatives. Luo and Zhang (2012) find that the VIX with different times to maturity subsumes all information contained in historical volatilities. However, they do not consider using the information of the entire term structure to predict future implied volatilities. Zhu and Zhang (2007) and Lin (2013) both show that modeling the variance term structure is important for VIX derivatives pricing. In our robustness check with the VIX futures term structure data, we demonstrate the power of using the adaptive methods to predict VIX futures directly.

Finally, our study can be included in the multi-factor stochastic volatility models literature. Christoffersen et al. (2008) and Christoffersen et al. (2009) show that a two-factor volatility model which breaks volatility into long-run and short-run components assists option price modeling. Our empirical results are consistent with these studies as we demonstrate that that the information from the entire term structure is useful in predicitng future implied volatilities.

The rest of the paper is organized as follows. Section 2 describes the data, providing strong evidence for persistence and against TSIV stationarity. Section 3 describes the ADNS for modeling and forecasting with the LAR model as a dynamic element. Section 4 reports the forecast results and Section 5 concludes.

2 Data

We use two data sets to verify the predictive accuracy of the adaptive models against alternative popular models. The first data set is the daily implied volatility data of S&P 500 index call options from 1996 to 2011, a total of 16 years, which is the same sample as explored in Guo et al. (2018). The data set is obtained from the Ivy DB OptionMetrics database. For the detailed forecast comparison with alternative models, we select the term structure data of relatively liquid call options with a delta of 0.5. The TSIV consists of ten different times to maturity (30, 60, 91, 122, 152, 182, 273, 365, 547 and 730 days) on each trading day. Any missing data of non-traded times to maturity on a specific trading day are interpolated and provided by OptionMetrics. For robustness check, we re-do all exercises on the TSIV data with two different deltas, 0.4 and 0.6, respectively. The second data set for robustness check is the daily VIX futures term structure from December 31, 2010 to September 1, 2017, which contains market information on the implied volatility across strike prices given a maturity. We directly estimate the term structure with NS interpolation and make prediction based on the adapative models.

Figure 1 plots five selected maturities from our data set. Along the time series dimension, the figure shows that the series behave differently in different periods, being relatively more tranquil in 2004-2007 than in 1996-2003, then becoming highly volatile in 2008-2011. Modeling time series with long spans across different regimes often incurs nonstationarity problems. Thus, we divide the 16-year sample into four subsamples, each of four years, to reflect possible changes in the regimes mentioned above. We select the 1-month (30 days) and 1-year (365 days) maturities to plot the autocorrelation function (ACF) for each subsample in Figure 2 and the partial ACF (PACF) for each subsample in Figure 3. These figures reveal that, even within the subsamples with mitigated changing regime problems, the ACFs still present slowly decaying patterns which are highly significant up to lag 50, and the PACFs are significant after the first two lags, indicating strong persistence. These features justify our application of the adaptive approach described in Section 3 to choose subsamples in a data-driven way to address the persistence problem.

[Figure 1. Time Plot of Selected Maturities of Implied Volatility]

[Figure 2. Sample Autocorrelations of 1-month and 1-year Implied Volatility]

[Figure 3. Sample Partial Autocorrelations of 1-month and 1-year Implied Volatility]

Figure 1 also reveals strong comovement across maturities. Comparing the five time series and their statistics, there is an upward sloping curve, on average, with a higher variation at the short end. However, the shape can be reversed to a downward slope during turmoil or in the crisis periods of 1998-1999, 2002-2003 and 2008-2009. The changing shapes can be further visualized three-dimensionally in Figure 4, where we plot the evolution of the volatility curve in three representative months. The first month, January 1996, is a normal period where the curve is upward sloping and positively humped; the second month, March 2004, is a period of moderate volatility where the curve is sometimes upward, flat, or even downward sloping; the third month, September 2008, is a special period during which the Lehman Brothers' bankrupty triggered a huge spike of the whole volatility curve with a sharp downward shape and negative hump. These shapes of the implied volatility term structure are qualitatively similar to the term structure of interest rates, or often called the yield curves.

[Figure 4. The Term Structure of Implied Volatility in Selected Periods]

3 Adaptive dynamic Nelson-Siegel model for the TSIV

There is a rich literature on the cross-sectional modeling of the yield curve, among which the Nelson-Siegel interpolation (the NS model) proposed by Nelson and Siegel (1987) and then developed by Diebold and Li (2006) as the dynamic NS (DNS) model is the most popular of the three-factor reduced-form models, and it is widely used by investors, researchers and policy makers. The NS model has satisfactorily modeled the term structure of implied volatility, see Chalamandaris and Tsekrekos (2011) and Guo et al. (2014), among others. Although cross-section information is helpful for forecasting the TSIV, such as in VAR or VECM models, the results of Guo et al. (2018) do not favor the DNS when compared to alternative models.

But Guo et al. (2018) assume the stationarity of all models used in the estimation sample and use predetermined window lengths to estimate parameters for prediction. When we look at the data, we find that the time series behaves differently in time, for mean, reversability or volatility. In reality, changes in market conditions, business cycles and policies alike influence the pricing and trading of assets, sometimes temporarily and sometimes fundamentally. The implication is that the time-changing features of the data should be modeled for forecast precision.

The LAR model used in Chen et al. (2010) for realized volatility is convenient for modeling a single IV. However, it is computationally burdensome to predict the whole TSIV using this univariate model.

We will, therefore, model the TSIV jointly with the ADNS proposed in Chen and Niu (2014) which considers cross-sectional comovement in the maturity dimension. That is, we model the dynamics of the three NS factors as an LAR model. The LAR model allows time-varying parameters globally as well as local homogeneity in subsamples where parameters can be approximately constant. The ADNS model has performed remarkably well in yield curve forecasting. Chen and Niu (2014) show that the ADNS can outperform the random walk, as well as many alternative yield curve models with predetermined window selection. The superior predictability increases along the forecast horizon, up to 12 months ahead at the monthly frequency. This paper extends the scope of the ADNS application to the TSIV at the daily frequency.

Both the LAR specification of a single IV and the ADNS specification for the TSIV will be assessed against more than a dozen alternative models in our empirical study. The following illustration of the adaptive approach will focus on the ADNS which nests the LAR as its state dynamics.

3.1 Cross-sectional modeling with Nelson-Siegel factors

In the framework of Nelson and Siegel (1987) and Diebold and Li (2006), the yield curve across maturities can be formulated by three factors, β_{1t}, β_{2t} and β_{3t}, named level, slope and curvature, respectively. This gives

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) + \epsilon_t(\tau), \quad \epsilon_t(\tau) \sim N(0, \sigma_\epsilon^2) \qquad (1)$$

where $y_t(\tau)$ denotes the yield curve with maturity τ (in days) at time t. Determining the exponentially decaying rate of the loadings for the slope and curvature factors, the shape parameter λ needs to be calibrated with a reasonable value for the goodness of fit in the cross section. Minimizing the squared errors of Equation (1) for the whole sample, we find that λ falls into a range where the peak of the curvature factor is around 2 to 4 months (60 to 120 days). Calibrating the value with the sample up to the end of 2001, the point before the out-of-sample forecast exercise, we find that $\lambda = 0.0196$ is optimal, which ensures peak curvature at the three-month maturity. The factor loadings with $\lambda = 0.0196$ are depicted in Figure 5. We continue with this λ value for the ADNS model forecast.

[Figure 5. Nelson-Siegel Factor Loadings]

Conditional on $\lambda = 0.0196$, the loadings of the factors in Equation (1) are fixed for each maturity. We apply OLS to extract the three factors at each time period. Figure 6 displays the dynamics of the extracted three factors.

[Figure 6. Time Evolution of Nelson-Siegel Factors Extracted from Implied Volatility]

The level factor β_{1t} traces the long maturity volatility y(730) well, as displayed in the last panel of Figure 1. The slope factor β_{2t} is negative on average, implying an upward sloping curve, but surges positively to produce a negative slope during periods of market turmoil or crisis. The curvature factor fluctuates around zero, with a positive value for a positive hump shape and a negative value for an inverted hump.

With the factors already extracted in the first place, we model each factor, β_{it}, $i = 1, 2$ and 3, with an LAR process. The factor forecast is then used to construct the forecast of the TSIV according to Equation (1).

3.2 Local autoregressive model for factor dynamics

To simplify this illustration, in what follows, we omit the subscript i when no specific factor is referred to. In modeling the TSIV at the daily frequency, we do not consider exogenous factors related to the business cycle at a lower frequency. But the modeling framework is sufficiently flexible to accomodate exogenous factors for future improvement. The properties and robustness of the LAR model are investigated and demonstrated through Monte Carlo simulations in Chen and Niu (2014).

3.2.1 The LAR model and estimator

When we model the factor at a particular time t, the LAR(1) is defined through a parameter set θ_t:

$$\beta_t = \theta_{0t} + \theta_{1t}\beta_{t-1} + \mu_t, \quad \mu_t \sim N(0, \sigma_t^2) \tag{2}$$

where $\theta_t = (\theta_{0t}, \theta_{1t}, \sigma_t)^\top$. The parameter set is indexed to time to reflect the assumption that it can be time varying. However, the key to LAR model estimation is to identify an interval of length m_t, $[t - m_t, t - 1]$, over which the process can be described by an autoregressive model and where the parameter θ_t stays approximately constant, or is locally homogenous. For a specified interval, the (quasi) maximum likelihood estimation can be used to estimate the parameters. The local maximum likelihood estimator $\tilde{\theta}_t$ is defined as:

$$\begin{aligned}\tilde{\theta}_t &= \arg\max_{\theta_t\in\Theta} L(\boldsymbol{\beta}; I_t, \theta_t) \\ &= \arg\max_{\theta_t\in\Theta}\left\{-m_t\log\sigma_t - \frac{1}{2\sigma_t^2}\sum_{s=t-m_t+1}^{t}(\beta_s - \theta_{0t} - \theta_{1t}\beta_{s-1})^2\right\}\end{aligned}$$

where Θ is the parameter space and $L(\boldsymbol{\beta}; I_t, \theta_t)$ is the local log-likelihood function.

3.2.2 The testing procedure for homogeneous intervals

At time t, the goal is to conduct a backward selection for the longest possible interval of local homogeneity among a finite set of candidates. We divide the sample into discrete increments of M periods ($M > 1$) between any two adjacent subsamples to obtain K_t candidate subsamples for any particular time point t, or

$$I_t^{(1)}, \cdots, I_t^{(K)} \text{ with } I_t^{(1)} \subset \cdots \subset I_t^{(K)},$$

where the shortest subsample, $I_t^{(1)}$, should be reasonably fit by an AR(1) model with constant parameters and local homogeneity is assumed to hold within $I_t^{(1)}$ by default.

Starting from the first subsample, $I_t^{(1)}$, the parameters are estimated with the maximum likelihood (ML) estimator $\tilde{\theta}_t^{(k)}$ for a specific interval $I_t^{(k)}$. This estimate may or may not be accepted as the homogenous adaptive estimator, which is denoted as $\hat{\theta}_t^{(k)}$, except for the first subsample where $\hat{\theta}_s^{(1)} = \tilde{\theta}_s^{(1)}$ by default. Acceptance or rejection, starting from $I_t^{(2)}$ onwards, is controllwhied by a likelihood ratio test defined in each following subsample $I_t^{(k)}$ as

$$T_t^{(k)} = \left|L(I_t^{(k)}, \tilde{\theta}_t^{(k)}) - L(I_t^{(k)}, \hat{\theta}_t^{(k-1)})\right|^{1/2}, \quad k = 2, \cdots, K \tag{3}$$

where $L(I_t^{(k)}, \tilde{\theta}_t^{(k)}) = \max_{\theta_t\in\Theta} L(\boldsymbol{\beta}; I_t^{(k)}, \theta_t)$ denotes the fitted likelihood under *hypothetical* homogeneity and $L(I_t^{(k)}, \hat{\theta}_t^{(k-1)}) = L(\boldsymbol{\beta}; I_t^{(k)}, \hat{\theta}_t^{(k-1)})$ refers to the likelihood in the current testing subsample using the parameter estimate from the previously *accepted* local homogeneous interval $I_t^{(k-1)}$.

The test statistic measures the difference between these two estimates, which is then to be compared with a set of critical values $\zeta_1, \cdots, \zeta_K$. If $T_s^{(k)} \leq \zeta_k$, then the difference is regarded as acceptable due to sampling randomness, and we accept the current subsample $I_t^{(k)}$ as being homogeneous and update the adaptive estimator $\hat{\theta}_t^{(k)} = \tilde{\theta}_t^{(k)}$. If $T_t^{(k)} > \zeta_k$, then this indicates that the model significantly changes, and the procedure terminates with the latest accepted subsample $I_t^{(k-1)}$ selected, such that $\hat{\theta}_t^{(k)} = \hat{\theta}_t^{(k-1)} = \tilde{\theta}_t^{(k-1)}$. For $\ell \geq k$, we denote $\hat{\theta}_s^{(\ell)} = \tilde{\theta}_s^{(k-1)}$, meaning that the adaptive estimator for an even longer subsample at time t is the ML estimate over the longest subsample of local homogeneity identified. The procedure is continued until the parameter change cannot be rejected or the longest subsample, $I_t^{(K)}$, is reached under local homogeneity.

3.2.3 Critical value calibration

The required set of critical values are calibrated empirically with Monte Carlo experiments, where an underlying AR(1) model, such as Equation (4) with reasonable constant parameter values θ^*, is used to simulate training samples with a length of $I^{(K)}$. This gives

$$\beta_t = \theta_0^* + \theta_1^* \beta_{t-1} + \mu_t, \quad \mu_t \sim N(0, \sigma^{*2}). \tag{4}$$

With constant parameters, the training sample is a globally homogeneous AR(1) time series with $\theta^* = (\theta_0^*, \theta_1^*, \sigma^*)$ for $t = 1, \cdots, T$. Since time homogeneity is fulfilled globally, the ML estimate $\tilde{\theta}_t^{(k)}$ in each subsample $I_t^{(k)}$, $k = 1, \cdots, K$, is optimal. The estimation error can be measured by the fitted log-likelihood ratio:

$$R_k = E_{\theta^*} \left| L\left(I_t^{(k)}, \tilde{\theta}_t^{(k)}\right) - L\left(I_t^{(k)}, \theta^*\right) \right|^{1/2}, \tag{5}$$

where R_k is computed numerically with a known value of θ^*.

In the backward testing procedure, the goal is to achieve an adaptive estimator $\hat{\theta}_t^{(k)}$ close to θ^*. Under a specific $\hat{\theta}_t^{(k)}$, a temporal difference, denoted as $D_t^{(k)}$, will arise between the ML estimator $\tilde{\theta}_t^{(k)}$ and the adaptive estimator $\hat{\theta}_t^{(k)}$:

$$D_t^{(k)} = \left| L\left(I_t^{(k)}, \tilde{\theta}_t^{(k)}\right) - L\left(I_t^{(k)}, \hat{\theta}_t^{(k)}\right) \right|^{1/2}.$$

Ideally, for $\hat{\theta}_t^{(k)}$ close to θ^*, the stochastic distance $D_t^{(k)}$ is bounded by the ideal estimation error R_k in Equation (5):

$$E_{\theta^*}\left(D_t^{(k)}\right) = E_{\theta^*} \left| L\left(I_t^{(k)}, \tilde{\theta}_t^{(k)}\right) - L\left(I_t^{(k)}, \hat{\theta}_t^{(k)}\right) \right|^{1/2} \leq R_k. \tag{6}$$

With this inequality as the risk bound, the critical values determining $E_{\theta^*}\left(D_t^{(k)}\right)$ can be computed numerically. To see this, note that for a given ζ_k, with which comparing the test statistic $T_s^{(k)}$, two outcomes can arise:

- If $T_t^{(k)} \leq \zeta_k$, such that we accept $\hat{\theta}_t^{(k)} = \tilde{\theta}_t^{(k)}$, then we have $D_t^{(k)} = 0$; or
- If $T_t^{(k)} > \zeta_k$, such that we set $\hat{\theta}_s^{(k)} = \hat{\theta}_s^{(k-1)} = \tilde{\theta}_s^{(k-1)}$, then we have $D_t^{(k)} = \left| L\left(I_t^{(k)}, \tilde{\theta}_t^{(k)}\right) - L\left(I_t^{(k)}, \hat{\theta}_t^{(k-1)}\right) \right|$

With simulated training samples, $E_{\theta^*}\left(D_t^{(k)}\right)$ can be computed given any specified ζ_k. Thus, ζ_k can be calibrated to the value when the risk bound, Equation (6), is reached.

Although the calibration of critical values relies on a set of hyperparameters (θ^*, K, M), Chen and Niu (2014) demonstrate that the calibration is quite robust against reasonable deviations to the true value θ^*, and to a wide range of K and M combinations. To make our calibration reasonable, we use the NS factors from 1996 to 2001 to calculate the ML estimate as θ^* for our training samples. The assumption is that the time-varying parameters fluctuate around this θ^* which can be regarded as a sample averaged over 1996 to 2001. We use θ^* to simulate N series of data, each of length $T = 180$, and calibrate the set of critical values as described above in this section. The same set of calibrated critical values are adopted for every time point throughout the real-time estimation and forecast from 2002 to 2011. In our procedure, at each point in time, we consider $K = 30$ subsamples for the test procedure, with the increment of $M = 6$ trading days between any adjacent subsamples, i.e., 180 trading days (about 8 or 9 calendar months) is the maximal subsample size. As it happens, $T = 180$ is a sufficiently large training sample length as the selected maximal intervals in our empirical study never reach it.

3.2.4 Estimation

Following the test and estimation procedure described in Section 3.2.3, we estimate the LAR(1) model for each trading day from 2001-12-31 onwards for each factor of the ADNS model. To easily visualize, the results and the time-varying features of the parameters, we plot the parameter estimates for the ADNS factors on the last trading day of each month in Figure 7. In each column, we plot the evolution of the three parameters of the LAR process of the NS factor: the intercept, θ_{0t}; the autoregressive coefficient, θ_{1t}; and the standard deviation of the error term, σ_t. The parameters fluctuate wildly. Changes in the standard deviation are in line with the volatile behavior of the implied volatility for different periods. Figure 7 shows the varying properties of the parameters along time, which supports the adaptive approach application.

[Figure 7. Parameter Evolution for Factor Dynamics in the ADNS Model]

3.3 The ADNS state-space framework for the TSIV

To summarize, the state-space representation of the ADNS model is the following:

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) + \epsilon_t(\tau),\ \epsilon_t(\tau) \sim (0, \sigma^2_{\epsilon,\tau}) \quad (7)$$

$$\beta_{it} = \theta_{0t}^{(i)} + \theta_{1t}^{(i)}\beta_{it-1} + \mu_t^{(i)},\ \mu_t^{(i)} \sim (0, \sigma^2_{it}), \quad i = 1, 2, 3. \quad (8)$$

Now, based on the extracted NS factors and locally estimated state dynamics, the TSIV forecast h steps ahead can be directly obtained as a combination of h-steps ahead forecasts of the Nelson-Siegel factors:

$$\hat{y}_{t+h/t}(\tau) = \hat{\beta}_{1,t+h/t} + \hat{\beta}_{2,t+h/t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + \hat{\beta}_{3,t+h/t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) \tag{9}$$

where the factor forecasts are obtained by

$$\hat{\beta}_{i,t+h/t} = \hat{\theta}_{0t} + \hat{\theta}_{1t}\hat{\beta}_{i,t}, \;\; i = 1, 2, 3. \tag{10}$$

Coefficient $\hat{\theta}_{jt}$ ($j = 0, 1, 2$) is obtained by regressing $\hat{\beta}_{i,t}$ on an intercept and $\hat{\beta}_{i,t-h}$. We estimate the LAR model for each specific forecast horizon.

4 LAR and ADNS forecasts and comparison with alternative models

In this section, we show the results of the LAR model for a single IV and the ADNS model forecast for the whole TSIV, in comparison with the random walk model and thirteen other models for multiple forecast horizons. The forecast comparison is made over a long time span of 10 years.

4.1 Forecast procedure

Using the TSIV data described in Section 2 from 1996 to 2011, we make a psuedo real-time out-of-sample forecast. The forecast period for model comparison is from 2002-1-2 to 2011-12-30, a total of 2519 trading days. We make 1-, 5-, and 20-day ahead forecasts. Here the forecast horizon is in units of trading days, such that the 5-day ahead forecast is one calendar week ahead, and the 20-day ahead forecast is about a calendar month ahead, etc., taking the holidays in between into consideration. So for forecasting the first period of 2002-1-2, the 1-day ahead forecast is based on a sample up to 2001-12-31, the 5-day ahead forecast is based on a sample up to 2001-12-24, and the 20-day ahead forecast is based on a sample up to 2001-12-3. We make this estimation and forecast exercise for each forecast horizon day after day until we reach the end of the sample.

At each estimation period t, we use data up to t. The three NS factors are first extracted using OLS for the available sample. Then, for each factor of the ADNS, we estimate a LAR(1) process with the identified longest homogenous interval for a specific forecast horizon. Based on the local adaptive estimator, we make 1-, 5- and 20-day ahead out-of-sample forecasts for the NS factors as described by Equation (10). The predicted TSIV of various maturities are then constructed by the forecast of the NS factors at the specific horizon according to Equation (9).

It is straight forward to forecast the single IV as a univariate time series with LAR.

4.2 Alternative models for comparison

For comparison, we choose forecasts from the random walk model and thirteen other alternative models as summarized in Table 1. Forecast of these alternative models have been examined in Guo et al. (2018) with recursive estimation, i.e., the estimation sample extends as the forecast exercise moves forward. Among these thirteen forecast results, there are ten single model forecasts and the other three are combination forecasts based on the single model forecast. For different forecast horizons h, the direct forecast is used with a regression on lag h information. Guo et al. (2018) also examine two combination forecasts using discounted mean squared prediction errors. We do not consider these combinations because their forecast period starts later than that of the base models to construct weights from the initial forecasts generated by the base models. Also, in their paper, these two combination forecasts do not outperform the other thirteen forecasts.

We now list the econometric representations of these alternative forecasts. First, the random walk model serves as a natural benchmark for all, with $\hat{y}_{t+h}(\tau) = y_t(\tau)$. Then, the thirteen alternative models are as follows.

(M1) Nelson-Siegel factors as univariate AR(1) processes, which is a simple DNS model.

(M2) Nelson-Siegel factors as multivariate VAR(1) processes.

These two specifications both belong to the traditional DNS models where the estimation is based on a predetermined regression window length.

(M3) Slope regression: $\hat{y}_{t+h}(\tau) - y_t(\tau) = c_0(\tau) + c_1(\tau)\left(y_t(730) - y_t(30)\right)$.

(M4) AR(1) on volatility levels: $\hat{y}_{t+h}(\tau) = c_0(\tau) + c_1(\tau) y_t(\tau)$.

(M5) VAR(1) on volatility levels: $\hat{y}_{t+h} = c_0 + c_1 y_t$, where the vector of $y_t = \begin{bmatrix} y_t(30) \\ y_t(91) \\ y_t(152) \\ y_t(365) \\ y_t(730) \end{bmatrix}$.

(M6) VAR(1) on volatility changes: $\hat{z}_{t+h} = c_0 + c_1 z_t$, where $\hat{z}_t = \begin{bmatrix} y_t(30) - y_{t-h}(30) \\ y_t(91) - y_{t-h}(91) \\ y_t(152) - y_{t-h}(152) \\ y_t(365) - y_{t-h}(365) \\ y_t(730) - y_{t-h}(730) \end{bmatrix}$.

(M7) ECM(1) with one common trend: $\hat{z}_{t+h} = c_0 + c_1 z_t$, where $\hat{z}_t = \begin{bmatrix} y_t(30) - y_{t-h}(30) \\ y_t(91) - y_t(30) \\ y_t(152) - y_t(30) \\ y_t(365) - y_t(30) \\ y_t(730) - y_t(30) \end{bmatrix}$.

(M8) ECM(1) with two common trends: $\hat{z}_{t+h} = c_0 + c_1 z_t$, where $\hat{z}_t = \begin{bmatrix} y_t(30) - y_{t-h}(30) \\ y_t(91) - y_{t-h}(91) \\ y_t(152) - y_t(30) \\ y_t(365) - y_t(30) \\ y_t(730) - y_t(30) \end{bmatrix}$.

(M9) AR(1) regression on the first three principal components of the TSIV. Each principal component is modeled as an AR(1) process, and a volatility forecast is generated by the principal components with loadings estimated from the principal component analysis for each period.

(M10) VAR(1) on empirical level, slope and curvature: $\hat{y}_{t+h} = c_0 + c_1 F_t$, where the empirical factors are constructed by the volatility of the representative maturities; $F_t = \begin{bmatrix} y_t(365) \\ y_t(365) - y_t(30) \\ 2y_t(122) - [y_t(365) + y_t(30)] \end{bmatrix}$.

Then three combination forecasts are considered based on the model forecast from model (1) to (10); $\hat{y}_{c,t+h}(\tau) = \sum_{k=1}^{10} w_{k,t}(\tau)\, \hat{y}_{k,t+h}(\tau)$.

(M11) The mean combination forecast: $w_{k,t}(\tau) = 1/10$.

(M12) The median combination forecast: the median of $\hat{y}_{k,t+h}(\tau)$, k = 1,...,10.

(M13) The trimmed mean combination forecast: $w_{k,t}(\tau) = 0$ for the smallest and largest forecasts and $w_{k,t}(\tau) = 1/8$ for the remaining forecasts.

4.3 Measures of forecast comparison

The measures for forecast performance include the forecast RMSE, the MAE, out-of-sample R^2, and the percentage of correct predictions of future directional changes. The first three evaluate forecast variance and fit around the realized value, and the fourth measures the ability to predict the direction of future changes. Moreover, for statistical significance, we use the DM test to evaluate outperformance of the adaptive models against alternative models M1 to M13, and use the CW test to compare the adaptive models with random walk under nested model framework.

4.4 Forecast results

4.4.1 Comparison by forecast RMSE, MAE and out-of-sample R^2

Table 2 reports the forecast RMSE and MAE of the benchmark random walk model, the LAR for a single IV and the ADNS for a TSIV. The left-hand column shows the RMSE results and the right-hand column shows the results for MAE, with each panel showing the results of the model for each maturity and forecast horizons (denoted by h) of 1, 5 and 20 trading days, respectively. Table 3 reports the relative performance of the LAR and the ADNS compared to the random walk, based on the results in Table 2.

The ratios in Table 3 are the most intuitive comparison of the relative strengh of the LAR and the ADNS versus the random walk. If a ratio is smaller than 1, then the number is in bold, indicating a better forecast by either the LAR or the ADNS. If the number is between 0.8 and 0.9, then the bold number is underlined; if the number is smaller than 0.8, then a grey shadow is added to visually emphasize the outperformance. The ratios unequivocally show that the LAR and the ADNS outperform random walk for both measures. The advantage increases as the forecast horizon increases, from a reduction of RMSE or MAE of a few percent for the 1-day ahead forecast, to a reduction of between 10% and 20% reduction for the 5-day ahead forecast, and a huge reduction of between 30% and 50% for the 20-day ahead forecast. Compared with the single LAR of each IV series, the ADNS model not only predicts all cross-sections with only three factor processes, but also improves the forecast performance of the RMSE measure.

[Table 2. Forecast RMSE and MAE of LAR, ADNS and Random Walk]
[Table 3. Ratios of Forecast RMSE and MAE of LAR and ADNS vs. Random Walk]

It may be argued that the random walk is not a good model for the volatility curve econometrically or economically, or that the ADNS superiority is not clearly shown. Hence, using the random walk model's performance as a benchmark against which to compare the thirteen sets of alternative forecasts will give a better assessment. Table 4 reports the ratio of the RMSE and the MAE of the thirteen alternative forecasts relative to the random walk model for the 1-, 5-, and 20-day ahead forecast horizons. Some multivariate factor models exhibit better prediction abilities of between 1% and 4% for the 1- and 5-day ahead forecast, such as M6, M7, and M8. The forecast combinations also perform relatively well for the 1- and 5-day ahead forecasts with a reduction of 1-3 % reduction compared to the randow walk forecast. But clearly none of these alternative models perform better than our adaptive models.

[Table 4. Forecast Comparison of Alternative Models with Random Walk]

Out-of-sample R^2 is another measure for predictibility. We compute the ratio of out-of-sample R^2 of M1-M13 and the adaptive models against random walk and the results are reported in Table 5. A number bigger than 1 indicates a better prediction of the compared model than random walk. The results show that the random walk is unbeatable by M1-M13 across forecast horizons and maturities. But the ADNS model can beat random walk for forecast horizons of $h = 20$ days across all maturities, and the LAR model performs even better for both $h = 5$ and $h = 20$ across all maturities.[1]

[Table 5. Out-of-sample R^2 comparison versus random walk]

4.4.2 Comparison of future sign change predictions

For successful trading strategies, it is often crucial to forecast future directional changes, i.e., whether volatility will go up or down. Since the random walk model assumes no change, it is not useful here. It is interesting to assess how well the adaptive models perform with respect to the alternative models in the directional prediction.

For each model, based on the model forecast of $\hat{y}_{t+h}(\tau)$, we can compute $\Delta\hat{y}_{t+h}(\tau)$. Then we compare the signs of the predicted change $\Delta\hat{y}_{t+h}(\tau)$ and the realized change $\Delta y_{t+h}(\tau)$. We document the correct sign prediction frequency for each model/maturity/horizon combination in Table 6.

[Table 6. Frequency of Correct Sign Prediction for Future Changes]

Table 6 shows that the thirteen alternative models have only a moderate performance with around 50% of correct predictions, and the best performance is 59% correct predictions for Model 7 predicting the 5-day ahead 30-day IV. In stark contrast, the adaptive models show a commanding advantage, with the frequency of correct prediction ranging from an average 63-64% across the TSIV for the 1-day ahead forecast, 73-74% for the 5-day ahead forecast, and 86-86% for the 20-day ahead forecast.

[1] Guo et al. (2018) show that even with the marginal forecast advantages with respect to the random walk, the alternative models are able to produce economically viable profits after controlling for reasonable transaction costs. The forecast advantages of M6-M8 of a reduction in the errors of a few percent, with respect to the random walk model, are actually trivial compared to the superior forecast of the ADNS.

4.4.3 Statistical test of forecast performance

The above results show performance of the adaptive models with various measures based on a large forecast sample, i.e., 2002-01-21 to 2011-12-30, a total of 2519 trading days. Most measures confirm their sizable advantage against all alternative models for horizons $h = 5$ and $h = 20$. To demonstrate further their statistical significance, we use the DM test to evaluate outperformance of the adaptive models against alternative models M1 to M13, and use the CW test to compare the adaptive models with random walk under nested model framework.

The DM test is proposed by Diebold and Li (1995) for forecast comparison between different models or model-free methods. Under the hypothesis of equal predictive accuracy, the loss differential between two forecasts follows a limiting normal distribution. The CW test is developed by Clark and West (2006, 2007) for forecasts between nested models. They point out that the DM test does not have a standard normal distribution when applied to forecast from nested models. The CW test addresses the distribution issue for nested models. The LAR and ADNS models can be nested with the random walk under two restrictions: 1) the intercept of the LAR is zero or the combined intercept of LARs of the three ADNS factors is zeros, and 2) the autoregressive coefficients of the LAR or the LARs underlying the ADNS are 1s.

Based on the test applicability, we compute the DM test statistics for LAR/ADNS against M1 to M13 and the CW test statistics for LAR/ADNS against the random walk. The results are shown for LAR in Table 7-a and for ADNS in Table 7-b. A negative (positive) value indicates better (worse) accuracy of LAR/ADNS against an alternative model in terms of quadratic loss measured by mean-squared predicted errors (MSPE). Both tests show that the adaptive forecasts outperform alternative models significantly for $h = 5$ and $h = 20$ across all maturities.

[Table 7. Test of forecasting performance against alternative models]

4.4.4 Stable subsample lengths

The reason that the ADNS model performs markably better is that the data generating process is time changing, as shown in Figure 7 with estimated parameter evolution. The conventional practice to include more observations may contribute to more robust estimates, with the trade-off of higher parameter uncertainty with frequently or wildly changing parameters. To what extent and for how long a local model can stay homogenous is crucial to determining the trade-off. Table 8 summarizes the average lengths detected for homogenous intervals for each factor and each forecast horizon. It shows that the first and third

NS factors are less stable, with average homogenous intervals of less than 30 trading days, while the second NS factor changes less frequently but only up to an average of 39 trading days. Another feature is that, as the forecast horizon lengthens, the lengths of the stable intervals tend to decline. The implication is that the adaptive procedure can be explored up to a limit, conditional on the frequency of the data.

[Table 8. Average Lengths of Homogenous Intervals Detected Adaptively]

4.5 Robustness check with alternative data

In order to understand whether the adaptive models are robust for different strikes and whether it works well with model-free implied variance, such as VIX futures, we make robustness checks with alternative data.

First, from the same data source of Ivy DB OptionMetrics, we select the TSIV from two different deltas, 0.4 and 0.6, respectively. The maturity structure and time span is the same as the TSIV at delta of 0.5. From the previous exercises, we already know that no alternative models among M1-M13 can dominate the random work. We re-do the forecast comparison with the random walk for the ease of comparison and report the CW test statistics in Table 9. The results echo those when delta is 0.5 that the adaptive model significantly outperform the random walk for forecast at $h = 5$ and $h = 20$ across all maturities.

Second, we forecast the VIX futures term structure with the adaptive models and compare the forecast accuracy with random walk prediction. In order to forecast the whole term structure, we interpolate the VIX futures term structure of discrete maturies of 1 to 10 month contracts with the NS models at each trading day. It should be noted that for a n-month contract in the database, its remaining maturity may be between $n - 1$ and n months until another new contract of n-month replaces it as the market benchmark for n-month VIX futures. Therefore, on each trading day, we compute the precise remaining maturity in the unit of day for each contract. Then we interpolate the whole term structure up to 10 months with the static NS model for each trading day. We use the sample from December 31, 2010 to April 30, 2012 as the training sample, then make adaptive estimation and prediction starting from May 2, 2012 to September 1, 2017. For the random walk model, we assume that each NS factors follows a random walk and forecast the future NS factors and term structure accordingly. For each forecast horizon, we compare the NS implied futures price with the VIX futures actually exist in the database at their precise maturity, and compute the CW test statistics for adaptive models against the random walk.

The lower panel of Table 9 shows the test statistics with VIX futures term structure. The results are more encouraging. The adaptive models not only significantly outperform the random walk at $h = 5$ and $h = 20$ across all maturities, but they also forecast significantly better than the random walk at h = 1 for medium to long term contracts. And the ADNS performs better than the LAR.

5 Conclusion

When modeling the TSIV, we show that the adaptive procedure, combined with local parsimoneous dynamic models, can generate markedly better prediction performance in comparison to the random walk and thirteen alternative forecast models. The forecast advantage of this approach originates from an ability to adaptively choose suitable homogeneous intervals, maximally including useful information while reducing parameter instability.

The ADNS model based on three NS factors across maturity, with each factor modeled as an LAR process, is more efficient than modeling each single IV with an LAR model. The ADNS model consistently outperforms the random walk model over the whole volatility curve for multiple forecast horizons up to 20 trading days ahead. The forecast errors of the random walk model, measured by RMSE and MAE, can be reduced by between 20% and 60% for the 5- to 20-day ahead forecasts by using the ADNS model. The outperformance is reasured with robustness check using TSIV at alternative deltas and the VIX futures term structure.

The successful application of the ADNS model to forecasting beyond the yield curve context shows the power of the LAR model when in combination with a parsimonious cross-sectional NS interpolation. The results not only demonstrate to investors, researchers and policy makers the importance of modeling the time-changing aspects of financial time series for forecast purposes, but also give an effective example of how to achieve this.

Acknowledgements

Ying Chen acknowledges the support of the Singapore Ministry of Education Academic Research Fund Tier 1 grant (FRC R-155-000-178-114), Institute of Data Science grant (R-155-000-185-646) at the National University of Singapore, and Humboldt Universitaet zu Berlin-National University of Singapore Joint Grant (R-155-000-199-114) and National Natural Science Foundation of China (NSFC) Grant (No. 71528008). Linlin Niu acknowledges the support of the NSFC Grants (Nos. 71528008 and 71273007) and the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk". Qian Han acknowledges the support of NSFC Grant (Nos. 71471153 and 7163000017) and the Fundamental Research Funds for the Central Universities (No. 20720181004).

Reference

1. Byoun, S., Kwok, C. C. and Park, H. Y. (2003). Expectations hypothesis of the term structure of implied volatility: Evidence from foreign currency and stock index options. Journal of Financial Econometrics, 1(1), 126-151.

2. Bollen, N. P. and Whaley, R. E. (2004). Does net buying pressure affect the shape of implied volatility functions?. The Journal of Finance, 59(2), 711-753.

3. Breeden, D. T. and Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. Journal of Business, 621-651.

4. Campa, J. and Chang, P. K. (1995). Testing the expectations hypothesis on the term structure of volatilities in foreign exchange options. The Journal of Finance, 50(2), 529-547.

5. Canina, L. and Figlewski, S. (1993). The informational content of implied volatility. Review of Financial Studies, 6(3), 659-681.

6. Chalamandaris, G. and Tsekrekos A. E. (2011). How important is the term structure in implied volatility surface modeling? Evidence from foreign exchange options. Journal of International Money and Finance, 30, 623-640.

7. Chen, R. B., Chen, Y. and Haerdle, W. (2014). TVICA - Time varying independent component analysis and its application to financial data. Computational Statistics & Data Analysis, 74, 95-109.

8. Chen, Y., Haerdle, W. and Pigorsch, U. (2010). Localized realized volatility modelling. Journal of American Statistical Association, 105, 1376-1393.

9. Chen, Y. and Niu, L. (2014). Adaptive dynamic Nelson–Siegel term structure model with applications. Journal of Econometrics, 180(1), 98-115.

10. Christoffersen, P., Heston, S., and Jacobs, K. (2009). The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. Management Science, 55(12), 1914-1932.

11. Christoffersen, P., Jacobs, K., Ornthanalai, C., and Wang, Y. (2008). Option valuation with long-run and short-run volatility components. Journal of Financial Economics, 90(3), 272-297.

12. Clark, T.E., West, K.D. (2006). Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. Journal of Econometrics, 138, 155-186.

13. Clark, T.E., West, K.D. (2007). Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics 138, 291-311.

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

15. Diebold, F. X. and Li, C. (2006). Forecasting the term structure of government bond yields. Journal of Econometrics, 130(2), 337-364.

16. Diebold, F., Mariano, R. (1995). Comparing predictive accuracy. Journal of Business and Economic Statistics, 13, 253-263.

17. Diz, F. and Finucane, T. J. (1993). Do the options markets really overreact? Journal of Futures Markets, 13(3), 299-312.

18. Giacomini, E., Haerdle, W. and Spokoiny, V. (2009), Inhomogeneous dependency modelling with time varying copulae. Journal of Business & Economic Statistics, 27, 224–234.

19. Goncalves, S. and Guidolin, M. (2006). Predictable dynamics in the S&P 500 index options implied volatility surface. The Journal of Business, 79(3), 1591-1635.

20. Granger, C. W. J. and Hyung, N. (2004). Occasional structural breaks and long memory with an application to the S&P500 absolute stock returns. Journal of Empirical Finance, 11, 399-421.

21. Guo, B., Han, Q. and Lin, H. (2018). Are there gains from using information over the surface of implied volatilities? Journal of Futures Markets, 38, 645–672.

22. Guo, B., Han, Q. and Zhao, B. (2014). The Nelson-Siegel model of the term structure of option implied volatility and volatility components. Journal of Futures Markets, 34(8), 788-806.

23. Haerdle, W., Hautsch, N. and Mihoci, A. (2015). Local adaptive multiplicative error models for high-frequency forecasts. Journal of Applied Econometrics, 30(4), 529-550.

24. Heynen, R., Kemna, A. and Vorst, T. (1994). Analysis of the term structure of implied volatilities. Journal of Financial and Quantitative Analysis, 29(01), 31-56.

25. Jiang, G. J. and Tian, Y. S. (2005). The model-free implied volatility and its information content. Review of Financial Studies, 18(4), 1305-1342.

26. Lin, Y. N. (2013). VIX option pricing and CBOE VIX term structure: A new methodology for volatility derivatives valuation. Journal of Banking & Finance, 37(11), 4432-4446.

27. Luo, X. and Zhang, J. E. (2012). The term structure of VIX. Journal of Futures Markets, 32(12), 1092-1123.

28. Mercurio, D. and Spokoiny, V. (2004). Statistical inference for time inhomogeneous volatility models. Annals of Statistics, 32, 577-602.

29. Mixon, S. (2007). The implied volatility term structure of stock index options. Journal of Empirical Finance, 14(3), 333-354.

30. Nelson C. R. and Siegel, A. F. (1987). Parsimonious modeling of yield curves. Journal of Business, 60, 473-89.

31. Neumann, M. and Skiadopoulos, G. (2013). Predictable dynamics in higher-order risk-neutral moments: Evidence from the S&P 500 options. Journal of Financial and Quantitative Analysis, 48(03), 947-977.

32. Spokoiny, V., Wang, W. and Haerdle, W. (2013). Local quantile regression (with discussion). Journal of Statistical Planning and Inference, 143, 1109–1129.

33. Stein, J. (1989). Overreactions in the options market. The Journal of Finance, 44(4), 1011-1023.

34. Xu, X. and Taylor, S. J. (1994). The term structure of volatility implied by foreign exchange options. Journal of Financial and Quantitative Analysis, 29(01), 57-74.

35. Zhu, Y. and Zhang, J. E. (2007). Variance term structure and VIX futures pricing. International Journal of Theoretical and Applied Finance, 10(01), 111-127.

Table 1. Alternative Models for Comparison

We choose the random walk model and other thirteen models as alternative models for forecast comparison.

1) The random walk model is the benchmark model in Guo et al. (2018).
2) The thirteen models which have the same forecast period for comparison are listed below:

M1:	Nelson-Siegel factors as univariate AR(1) processes, with $\lambda = 0.0147$
M2:	Nelson-Siegel factors as multivariate VAR(1) processes, with $\lambda = 0.0147$
M3:	Slope regression
M4:	AR(1) on volatility levels
M5:	VAR(1) on volatility levels
M6:	VAR(1) on volatility changes
M7:	ECM(1) with one common trend
M8:	ECM(1) with two common trends
M9:	AR(1) regression on three principal components
M10:	VAR(1) on empirical level, slope and curvature
M11:	mean combination forecast
M12:	median combination forecast
M13:	trimmed mean combination forecast

Table 2. Forecast RMSE and MAE of LAR, ADNS and Random Walk

Model	y(n)\h	RMSE 1	RMSE 5	RMSE 20	MAE 1	MAE 5	MAE 20
RW	y(30)	0.0197	0.0317	0.0515	0.0125	0.0207	0.0327
	y(60)	0.0149	0.0246	0.0417	0.0095	0.0163	0.0265
	y(91)	0.0126	0.0211	0.0362	0.0081	0.0141	0.0235
	y(122)	0.0108	0.0184	0.0319	0.0070	0.0123	0.0209
	y(152)	0.0095	0.0164	0.0289	0.0062	0.0111	0.0192
	y(182)	0.0086	0.0150	0.0267	0.0056	0.0102	0.0179
	y(273)	0.0072	0.0126	0.0228	0.0047	0.0087	0.0157
	y(365)	0.0064	0.0115	0.0210	0.0042	0.0079	0.0145
	y(547)	0.0052	0.0096	0.0181	0.0034	0.0067	0.0129
	y(730)	0.0047	0.0088	0.0170	0.0031	0.0062	0.0123
LAR	y(30)	0.0181	0.0277	0.0398	0.0112	0.0166	0.0193
	y(60)	0.0139	0.0214	0.0323	0.0086	0.0128	0.0151
	y(91)	0.0118	0.0186	0.0276	0.0074	0.0113	0.0135
	y(122)	0.0100	0.0161	0.0235	0.0063	0.0098	0.0114
	y(152)	0.0088	0.0144	0.0214	0.0056	0.0089	0.0104
	y(182)	0.0080	0.0130	0.0187	0.0052	0.0080	0.0092
	y(273)	0.0067	0.0110	0.0153	0.0044	0.0068	0.0077
	y(365)	0.0060	0.0100	0.0136	0.0039	0.0062	0.0068
	y(547)	0.0048	0.0081	0.0116	0.0031	0.0051	0.0059
	y(730)	0.0043	0.0072	0.0098	0.0028	0.0046	0.0050
ADNS	y(30)	0.0181	0.0272	0.0371	0.0112	0.0162	0.0193
	y(60)	0.0140	0.0216	0.0292	0.0087	0.0130	0.0153
	y(91)	0.0119	0.0184	0.0248	0.0075	0.0112	0.0134
	y(122)	0.0103	0.0161	0.0214	0.0065	0.0099	0.0117
	y(152)	0.0090	0.0142	0.0189	0.0057	0.0088	0.0104
	y(182)	0.0082	0.0129	0.0172	0.0052	0.0081	0.0095
	y(273)	0.0068	0.0108	0.0140	0.0045	0.0069	0.0080
	y(365)	0.0062	0.0097	0.0126	0.0041	0.0063	0.0072
	y(547)	0.0049	0.0081	0.0107	0.0032	0.0052	0.0059
	y(730)	0.0046	0.0076	0.0101	0.0032	0.0050	0.0057

Notes: 1) *h* denotes the forecast horizon in trading day units. The forecast period is from 2002-01-21 to 2011-12-30, a total of 2519 trading days.
2) In comparison to the random walk model, forecast RMSE and MAE results are reported for the two adaptive models: the LAR model for each single IV, and the ADNS for the whole TSIV with three factors, with each factor based on an LAR process.

Table 3. Ratios of Forecast RMSE and MAE of LAR and ADNS vs. Random Walk

	$y_{(n)}$\h	RMSE ratio vs. R.W. 1	5	20	MAE ratio vs. R.W. 1	5	20
LAR	y(30)	**0.918**	**0.873**	**0.773**	**0.894**	**0.800**	**0.591**
	y(60)	**0.929**	**0.871**	**0.775**	**0.903**	**0.785**	**0.570**
	y(91)	**0.934**	**0.880**	**0.762**	**0.917**	**0.799**	**0.573**
	y(122)	**0.930**	**0.875**	**0.736**	**0.908**	**0.790**	**0.544**
	y(152)	**0.930**	**0.879**	**0.739**	**0.912**	**0.801**	**0.543**
	y(182)	**0.932**	**0.866**	**0.699**	**0.917**	**0.781**	**0.513**
	y(273)	**0.941**	**0.870**	**0.672**	**0.931**	**0.779**	**0.489**
	y(365)	**0.938**	**0.874**	**0.649**	**0.926**	**0.787**	**0.471**
	y(547)	**0.926**	**0.850**	**0.641**	**0.910**	**0.757**	**0.457**
	y(730)	**0.913**	**0.812**	**0.574**	**0.896**	**0.732**	**0.409**
	Average	**0.929**	**0.865**	**0.702**	**0.911**	**0.781**	**0.516**
ADNS	y(30)	**0.920**	**0.858**	**0.721**	**0.899**	**0.782**	**0.592**
	y(60)	**0.940**	**0.877**	**0.701**	**0.917**	**0.797**	**0.577**
	y(91)	**0.944**	**0.870**	**0.686**	**0.928**	**0.797**	**0.569**
	y(122)	**0.955**	**0.873**	**0.671**	**0.935**	**0.800**	**0.560**
	y(152)	**0.951**	**0.863**	**0.652**	**0.927**	**0.791**	**0.541**
	y(182)	**0.949**	**0.859**	**0.643**	**0.932**	**0.793**	**0.533**
	y(273)	**0.951**	**0.852**	**0.615**	**0.945**	**0.790**	**0.508**
	y(365)	**0.973**	**0.848**	**0.601**	**0.969**	**0.790**	**0.493**
	y(547)	**0.941**	**0.843**	**0.590**	**0.921**	**0.776**	**0.458**
	y(730)	**0.991**	**0.856**	**0.593**	1.013	**0.802**	**0.464**
	Average	**0.952**	**0.860**	**0.647**	**0.938**	**0.792**	**0.530**

Notes: 1) *h* denotes the forecast horizon in trading day units. The forecast period is from 2002-01-21 to 2011-12-30, a total of 2519 trading days.

2) The ratios of forecast RMSE and MAE of the two adaptive models versus the random walk model are reported. The two adaptive models are: the LAR model for each single IV, and the ADNS for the whole TSIV with three factors, with each factor based on an LAR process.

3) For the forecast ratio comparison, a number smaller than 1 is marked in bold, indicating a better forecast than the adaptive model. If the number is between 0.8 and 0.9, then the bold number is underlined; if the number is smaller than 0.8, then a grey shadow is added for emphasis.

Table 4. Forecast Comparison of Alternative Models with Random Walk

4-a. RMSE Ratios versus Random Walk

h = 1

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13
y(30)	1.001	**0.996**	1.000	1.001	**0.998**	**0.974**	**0.974**	**0.973**	**0.991**	**0.996**	**0.979**	**0.984**	**0.980**
y(91)	1.043	1.015	**0.996**	**1.000**	**0.999**	**0.975**	**0.981**	**0.979**	1.013	1.019	**0.989**	**0.991**	**0.989**
y(152)	1.030	1.008	**0.997**	**1.000**	1.001	**0.977**	**0.981**	**0.979**	1.014	1.022	**0.990**	**0.992**	**0.991**
y(365)	1.084	1.027	**1.000**	1.000	**0.999**	**0.977**	**0.980**	**0.979**	1.015	**0.999**	**0.988**	**0.989**	**0.987**
y(730)	1.077	1.059	1.001	1.001	1.002	**0.984**	**0.988**	**0.987**	1.069	1.226	1.003	**0.996**	**0.999**

h = 5

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13
y(30)	**0.992**	**0.992**	1.001	**0.999**	**0.996**	**0.958**	**0.965**	**0.962**	**0.983**	**0.997**	**0.972**	**0.978**	**0.974**
y(91)	1.033	1.019	**0.998**	**1.000**	1.009	**0.976**	**0.980**	**0.976**	1.009	1.014	**0.990**	**0.996**	**0.991**
y(152)	1.017	1.004	1.003	1.001	1.010	**0.978**	**0.982**	**0.978**	**0.997**	1.008	**0.989**	**0.995**	**0.990**
y(365)	1.049	1.012	1.005	1.001	1.012	**0.984**	**0.991**	**0.987**	1.002	1.009	**0.992**	**0.997**	**0.994**
y(730)	1.038	1.032	1.006	1.003	1.015	**0.994**	1.002	**0.997**	1.034	1.079	1.002	1.004	1.002

h = 20

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13
y(30)	**0.991**	**1.000**	1.003	**0.989**	1.010	1.021	1.005	1.003	**0.982**	1.006	**0.984**	**0.995**	**0.987**
y(91)	1.026	1.021	1.016	1.000	1.025	1.022	1.021	1.022	**0.999**	1.018	1.002	1.013	1.005
y(152)	1.030	1.020	1.026	1.005	1.031	1.032	1.026	1.026	**0.997**	1.024	1.008	1.017	1.012
y(365)	1.051	1.023	1.026	1.007	1.033	1.037	1.034	1.036	**0.998**	1.028	1.012	1.020	1.015
y(730)	1.038	1.023	1.018	1.008	1.024	1.041	1.038	1.037	1.005	1.032	1.008	1.014	1.009

4-b. MAE Ratios versus Random Walk

h = 1

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13
y(30)	1.006	1.001	1.000	1.004	1.001	**0.980**	**0.979**	**0.980**	**0.991**	**0.999**	**0.981**	**0.987**	**0.982**
y(91)	1.070	1.031	**0.996**	1.001	1.003	**0.980**	**0.988**	**0.982**	1.024	1.035	**0.994**	**0.994**	**0.993**
y(152)	1.031	1.005	**0.999**	1.000	1.002	**0.984**	**0.987**	**0.986**	1.010	1.020	**0.989**	**0.991**	**0.989**
y(365)	1.125	1.050	1.004	1.000	1.003	**0.988**	**0.990**	**0.989**	1.023	1.002	**0.994**	**0.994**	**0.992**
y(730)	1.099	1.088	1.003	1.001	1.005	**0.994**	**0.999**	**0.997**	1.101	1.320	1.013	**0.998**	1.006

h = 5

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13
y(30)	1.005	1.001	1.001	1.014	**1.000**	**0.965**	**0.961**	**0.968**	**0.989**	1.003	**0.975**	**0.982**	**0.976**
y(91)	1.053	1.027	**0.995**	1.011	1.015	**0.973**	**0.978**	**0.975**	1.020	1.022	**0.994**	1.000	**0.994**
y(152)	1.038	1.010	1.004	1.012	1.017	**0.981**	**0.981**	**0.982**	1.011	1.016	**0.995**	1.001	**0.996**
y(365)	1.075	1.026	1.005	1.008	1.016	**0.987**	**0.988**	**0.984**	1.012	1.013	**0.997**	**1.000**	**0.997**
y(730)	1.049	1.030	1.007	1.009	1.014	1.002	**1.000**	**0.994**	1.038	1.086	1.001	1.001	1.001

h = 20

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13
y(30)	1.045	1.039	1.005	1.037	1.049	**1.000**	1.024	1.024	1.011	1.038	1.000	1.024	1.005
y(91)	1.086	1.053	1.004	1.034	1.059	1.019	1.038	1.031	1.034	1.044	1.016	1.037	1.019
y(152)	1.079	1.035	1.017	1.027	1.046	1.024	1.031	1.030	1.023	1.035	1.015	1.030	1.017
y(365)	1.093	1.023	1.023	1.017	1.026	1.036	1.028	1.028	1.009	1.016	1.010	1.021	1.011
y(730)	1.049	**0.998**	1.018	1.013	1.003	1.042	1.031	1.028	**0.997**	**0.999**	**0.998**	1.003	**0.998**

Notes: 1) h denotes the forecast horizon in trading day units. The forecast period is from 2002-01-21 to 2011-12-30, a total of 2519 trading days.

2) For the forecast ratio comparison between an alternative model and the random walk model, a number **smaller** than 1 is marked in bold, indicating a **better** forecast than the random walk.

Table 5. Out-of-sample R^2 Comparison versus Random Walk

h = 1

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	LAR	ADNS
y(30)	1.000	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	0.995
y(91)	0.999	0.999	0.999	0.998	0.998	1.000	1.000	1.000	0.999	0.998	1.000	1.000	1.000	0.998	0.995
y(152)	0.999	0.999	0.999	0.999	0.999	1.000	1.000	1.000	0.999	0.999	1.000	1.000	1.000	0.999	0.997
y(365)	0.999	0.999	0.999	0.998	0.999	1.000	1.000	1.000	0.999	0.999	1.000	1.000	1.000	0.999	0.998
y(730)	0.999	0.999	0.999	0.998	0.998	1.000	0.999	0.999	0.999	0.995	0.999	0.999	0.999	0.999	0.998

h = 5

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	LAR	ADNS
y(30)	0.999	0.998	0.999	0.999	1.000	**1.008**	**1.006**	**1.006**	**1.003**	0.999	**1.005**	**1.003**	**1.005**	**1.039**	0.998
y(91)	0.999	0.997	0.999	0.993	0.996	**1.003**	**1.001**	0.992	0.998	0.997	1.000	0.999	0.999	**1.019**	0.991
y(152)	0.999	0.998	0.999	0.996	0.998	1.001	1.000	1.000	1.000	0.998	1.000	0.999	1.000	**1.016**	0.997
y(365)	0.999	0.999	0.999	0.996	0.998	1.000	0.999	0.999	1.000	0.999	1.000	0.999	1.000	**1.014**	**1.005**
y(730)	0.999	0.999	0.999	0.998	0.998	1.000	0.999	0.999	0.998	0.995	0.999	0.999	0.999	**1.012**	**1.007**

h = 20

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	LAR	ADNS
y(30)	**1.004**	0.980	0.997	**1.007**	0.994	0.981	0.989	0.990	**1.015**	0.990	**1.012**	**1.001**	**1.009**	**1.319**	**1.027**
y(91)	0.997	0.980	0.986	0.985	0.983	0.987	0.984	1.004	0.998	0.985	0.999	0.994	0.997	**1.200**	**1.000**
y(152)	0.996	0.980	0.983	0.986	0.986	0.987	0.983	0.982	0.998	0.983	0.993	0.988	0.991	**1.161**	**1.001**
y(365)	0.997	0.985	0.985	0.988	0.988	0.989	0.985	0.984	1.000	0.986	0.994	0.991	0.993	**1.124**	**1.005**
y(730)	0.998	0.986	0.985	0.995	0.987	0.991	0.982	0.982	0.998	0.984	0.994	0.990	0.992	**1.198**	**1.004**

Notes: 1) *h* denotes the forecast horizon in trading day units. The forecast period is from 2002-01-21 to 2011-12-30, a total of 2519 trading days.

2) For the forecast ratio comparison between an alternative model and the random walk model, a number **bigger** than 1 is marked in bold, indicating a **better** forecast than the random walk.

Table 6. Frequencies of Correct Sign Prediction for Future Changes (Unit: %)

h = 1

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	LAR	ADNS
y(30)	51.7	53.0	45.4	51.9	53.3	56.6	57.2	57.4	52.4	53.1	55.2	53.2	54.7	**65.0**	64.9
y(91)	50.7	51.3	53.2	52.0	52.4	56.5	55.5	56.9	50.7	50.9	54.5	53.4	54.1	63.4	**64.2**
y(152)	50.3	52.1	51.8	51.4	51.3	56.2	55.4	55.9	51.1	51.3	53.6	53.2	53.4	**63.5**	63.3
y(365)	51.5	51.6	51.1	52.0	51.5	55.0	54.0	54.4	52.0	52.4	52.4	52.7	52.4	**63.2**	59.9
y(730)	49.0	49.9	50.7	52.3	52.8	55.7	55.2	55.7	49.8	50.1	50.9	53.3	51.0	**65.9**	59.0
Average	50.7	51.6	51.2	51.8	51.7	56.0	54.9	55.2	51.0	51.5	53.2	52.9	53.1	**63.9**	62.7

h = 5

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	LAR	ADNS
y(30)	52.8	53.0	46.0	52.9	53.1	57.7	58.6	57.2	53.1	52.7	55.6	53.4	54.7	71.9	**73.6**
y(91)	52.4	52.8	53.1	52.0	51.4	57.3	56.1	56.5	52.0	51.1	53.5	52.8	53.5	**73.0**	72.4
y(152)	51.2	51.3	51.1	51.6	51.6	55.1	56.6	55.1	51.7	50.3	53.1	52.2	52.8	72.4	**73.6**
y(365)	51.6	52.6	50.0	51.4	51.7	53.4	55.5	55.4	51.8	51.8	53.9	53.1	53.4	**72.7**	72.4
y(730)	47.6	50.0	49.1	51.0	51.4	51.8	53.0	53.8	51.4	50.6	51.1	51.8	51.3	**77.3**	70.8
Average	51.2	51.9	50.7	51.7	51.6	54.9	55.4	55.1	51.8	51.1	53.2	52.3	53.0	**73.7**	72.6

h = 20

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	LAR	ADNS
y(30)	53.9	53.2	43.0	55.2	53.1	54.2	53.2	53.5	54.2	54.1	54.0	53.7	53.8	82.8	**82.9**
y(91)	53.1	53.1	52.4	54.4	52.6	50.2	51.9	52.1	52.8	54.4	53.2	52.8	53.1	**85.4**	83.9
y(152)	51.7	55.0	49.3	53.3	53.6	49.1	51.0	50.7	54.4	55.5	52.6	53.0	52.8	**86.9**	85.4
y(365)	52.0	54.5	45.7	54.0	53.4	44.9	49.0	49.8	54.1	52.6	54.5	53.3	53.9	**89.5**	87.6
y(730)	51.0	55.5	39.4	54.7	55.9	43.7	46.4	47.0	56.0	56.1	54.5	53.8	54.2	**90.2**	88.6
Average	52.3	54.1	47.2	54.2	53.4	48.2	50.0	50.5	54.0	54.1	53.6	53.1	53.5	**87.1**	85.9

Notes: 1) *h* denotes the forecast horizon in trading day units. The forecast period is from 2002-01-21 to 2011-12-30, a total of 2519 trading days.

2) The frequencies (%) of correct sign prediction for each model are reported for different maturities of IV and different forecast horizons. In each row of a given maturity and forecast horizon, the best forecast (the highest number) is marked in bold.

Table7. Test of Forecasting Performance against Alternative Models

7-a. LAR against M1-M13 with DM Test and against Random Walk with CW Test

h = 1

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	R.W.(CW)
y(30)	1.63**	1.94**	4.28***	3.78***	1.41*	1.96**	3.43***	3.22***	4.01***	2.18**	4.79***	4.01***	4.69***	0.69
y(91)	3.18***	2.35***	3.86***	3.90***	-4.21***	-1.91**	2.98***	3.25***	-1.12	-2.26**	3.27***	3.44***	3.33***	1.67**
y(152)	5.19***	3.89***	5.10***	5.01***	1.82**	3.57***	4.80***	4.57***	3.19***	0.68	5.46***	5.50***	5.46***	2.34**
y(365)	5.66***	5.03***	5.27***	5.47***	-6.12***	-2.27*	5.30***	5.34***	1.41*	5.09***	5.10***	5.84***	5.48***	3.33***
y(730)	6.29***	4.86***	6.58***	6.44***	-5.56***	-5.90***	3.95***	4.15***	-5.14***	-17.28***	1.66**	5.45***	3.39***	5.53***

h = 5

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	R.W.(CW)
y(30)	-11.69***	-11.01***	-10.08***	-9.79***	-11.19***	-11.17***	-9.36***	-9.29***	-10.12***	-11.06***	-9.50***	-10.05***	-9.60***	-9.48***
y(91)	-11.43***	-12.00***	-10.75***	-9.88***	-13.22***	-12.59***	-10.44***	-14.26***	-11.92***	-12.02***	-11.00***	-11.32***	-11.07***	-5.42***
y(152)	-12.20***	-12.57***	-11.84***	-11.10***	-12.72***	-12.10***	-11.32***	-11.37***	-11.99***	-12.19***	-11.47***	-11.73***	-11.60***	-6.10***
y(365)	-12.15***	-12.40***	-11.97***	-11.54***	-14.51***	-13.05***	-11.68***	-11.70***	-12.27***	-12.22***	-11.73***	-11.73***	-11.79***	-6.96***
y(730)	-13.02***	-13.15***	-12.97***	-12.59***	-14.31***	-13.94***	-13.16***	-13.06***	-14.30***	-16.35***	-13.17***	-13.07***	-13.13***	-7.11***

h = 20

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	R.W.(CW)
y(30)	-26.70***	-25.68***	-24.08***	-23.29***	-25.66***	-25.81***	-23.89***	-22.97***	-25.04***	-25.55***	-23.83***	-24.82***	-23.92***	-13.56***
y(91)	-28.62***	-28.260***	-25.66***	-26.65***	-29.09***	-28.10***	-26.56***	-26.92***	-28.58***	-27.81***	-26.67***	-27.42***	-26.76***	-14.29***
y(152)	-29.37***	-28.70***	-27.23***	-27.98***	-29.84***	-28.21***	-27.39***	-27.15***	-29.40***	-28.11***	-27.55***	-28.06***	-27.63***	-15.96***
y(365)	-30.44***	-30.07***	-29.57***	-30.66***	-31.44***	-29.96***	-29.38***	-29.14***	-30.75***	-29.38***	-29.57***	-29.69***	-29.61***	-16.72***
y(730)	-33.27***	-32.82***	-32.63***	-33.96***	-33.17***	-32.09***	-33.06***	-32.93***	-33.69***	-32.76***	-32.87***	-32.77***	-32.86***	-18.95***

7-b. ADNS against M1-M13 with DM Test and against Random Walk with CW Test

h = 1

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	R.W. (CW)
y(30)	2.45***	1.99**	3.21***	1.82**	2.13**	2.97***	2.62***	2.65***	3.03***	2.17**	3.33***	2.83**	3.26***	-1.64**
y(91)	5.19***	4.51***	4.53***	5.48***	1.93***	3.16***	5.04***	4.71***	4.16***	3.47***	5.16***	5.06***	5.13***	0.07*
y(152)	5.33***	4.71***	4.93***	5.33***	3.55***	4.52***	5.06***	4.98***	4.98***	3.39***	5.35***	5.35***	5.33***	0.86
y(365)	5.44***	5.85***	4.89***	5.66***	3.45***	5.23***	6.11***	6.05***	5.54***	5.48***	6.38***	6.10***	6.30***	0.10
y(730)	7.83***	7.79***	7.73***	7.94***	5.90***	5.66***	8.04***	7.78***	6.40***	-4.48**	8.08***	8.19***	8.26***	2.01**

h = 5

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	R.W. (CW)
y(30)	-8.03**	-7.33***	-6.74***	-5.84***	-7.53***	-7.45***	-5.75***	-5.69***	-6.65***	-7.39***	-6.02***	-6.49***	-6.12***	-7.62***
y(91)	-5.90***	-6.28***	-5.23***	-4.35***	-7.73***	-6.93***	-4.82***	-8.13***	-6.43***	-6.51***	-5.44***	-5.73***	-5.52***	-7.73***
y(152)	-5.72***	-6.05***	-5.42***	-4.62***	-6.46***	-5.66***	-4.85***	-4.89***	-5.62***	-5.89***	-5.06***	-5.29***	-5.17***	-8.32***
y(365)	-7.73***	-8.07***	-7.55***	-6.86***	-10.96***	-8.97***	-7.11***	-7.13***	-7.96***	-7.86***	-7.33***	-7.36***	-7.38***	-10.54***
y(730)	-2.45***	-2.71***	-2.66***	-2.26***	-3.22***	-3.19***	-2.82***	-2.80***	-3.09***	-4.48***	-2.61***	-2.66***	-2.63***	-11.52***

h = 20

y(n)\Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	R.W. (CW)
y(30)	-7.19**	-6.72***	-5.70***	-5.01***	-6.83***	-6.82***	-5.04***	-4.98***	-5.72***	-6.78***	-5.10***	-5.68***	-5.22***	-9.60***
y(91)	-3.98***	-4.52***	-3.35***	-2.38***	-6.12***	-5.26***	-2.93***	-6.83***	-4.57***	-4.76***	-3.55***	-3.87***	-3.64***	-9.61***
y(152)	-3.52***	-3.98***	-3.28***	-2.39***	-4.41***	-3.54***	-2.68***	-2.74***	-3.39***	-3.82***	-2.85***	-3.11***	-2.97***	-10.18***
y(365)	-6.02***	-6.48***	-5.95***	-5.25***	-9.23***	-7.33***	-5.58***	-5.60***	-6.14***	-6.30***	-5.66***	-5.74***	-5.73***	-12.44***
y(730)	-4.44***	-4.84***	-4.53***	-4.10***	-6.04***	-5.93***	-4.84***	-4.74***	-5.99***	-9.02***	-4.74***	-4.70***	-4.71***	-11.88***

Notes: 1) *h* denotes the forecast horizon in trading day units. The forecast period is from 2002-01-21 to 2011-12-30, a total of 2519 trading days.

2) Diebold-Mariano (DM) test statistics (Diebold and Mariano, 1995) are provided for LAR/ADNS against M1 to M13. Clark-West (CW) test statistics are provided for LAR/ADNS against the random walk model. For both test statistics, a negative (positive) value indicates better (worse) accuracy of LAR/ADNS against an alternative model in terms of quadratic loss measured by mean-squared predicted errors (MSPE). Significance of 10%, 5% and 1% are marked by *, **, and ***, respectively.

Table 8. Average Lengths of Homogenous Intervals Detected Adaptively

(Unit: trading days)

	$h = 1$	$h = 5$	$h = 20$
NS1	30	28	18
NS2	39	36	31
NS3	19	26	22

Notes: *h* denotes the forecast horizon.

Table 9. Robustness Check with Alternative Data on Forecast Performance against Random Walk using CW Test

		LAR			ADNS		
	y(n)\h	1	5	20	1	5	20
IVs Delta=0.4	y(30)	-0.90	-7.78***	-14.01***	-2.39***	-6.56***	-13.86***
	y(60)	-0.12	-8.35***	14.14***	-0.22	-6.48***	-13.75***
	y(91)	0.21	-9.24***	-14.83***	0.12	-7.35***	-14.71***
	y(122)	0.06	--9.56***	-15.36***	0.29	-7.53***	-15.10***
	y(152)	0.79	-9.94***	-16.34***	0.35	-7.89***	-15.37***
	y(182)	0.92	-10.96***	16.59***	0.29	-8.56***	-16.14***
	y(273)	1.00	-11.27***	18.56***	-0.35	-9.03***	-18.11***
	y(365)	-0.32	-11.03***	-20.88***	-0.19	-8.99***	-18.71***
	y(547)	0.23	-11.72***	-20.88***	3.60***	-9.94***	-20.21***
	y(730)	0.74	-12.84***	-22.14***	2.48	-11.26***	-21.38***
	y(n)\h	1	5	20	1	5	20
IVs Delta=0.6	y(30)	-2.11**	-8.46***	-13.67***	-3.78***	-6.68***	-11.97***
	y(60)	-0.20	-8.73***	-13.75***	-1.50*	-6.57***	-11.79***
	y(91)	-0.02	-8.97***	-13.67***	-1.46*	-7.16***	-12.27***
	y(122)	0.15	-9.48***	-14.73***	-1.48*	-7.12***	-13.20***
	y(152)	0.75	-9.47***	-15.27***	-1.25	-7.54***	-13.58***
	y(182)	0.03	-10.59***	-15.61***	-0.92	-8.41***	-14.12***
	y(273)	0.29	-11.70***	-17.85***	-2.13**	-9.72***	-16.68***
	y(365)	0.51	-12.31***	-18.44***	-1.78**	-10.18***	-17.84***
	y(547)	-0.36	-13.29***	-20.52***	0.97	-11.63***	-19.66***
	y(730)	0.09	-13.83***	-21.82***	1.75**	-12.25***	-21.00***
	y(m)\h	1	5	20	1	5	20
VIX Futures	y(1m)	-1.73**	-10.88***	-12.49***	-3.34***	-12.96***	-13.07***
	y(2m)	0.53	-9.96***	-12.31***	-1.20	-11.53***	-12.64***
	y(3m)	0.07	-10.94***	-14.46***	-2.46**	-11.84***	-13.69***
	y(4m)	-0.83	-11.717***	-15.22***	-3.94***	-14.78***	-14.56***
	y(5m)	0.29	-12.04***	-15.96***	-4.17***	-15.58***	-16.01***
	y(6m)	-1.35*	-10.27***	-15.91***	-4.61***	-14.75***	-15.28***
	y(7m)	-2.41***	-10.60***	-16.66***	-4.98***	-15.30***	-16.38***
	y(8m)	-2.94***	-12.95***	-17.95***	-4.97***	-14.54***	-19.00***
	y(9m)	-3.01***	-11.04***	-16.98***	-5.76***	-14.43***	-18.44***
	y(10m)	-2.91***	-9.88***	-16.58***	-6.10***	-13.46***	-18.20***

Notes: 1) h denotes the forecast horizon in trading day units. The forecast period for the IVs data is from 2002-01-21 to 2011-12-30. The forecast period for VIX future is from 2012-05-02 to 2017-09-01.
2) Clark-West (CW) test statistics are provided for LAR/ADNS against the random walk model. A negative (positive) value indicates better (worse) accuracy. Significance of 10%，5% and 1% are marked by *, **, and ***, respectively.

IRTG 1792 Discussion Paper Series 2018

For a complete list of Discussion Papers published, please visit irtg1792.hu-berlin.de.

001 "Data Driven Value-at-Risk Forecasting using a SVR-GARCH-KDE Hybrid" by Marius Lux, Wolfgang Karl Härdle and Stefan Lessmann, January 2018.
002 "Nonparametric Variable Selection and Its Application to Additive Models" by Zheng-Hui Feng, Lu Lin, Ruo-Qing Zhu asnd Li-Xing Zhu, January 2018.
003 "Systemic Risk in Global Volatility Spillover Networks: Evidence from Option-implied Volatility Indices " by Zihui Yang and Yinggang Zhou, January 2018.
004 "Pricing Cryptocurrency options: the case of CRIX and Bitcoin" by Cathy YH Chen, Wolfgang Karl Härdle, Ai Jun Hou and Weining Wang, January 2018.
005 "Testing for bubbles in cryptocurrencies with time-varying volatility" by Christian M. Hafner, January 2018.
006 "A Note on Cryptocurrencies and Currency Competition" by Anna Almosova, January 2018.
007 "Knowing me, knowing you: inventor mobility and the formation of technology-oriented alliances" by Stefan Wagner and Martin C. Goossen, February 2018.
008 "A Monetary Model of Blockchain" by Anna Almosova, February 2018.
009 "Deregulated day-ahead electricity markets in Southeast Europe: Price forecasting and comparative structural analysis" by Antanina Hryshchuk, Stefan Lessmann, February 2018.
010 "How Sensitive are Tail-related Risk Measures in a Contamination Neighbourhood?" by Wolfgang Karl Härdle, Chengxiu Ling, February 2018.
011 "How to Measure a Performance of a Collaborative Research Centre" by Alona Zharova, Janine Tellinger-Rice, Wolfgang Karl Härdle, February 2018.
012 "Targeting customers for profit: An ensemble learning framework to support marketing decision making" by Stefan Lessmann, Kristof Coussement, Koen W. De Bock, Johannes Haupt, February 2018.
013 "Improving Crime Count Forecasts Using Twitter and Taxi Data" by Lara Vomfell, Wolfgang Karl Härdle, Stefan Lessmann, February 2018.
014 "Price Discovery on Bitcoin Markets" by Paolo Pagnottoni, Dirk G. Baur, Thomas Dimpfl, March 2018.
015 "Bitcoin is not the New Gold - A Comparison of Volatility, Correlation, and Portfolio Performance" by Tony Klein, Hien Pham Thu, Thomas Walther, March 2018.
016 "Time-varying Limit Order Book Networks" by Wolfgang Karl Härdle, Shi Chen, Chong Liang, Melanie Schienle, April 2018.
017 "Regularization Approach for NetworkModeling of German EnergyMarket" by Shi Chen, Wolfgang Karl Härdle, Brenda López Cabrera, May 2018.
018 "Adaptive Nonparametric Clustering" by Kirill Efimov, Larisa Adamyan, Vladimir Spokoiny, May 2018.
019 "Lasso, knockoff and Gaussian covariates: a comparison" by Laurie Davies, May 2018.

IRTG 1792, Spandauer Straße 1, D-10178 Berlin
http://irtg1792.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the IRTG 1792.

IRTG 1792 Discussion Paper Series 2018

For a complete list of Discussion Papers published, please visit irtg1792.hu-berlin.de.

020 "A Regime Shift Model with Nonparametric Switching Mechanism" by Haiqiang Chen, Yingxing Li, Ming Lin and Yanli Zhu, May 2018.
021 "LASSO-Driven Inference in Time and Space" by Victor Chernozhukov, Wolfgang K. Härdle, Chen Huang, Weining Wang, June 2018.
022 " Learning from Errors: The case of monetary and fiscal policy regimes" by Andreas Tryphonides, June 2018.
023 "Textual Sentiment, Option Characteristics, and Stock Return Predictability" by Cathy Yi-Hsuan Chen, Matthias R. Fengler, Wolfgang Karl Härdle, Yanchu Liu, June 2018.
024 "Bootstrap Confidence Sets For Spectral Projectors Of Sample Covariance" by A. Naumov, V. Spokoiny, V. Ulyanov, June 2018.
025 "Construction of Non-asymptotic Confidence Sets in 2 -Wasserstein Space" by Johannes Ebert, Vladimir Spokoiny, Alexandra Suvorikova, June 2018.
026 "Large ball probabilities, Gaussian comparison and anti-concentration" by Friedrich Götze, Alexey Naumov, Vladimir Spokoiny, Vladimir Ulyanov, June 2018.
027 "Bayesian inference for spectral projectors of covariance matrix" by Igor Silin, Vladimir Spokoiny, June 2018.
028 "Toolbox: Gaussian comparison on Eucledian balls" by Andzhey Koziuk, Vladimir Spokoiny, June 2018.
029 "Pointwise adaptation via stagewise aggregation of local estimates for multiclass classification" by Nikita Puchkin, Vladimir Spokoiny, June 2018.
030 "Gaussian Process Forecast with multidimensional distributional entries" by Francois Bachoc, Alexandra Suvorikova, Jean-Michel Loubes, Vladimir Spokoiny, June 2018.
031 "Instrumental variables regression" by Andzhey Koziuk, Vladimir Spokoiny, June 2018.
032 "Understanding Latent Group Structure of Cryptocurrencies Market: A Dynamic Network Perspective" by Li Guo, Yubo Tao and Wolfgang Karl Härdle, July 2018.
033 "Optimal contracts under competition when uncertainty from adverse selection and moral hazard are present" by Natalie Packham, August 2018.
034 "A factor-model approach for correlation scenarios and correlation stress-testing" by Natalie Packham and Fabian Woebbeking, August 2018.
035 "Correlation Under Stress In Normal Variance Mixture Models" by Michael Kalkbrener and Natalie Packham, August 2018.
036 "Model risk of contingent claims" by Nils Detering and Natalie Packham, August 2018.
037 "Default probabilities and default correlations under stress" by Natalie Packham, Michael Kalkbrener and Ludger Overbeck, August 2018.
038 "Tail-Risk Protection Trading Strategies" by Natalie Packham, Jochen Papenbrock, Peter Schwendner and Fabian Woebbeking, August 2018.

IRTG 1792, Spandauer Straße 1, D-10178 Berlin
http://irtg1792.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the IRTG 1792.

IRTG 1792 Discussion Paper Series 2018

For a complete list of Discussion Papers published, please visit irtg1792.hu-berlin.de.

039 "Penalized Adaptive Forecasting with Large Information Sets and Structural Changes" by Lenka Zbonakova, Xinjue Li and Wolfgang Karl Härdle, August 2018.
040 "Complete Convergence and Complete Moment Convergence for Maximal Weighted Sums of Extended Negatively Dependent Random Variables" by Ji Gao YAN, August 2018.
041 "On complete convergence in Marcinkiewicz-Zygmund type SLLN for random variables" by Anna Kuczmaszewska and Ji Gao YAN, August 2018.
042 "On Complete Convergence in Marcinkiewicz-Zygmund Type SLLN for END Random Variables and its Applications" by Ji Gao YAN, August 2018.
043 "Textual Sentiment and Sector specific reaction" by Elisabeth Bommes, Cathy Yi-Hsuan Chen and Wolfgang Karl Härdle, September 2018.
044 "Understanding Cryptocurrencies" by Wolfgang Karl Härdle, Campbell R. Harvey, Raphael C. G. Reule, September 2018.
045 "Predicative Ability of Similarity-based Futures Trading Strategies" by Hsin-Yu Chiu, Mi-Hsiu Chiang, Wei-Yu Kuo, September 2018.
046 "Forecasting the Term Structure of Option Implied Volatility: The Power of an Adaptive Method" by Ying Chen, Qian Han, Linlin Niu, September 2018.

IRTG 1792, Spandauer Straße 1, D-10178 Berlin
http://irtg1792.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the IRTG 1792.