The Journal of Financial Research • Vol. XXXII, No. 3 • Pages 231–259 • Fall 2009

THE ECONOMIC VALUE OF USING REALIZED VOLATILITY IN FORECASTING FUTURE IMPLIED VOLATILITY

Wing Hong Chan
Wilfrid Laurier University

Ranjini Jha
University of Waterloo

Madhu Kalimipalli
Wilfrid Laurier University
City University of Hong Kong

Abstract

We examine the economic benefits of using realized volatility to forecast future implied volatility for pricing, trading, and hedging in the S&P 500 index options market. We propose an encompassing regression approach to forecast future implied volatility, and hence future option prices, by combining historical realized volatility and current implied volatility. Although the use of realized volatility results in superior performance in the encompassing regressions and out-of-sample option pricing tests, we do not find any significant economic gains in option trading and hedging strategies in the presence of transaction costs.

JEL Classification: G10, G14

I. Introduction

We examine the incremental economic value of using intraday realized volatility (RV) in forecasting implied volatility (IV) in the context of option pricing, volatility trading, and hedging in the S&P 500 index options markets. We employ an

The authors thank the editor, Gerald Gay, and the referee, George Jiang, for their valuable comments that greatly improved the quality of this article. The authors also thank seminar participants at the Bank of Canada, McMaster University, University of Houston, University of Toronto, York University, Northern Finance Association Meetings 2004, Financial Management Association Meetings 2004, Annual Summer Conference at the Indian School of Business (ISB) 2005, the Eastern Finance Association Meetings 2005, and the Financial Management Association European Conference 2007 for their suggestions. This article received the Best Paper Award (Derivatives Category) at the Northern Finance Association Meetings 2004. The authors gratefully acknowledge Phelim Boyle, Melanie Cao, Peter Christoffersen, Kris Jacobs, Ravi Jagannathan, Mark Kamstra, Raymond Kan, Michael King, Praveen Kumar, Thomas McCurdy, Raul Susmel, and Yisong Tian for helpful comments that have significantly improved the article. This research is supported by the Social Sciences and Humanities Research Council of Canada.

encompassing regression approach to forecast future IVs by combining historical RV and current IV. Forecasting IVs is of interest to option market participants, who routinely formulate volatility and option price forecasts for trading and hedging purposes. We find that although the combination models outperform stand-alone IV models statistically in the encompassing regression and out-of-sample option pricing tests, they do not render significant economic gains in option trading and hedging strategies in the presence of transaction costs. Our study implies that the statistical superiority of RV may not result in economic benefits for trading and risk management decisions in the options markets in the presence of market frictions.

RV refers to the volatility measure based on high-frequency intraday returns. Many studies document the statistical superiority of RV;[1] however, only few examine its economic implications. Fleming, Kirby, and Ostdiek (2003) find that RV has significant economic value in volatility timing in asset allocation decisions in the equity and bond markets, and Giot and Laurent (2004) study value-at-risk strategies. We add to this literature by empirically examining the economic relevance of high-frequency volatility measures for out-of-sample pricing, trading, and hedging in the index options market.

We also provide a parsimonious specification to forecast future IVs, and hence future option prices, by combining information from historical RV with the forward-looking option IVs. Although numerous studies focus on predicting future RV using options data,[2] few look at predicting IV. The IV forecasts provide vital information to option traders in formulating their conditional volatility and option price forecasts.

II. Data and Summary Statistics

Following the previous literature (e.g., Bakshi, Cao, and Chen 1997; Christoffersen and Jacobs 2004), we examine S&P 500 index (SPX) options. SPX options are the second most liquid options in terms of trading volume and most liquid in terms of open interest in the U.S. markets; furthermore, SPX options are easier to value as they are European in nature and have no wild card options.

The SPX intraday option data for the sample period, January 1996 to December 2002, are obtained from the Chicago Board Options Exchange (CBOE).

[1]See Andersen and Benzoni (2009) for an overview of the realized volatility concept. Previous research shows that RV is log normally distributed and RV standardized daily returns are approximately normal. RV is found to have long-memory properties and is best described by a fractionally integrated autoregressive process. These findings are consistent across asset classes (e.g., for exchange rates, see Andersen, Bollerslev, Diebold, and Labys 2001, 2003; for individual stocks, see Andersen, Bollerslev, Diebold, and Ebens 2001).

[2]For a recent daily volatility study, see Jiang and Tian (2005), and for recent intraday volatility research, see Pong et al. (2004), Martens and Zein (2004), and Ghysels, Santa-Clara, and Valkanov (2006).

As in Nandi (1998), the S&P 500 daily dividend is used as a proxy for the underlying dividends. We choose the final trading hour window for options each day, as it tends to have the most active trading (e.g., Jiang and Tian 2005). For each option trading between 2 p.m. and 3 p.m. every day, we identify the corresponding underlying spot index data. We use the earliest occurring spot transaction before a given option trade as its matching spot value.

Eurodollar interest rates with maturities of 1 month to 12 months are obtained from Datastream, and interest rates for any given option maturity option are obtained by linearly interpolating the two closest maturities (time to maturity is annualized using the 365 days convention).

Filters are applied separately to the call and put option data, and options that meet the following criteria are dropped: (1) options that violate the arbitrage bounds (i.e., upper and lower boundaries), (2) options below the minimum tick size of 3/8, (3) options with maturities less than 6 days, and (4) options with IV above 2. Based on these criteria, our initial sample consists of 86,279 calls and 110,274 puts.

Table 1 provides the following summary statistics for the options data: average price, moneyness, maturity, and IV for options in each moneyness-maturity subgroup for the daily 2 p.m.–3 p.m. window data for January 1996 to December 2002. We report summary statistics for subsamples relating to short-term (6–60 days), medium-term (61–180 days), and long-term ($>$180 days) maturities and three moneyness levels. Out-of-the-money (OTM) options are those with moneyness, that is, strike price/spot price or X/S above 1.03 for calls and below 0.97 for puts; at-the-money (ATM) options, X/S between 0.97 and 1.03, and in-the-money (ITM) options, X/S below 0.97 for calls and above 1.03 for puts.

The moneyness and maturity related biases are evident from Table 1. For example, OTM-put IVs are higher than ITM-put IVs, especially for shorter maturities. Most of the option trading seems to be concentrated in short-term ATM and OTM calls and puts. Trades for OTM puts exceed those of ATM puts, indicating that the former have been widely used as crash insurance. In general, trading is thin for ITM options and for options with maturities beyond 60 days. In our subsequent analysis, we exclude ITM options and all options with medium and long maturities, as thin trading could mask the informativeness of prices for these trades and exacerbate the bid–ask spreads. Our final option data sample consists of 60,825 calls and 76,176 puts.

The spot market data consist of both intraday and daily S&P 500 index return data for the seven-year period from January 1996 to December 2002, for a total of 1,738 trading days. The intraday data are obtained from CBOE and the daily data are extracted from Datastream. The intraday data are used to calculate daily RV as the sum of squared returns at finely sampled intervals over a given day added to the square of the previous overnight return as in Blair, Poon, and Taylor (2001). Choosing a very high sampling frequency may introduce a bias in

TABLE 1. Sample Statistics for the S&P 500 Index Option Data (January 1996–December 2002).

		Call Options					Put Options			
		Maturity					Maturity			
		6–60 Days	60–180 Days	≥180 Days	All Maturities		6–60 Days	60–180 Days	≥180 Days	All Maturities
No. of obs.		29,978	5,678	773	36,429		45,144	11,406	1,509	58,059
Average price		5.11	12.81	26.28	6.27		5.07	11.14	21.39	6.01
Average maturity	OTM: $X/S \geq 1.03$	28.3	94.58	244.04	41	OTM: $X/S \leq 0.97$	28.24	99	231.99	43.82
Average moneyness		1.06	1.1	1.15	1.07		0.9	0.85	0.84	0.89
Average IV		0.19	0.19	0.23	0.19		0.29	0.27	0.27	0.29
No. of obs.		31,847	2,290	93	34,230		31,032	3,069	273	34,374
Average price		14.67	35.64	68.85	13.76		14.68	32.98	54.46	14.27
Average maturity	ATM: $0.97 <$	21.44	88.8	241.9	21.04	ATM: $0.97 <$	22.11	92.36	244.11	24.86
Average moneyness	$X/S < 1.03$	1.01	1.01	1.01	1.01	$X/S < 1.03$	0.99	0.99	1	0.99
Average IV		0.19	0.2	0.18	0.2		0.19	0.2	0.29	0.19
No. of obs.		1,162	231	18	1,411		782	373	95	1,250
Average price		75.34	126.59	120.66	81.2		64.53	101.74	108.7	77.68
Average maturity	ITM: $X/S \leq 0.97$	26.07	91.84	231.56	35.34	ITM: $X/S \geq 1.03$	27.85	98.87	250.22	62.6
Average moneyness		0.94	0.91	0.92	0.94		1.05	1.08	1.07	1.06
Average IV		0.27	0.31	0.32	0.28		0.21	0.21	0.27	0.22
No. of obs.		62,987	8,199	884			76,958	14,848	1,877	
Average price		10.96	22.15	32.57			9.09	17.77	30.62	
Average maturity	All moneyness	24.65	92.86	243.39		All moneyness	25.59	97.47	234.68	
Average moneyness		1.04	1.07	1.13			0.94	0.89	0.87	
Average IV		0.19	0.2	0.22			0.25	0.26	0.28	

Note: For each option that trades between 2 p.m. and 3 p.m. every day, we identify the corresponding spot index value based on the most recent spot transaction. We obtain the underlying interest rate by interpolating two Eurodollar rates with the closest maturities (time to maturity is annualized using the 365 days convention). We apply the following filters separately to the call and put option data: (1) options that violate the arbitrage bounds are dropped, (2) options below the minimum tick size of 3/8 are dropped, (3) options below six-day maturities are eliminated, and (4) options with implied volatility (IV) higher than 2 are dropped. The IVs are calculated using the IV surface as described in Appendix A. For each moneyness subsample, we report the number of observations and averages for option price, maturity, moneyness, and volatility for calls and puts. OTM is out of the money; ATM is at the money; ITM is in the money.

the variance estimate because of market microstructure effects (bid-ask bounces, price discreteness, or nonsynchronous trading).[3]

We examine the volatility signature plot for RV over a range of sampling frequencies (not shown). We find that the average RV stabilizes at a 15-minute interval and hence choose 15-minute sampling intervals to construct the RV measure to minimize the measurement error bias. The autocorrelation function of RV (not reported) further suggests the presence of long-memory dynamics as inherent in the autoregressive fractionally integrated moving average (ARFIMA) models.

III. Obtaining Volatility Forecasts

Stand-Alone Historical Volatility Forecasts

Model 1: RV. Here we compute volatility for a given day as a 20-day trailing arithmetic average of past daily RVs. RV for a given day is obtained as the sum of squared returns for that day, discretely sampled at Δ-minute intervals. Specifically, RV on day t is given by $RV_t = \sum_{i=1}^{1/\Delta} r_{t+j\Delta,\Delta}^2$, where $r_{t+\Delta,\Delta}$ refers to Δ-minute sampled returns, obtained as $\ln p_{t+\Delta} - \ln p_t$.

Model 2: ARFIMA Model. The ARFIMA model is

$$(1 - L)^d \ln \sigma_t^2 = \omega + \gamma \ln \sigma_{t-1}^2 + u_t.$$

The model is used following the previous evidence that the RV is a highly persistent process with slowly decaying autocorrelations (e.g., Andersen et al. 2003). Previous studies find that RV has no unit root but is fractionally integrated with a slow rate of mean reversion. Because the logarithmic RV series is observed to be homoskedastic and approximately unconditionally Gaussian, we use the ARFIMA (1, d, 0) representation of the logarithmic RV.

Model 3: ARFIMA Model with Leverage Effect (ARFIMA-L). The ARFIMA-L model is

$$(1 - L)^d \ln \sigma_t^2 = \omega + \gamma \ln \sigma_{t-1}^2 + \phi \ln \sigma_{t-1}^2 I(r_{t-1} < 0) + u_t.$$

This model incorporates the leverage component and thereby accounts for the asymmetric response of volatility to positive versus negative news shocks.

[3]Aït-Sahalia, Mykland, and Zhang (2005) show that in the presence of market microstructure noise, realized volatility calculated from ultra high frequency returns only captures the noise instead of return volatility. Hence, the authors show that it is optimal to sample less often than one would otherwise do so.

Model 4: ARFIMA Model with Leverage Effect and Jumps (ARFIMA-LJ). The ARFIMA-LJ model is

$$(1-L)^d \ln \sigma_t^2 = \omega + \gamma \ln \sigma_{t-1}^2 + \phi \ln \sigma_{t-1}^2 I(r_{t-1} < 0) + \beta_j J_{t-1} + u_t,$$

where J refers to the jump measure based on the theory of quadratic variation (see Andersen, Bollerslev, and Diebold 2007)[4] and is obtained as

$$J_t = \max[RV_t - BV_t, 0] \,.$$

BV is the bipower variation measure expressed as

$$BV_t = \mu_1^{-2} \sum_{j=2}^{1/\Delta} \left| r_{t+j\Delta,\Delta}^2 \right| \left| r_{t+(j-1)\Delta,\Delta}^2 \right| \quad \text{with} \quad \mu_1 = \sqrt{2/\pi}.$$

Model 4 incorporates the leverage and jump components. According to the theory of quadratic variation, RV approaches true volatility as the sampling frequency tends to infinity; however, in the presence of jumps, as the sampling frequency increases, RV tends to the true volatility that now has both continuous and jump components. Barndorff-Nielsen and Shephard (2003, 2004, 2006) develop the related theory in a series of studies and show that the jump component can be isolated using a nonparametric method, which involves defining a BV measure and imposing a nonnegative constraint on the difference between RV and BV. The jump factor obtained captures the effect of sudden market news events.

Incorporating the jump component helps reduce the effect of lagged RV on days in which the jump component is active, thereby lowering the persistence of the RV process. Therefore, whereas the ARFIMA model captures the long-memory property of the daily data, the leverage and jump components help explain the negative skewness and possible fat tails in the returns.

All the ARFIMA models are based on daily realized volatilities. Furthermore, the calculation of realized volatilities includes the overnight adjustment (see Blair, Poon, and Taylor, 2001). Specifically, the first return on a given day is calculated as the log difference of the first trade that day and the closing index from the previous trading day. Similarly, the second return on a given day is obtained as the log difference between the first trade and the trade at the 15-minute grid.

Model 5: Glosten, Jagannathan, and Runkle GARCH(1,1) Model (GJR GARCH-L). The Glosten, Jagannathan, and Runkle (GJR) (1993) model is used as a benchmark historical volatility specification.

[4]Alternative tests of jumps proposed in the literature include Aït-Sahalia (2002), Aït-Sahalia and Jacod (2009), Andersen, Bollerslev, and Dobrev (2007), Jiang and Oomen (2008), and Lee and Mykland (2008).

$$r_t = \mu + \varepsilon_t, \quad h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-1}^2 I(\varepsilon_{t-1} < 0) + \beta h_{t-1}$$
$$\varepsilon_t = \sqrt{h_t} z_t, \quad z_t \sim N(0, 1),$$

where r_t is the daily return on day t, h_t is the conditional volatility on day t, and $I(.)$ is an indicator function equal to 1 if the argument is true, and 0 otherwise.

We estimate each model based on a four-year rolling window for the sample period 1996–2002 and generate multistep out-of-sample forecasts on a given day for each option according to the underlying maturity horizon. We examine 1-day-, 5-day-, and 20-day-ahead volatility forecasts corresponding to daily, weekly, and monthly forecasts. Our out-of-sample analysis is conducted on the last three years of option data (i.e., 2000–2002).

IV Forecasts

We next obtain IV using the implied volatility function approach (details in Appendix A). The volatility surface is defined in log form as a polynomial function of maturity and moneyness, and its parameters are obtained by minimizing the difference between the market and Black–Scholes model prices observed during the 2 p.m.–3 p.m. window every trading day (see Christoffersen and Jacobs 2004; Christoffersen 2003, p. 138).

We report the summary statistics for daily volatility forecasts for the competing models for the call option data in Table 2. From Panel A, we observe that, in general, the average volatility forecasts from the RV and ARFIMA models are lower than those from the IV and GARCH-L models for both OTM and ATM options. As Bakshi and Kapadia (2003) argue, the index option IVs can be upward biased on account of the negative volatility risk premium. Though the untabulated results for puts are qualitatively similar, the put IV forecasts are much higher compared to the call IV forecasts because of high put prices (especially OTM puts), possibly from hedging demand. In Panel B, we observe that the volatility forecasts from the IV model have a higher correlation with volatility forecasts from the ARFIMA models than with the GARCH-L model, implying that long-memory intraday volatility measures are more sensitive to market new events compared to low-frequency GARCH models.

Combining Historical Volatility and IV Forecasts

The availability of high-frequency data can help provide superior volatility forecasts to option traders by combining intraday historical and IV forecasts. Whereas historical volatility has clustering effects in daily or weekly data and long-memory properties in intraday data (see Andersen, Bollerslev, Diebold, and Ebens 2001), IV has forward-looking market information implicit in the IV surface. For example, Clemen (1989) and Diebold and Lopez (1996) show that when each candidate

TABLE 2. Descriptive Statistics and Correlations of Daily Volatility Forecasts (S&P 500 Index Call Options: January 2000–December 2002).

	IV	GARCH-L	RV	ARFIMA	ARFIMA-L	ARFIMA-LJ
Panel A. Descriptive Statistics of Volatility Forecasts						
OTM						
Mean	0.21	0.22	0.17	0.18	0.18	0.18
Median	0.20	0.20	0.16	0.17	0.17	0.17
Standard dev.	0.05	0.04	0.04	0.04	0.04	0.04
Maximum	0.43	0.42	0.34	0.32	0.31	0.31
Minimum	0.09	0.14	0.08	0.12	0.12	0.12
ATM						
Mean	0.24	0.22	0.17	0.18	0.18	0.18
Median	0.22	0.20	0.16	0.18	0.18	0.18
Standard dev.	0.06	0.04	0.04	0.04	0.04	0.04
Maximum	0.47	0.42	0.34	0.32	0.31	0.31
Minimum	0.11	0.14	0.08	0.12	0.12	0.12
Panel B. Correlations among One-Day-Ahead Volatility Forecasts						
OTM						
IV	1.00	0.74	0.66	0.81	0.81	0.81
GARCH-L	0.74	1.00	0.66	0.69	0.70	0.70
RV	0.66	0.66	1.00	0.77	0.77	0.77
ARFIMA	0.81	0.69	0.77	1.00	0.99	0.99
ARFIMA-L	0.81	0.70	0.77	0.99	1.00	0.99
ARFIMA-LJ	0.81	0.70	0.77	0.99	0.99	1.00
ATM						
IV	1.00	0.69	0.64	0.77	0.77	0.78
GARCH-L	0.69	1.00	0.65	0.69	0.70	0.70
RV	0.64	0.65	1.00	0.78	0.78	0.78
ARFIMA	0.77	0.69	0.78	1.00	0.99	0.99
ARFIMA-L	0.77	0.70	0.78	0.99	1.00	0.99
ARFIMA-LJ	0.78	0.70	0.78	0.99	0.99	1.00

Note: The sample consists of out-of-the-money (OTM) call options defined as $X/S \geq 1.03$ and at-the-money (ATM) call options defined as $0.97 < X/S < 1.03$ with maturity ranging from 6 to 60 days. For the two subsamples, we calculate out-of-sample daily volatility forecasts for the implied volatility (IV) and five other model specifications described in Section III: generalized autoregressive heteroskedasticity model with leverage (GARCH-L), realized volatility (RV), autoregressive fractionally integrated moving average (ARFIMA), ARFIMA model with leverage effect (ARFIMA-L), and the ARFIMA model with leverage and jump (ARFIMA-LJ). In Panel A we report the corresponding summary statistic, and in Panel B we present correlations for one-day-ahead volatility forecasts from different models.

model is potentially misspecified, and pooling of underlying information sets is not feasible, individual model forecasts can be combined to generate more informative composite forecasts. One way of combining option IV and historical volatility is by employing encompassing regressions, which have been widely used to study the information content of the options market.[5]

[5]For alternative ways of combining data, see Blair, Poon, and Taylor (2001), Donaldson and Kamstra (2005), Koopman, Jungbacker, and Hol (2005), and Engle and Gallo (2006).

We therefore express the encompassing regression with IV as the dependent variable as below.

$$ln\left(\sigma_{t,T}^{IV}\right) = \alpha + \beta_1 ln\left(\sigma_{t,T}^{FIV}\right) + \beta_2 ln\left(\sigma_{t,T}^{FHV}\right) + e_{t,T} \quad (1)$$

We target the IV for a $T-t$ maturity option, conditional on the observed market option price at time t. The variable $\sigma_{t,T}^{FIV}$ is the forecast of IV on day t using the parameter vector for the IV function estimated on day $t-1$, and the variable $\sigma_{t,T}^{FHV}$ is the forecast of historical volatility conditional on day $t-1$ information, obtained by averaging the daily (annualized) volatility forecasts from a given historical volatility model over the life of the option. Note that both independent variables are conditional on information up to day $t-1$. The log volatility specification imposes a nonnegative constraint lowering the effect of outliers.

The specification described in equation (1) is similar to a stochastic volatility specification with the IV being part of the latent volatility process. This specification implies that if option prices reflect volatility information, IV should not only predict future volatility but should be a function of past volatility because past and future volatilities are positively related. Christensen and Prabhala (1998, p. 140) suggest that this specification "provides a constructive (though nontheoretic) means of forecasting future implied volatilities, and hence future option prices, using variables in the market's information set."

By focusing on future IV in the encompassing regressions, we can meaningfully investigate the economic relevance of intraday RV measure for option markets. Because there is a one-to-one mapping between IV and option prices we can also generate option price forecasts for option trading by targeting IV at time $t+1$.

Generating Combination Forecasts

In addition to the six stand-alone volatility measures (i.e., RV, ARFIMA, ARFIMA-L, ARFIMA-LJ, GARCH-L, and IV), we consider the following volatility combinations, that are denoted by a plus (+) sign: IV+GARCH-L, IV+RV, IV+ARFIMA, IV+ARFIMA-L, and IV+ARFIMA-LJ for out-of-sample option pricing, trading, and hedging tests. We adopt the regression combination method of Granger and Ramanathan (1984) to generate rolling volatility forecasts, as described in the following:

- Consider an option (from a specific moneyness category) with a remaining maturity of $T-t$ days on a given day t. Suppose that we are forecasting at time t the IV for the underlying maturity (i.e., $T-t$) and denote it as $\hat{\sigma}_{t,T}^{IV}$.
- Using a given time-series model, we first obtain the average volatility forecast on day t for the $T-t$ ahead period, conditional on the information

set on day $t-k$, where $k = 1$, 5, or 20 days, corresponding to the daily, weekly, or monthly forecasting horizon. We denote this historical volatility forecast as $\sigma^{FHV}_{t,T|\Omega_{t-k}}$.

- Next, we obtain the IV forecast on day t for the $T-t$ day ahead period, using the IV function for period $t-k$, where the forecasting horizon $k = 1$, 5, or 20 days, for a given moneyness/maturity option group (as described in Appendix A). We denote this IV forecast as $\sigma^{FIV}_{t,T|\Omega_{t-k}}$. For example, $\sigma^{FIV}_{t,T|\Omega_{t-1}}$ for a 30-day ATM call option represents a $T-t$ day ahead forecast of IV on day t based on the estimated IV function using all the short-term (i.e., 6- to 60-day) ATM options traded on day $t-1$.
- Finally, we make a combined IV forecast at time t, based on the historical and IV forecasts using the forecasting equation

$$\left(\hat{\sigma}^{IV}_{t,T}\right) = \hat{\delta}_0 + \hat{\delta}_1 \left(\sigma^{FIV}_{t,T|\Omega_{t-k}}\right) + \hat{\delta}_2 \left(\sigma^{FHV}_{t,T|\Omega_{t-k}}\right), \tag{2}$$

where the dependent variable $(\hat{\sigma}^{IV}_{t,T})$ represents the day t combined volatility forecast for the option (whose maturity is $T - t$ days), conditional on information at day $t-k$. The delta coefficients are calibrated from the historical option market IV and spot market volatility data. Specifically, the parameter vector $\hat{\delta}$: $\hat{\delta} = (\hat{\delta}_0, \hat{\delta}_1, \hat{\delta}_2)$ for a given moneyness/maturity group is obtained by estimating the following regression

$$\left(\sigma^{IV}_{t,T}\right) = \delta_0 + \delta_1 \left(\sigma^{FIV}_{t,T|\Omega_{t-k}}\right) + \delta_2 \left(\sigma^{FHV}_{t,T|\Omega_{t-k}}\right) + e_{t,T} \tag{3}$$

over the previous 20-trading-day window of volatility data, over that specific moneyness-maturity group.[6] For each day t in the sample, forecasts of IV are made using a rolling 20-day window estimation.

Econometric Issues

Encompassing regression tests have to address several methodological issues such as possible measurement errors in the instrumental variables, telescoping errors, orthogonal errors in forecasting regressions, and biases in regression tests (for a discussion of these issues, see, e.g., Fleming 1998; Christensen and Prabhala 1998; Christensen, Hansen, and Prabhala 2002; Neely 2009).

Measurement Errors. The independent variable in our encompassing regressions (i.e., the current IV) can contain measurement errors for the following reasons. First, when a specific option pricing model is used, model misspecification can lead to biases in IV. For example, Jiang and Tian (2005) show that implementing

[6]Our results hold when we use longer estimation windows for the volatility data.

a model-free IV using different strike options can yield a more consistent measure. Jiang and Tian (2007) further demonstrate that the performance of the model-free IV approach is contingent on the underlying numerical implementation of the measure. These authors show that the CBOE VIX procedure can lead to systematic bias of IV and can be improved by using a smoothing procedure that minimizes truncation and discretization errors. Second, the choice of particular option contracts used in calculating IV can result in errors. Finally, the options market is less liquid compared to the underlying SPX; thus, market microstructure noise can cause error. Potential nonsynchronous measurement of option and index levels, and noncontinuous prices could lead to error. In general, these errors cause inconsistent least square regression estimates and could lead to spurious regression results in small samples (Fleming 1998).

Because we focus on the most liquid segment of the SPX options market, namely, short-term ATM and OTM calls and puts, we expect to minimize the effect of measurement errors arising for liquidity reasons. In estimating the IV series, we employ the IV surface (as in Christoffersen and Jacobs 2004) involving an end-of-day 2 p.m.–3 p.m. window of option prices. Christoffersen and Jacobs (2004) show that the IV function can be used as a robust benchmark for forecasting purposes, and Dumas, Fleming, and Whaley (1998) point out that the ad hoc or practitioner Black-Scholes model outperforms many of the deterministic volatility models. As in Fleming (1998), using a window of transactions minimizes the effects of bid-ask spreads and noncontinuous option prices; because the window is centered around the stock market close, the errors relating to infrequent trading and timing mismatch are minimized.

Overlapping Errors. Simultaneous trading in contracts with overlapping expirations creates cross-correlations in the time series of forecasting errors. Because daily observations on IV involve sequential forecasts for overlapping periods, there is an overlapping period between current IV and future IV. This leads to serial dependence in the time series of forecast errors.

Previous studies deal with the problem by aggregating and excluding data to create nonoverlapping observations. Unfortunately, such procedures can severely reduce the power of statistical tests, as implementing nonoverlapping sample estimation requires long time series of historical option and spot data to ensure adequate sample size. For example, Christensen and Prabhala (1998) use more than 11 years of monthly data, and Christensen and Nielsen (2006) use more 16 years of monthly data for forecasting RV.

Because our objective is to forecast IVs from the perspective of option market participants, we expect the most recent history of option market information, and hence most recent time series, to be more informative. Hence, we maximize the amount of information obtained from the data by adopting an estimation procedure that uses the entire data sample and corrects for the time dependence directly. We estimate the encompassing regressions using the Hansen (1982) generalized method of moments (GMM) approach and employ a heteroskedasticity-consistent variance–covariance matrix to minimize potential telescoping overlapping data problems (see Fleming 1998; Neely 2009).

Overall, our approach minimizes the problems relating to misspecification and estimation errors. Any remaining measurement errors could lead to a downward bias in the standard errors in the encompassing regression, which we attempt to minimize by using the Hansen GMM approach. Because the estimated coefficients are unbiased, they are still suitable for forming point volatility forecasts for the option pricing, trading, and hedging exercises.

IV. Encompassing Regression Results

Here we employ the encompassing regression (1) to understand whether IV and hence option prices are predictable. If the information content in the volatility forecasting model is subsumed in the IV forecast, the testable null hypothesis is $\beta_2 = 0$. If the IV forecast is an unbiased estimate of future IV, the testable null hypothesis is $\beta_1 = 1$ and $\beta_2 = 0$.

The results for the unbiasedness tests for S&P 500 (OTM and ATM) call options are provided in Table 3. We report only call option results for brevity. We also report Hansen's (1982) t-statistics for individual coefficients and the F-statistics for the joint null hypothesis: $\alpha = 0$, $\beta_1 = 1$, $\beta_2 = 0$, where the corresponding p-values are robust to overlapping observations. In addition to the adjusted R^2 we report the nonnested Diebold–Mariano (DM) test statistic and corresponding p-values for each encompassing regression evaluated against the IV benchmark.

From the first rows of Panels A and B in Table 3, we observe that IV has significant information content for the next-period IV. The coefficient for IV (and its t-statistic) declines, and the coefficient for historical volatility (and its t-statistic) increases monotonically from RV to ARFIMA-LJ specifications. The adjusted R^2 improves when IV is augmented by ARFIMA forecasts. The ARFIMA volatility measure has a first-order effect in terms of the adjusted R^2, and leverage and the jump terms only have second-order effects. Based on the coefficient (and t-statistic) of the historical volatility and the adjusted R^2 measure, ARIMA-LJ forecasts seem to have the highest additional information content.

Second, the Diebold–Mariano test statistics are significant, indicating that IV does not subsume the information contained in the volatility forecast measures for all the other models. The F-statistic is significant for all of the models and indicates that the unbiasedness hypothesis is rejected for OTM and ATM options. This implies that IV alone at time t cannot forecast IV for the next period and RV has incremental information content in predicting IV. The results hold for both OTM and ATM options, and suggest that current IV is a biased predictor of the next-period IV. The results for puts are qualitatively similar, although IV has greater

TABLE 3. Encompassing Regression Tests with Actual Implied Volatility as the Dependent Variable (S&P 500 Index Call Options: January 2000–December 2002).

Volatility Forecast Model	α	β_1	β_2	Adjusted R^2	Diebold-Mariano Test Statistic	F-test
Panel A. OTM: $X/S \geq 1.03$ ($N = 724$)						
IV	−0.217	0.857		0.778		38.15
	(19.099)	(120.979)				(.0001)
GARCH-L	−0.16	0.805	0.091	0.781	9.25	29.11
	(9.606)	(77.593)	(8.536)		(.0001)	(.0001)
RV	−0.116	0.786	0.120	0.785	19.63	34.19
	(7.445)	(80.889)	(13.895)		(.00)	(.00)
ARFIMA	−0.032	0.667	0.282	0.798	13.6	52.97
	(2.358)	(77.748)	(35.676)		(.0001)	(.0001)
ARFIMA-L	−0.016	0.655	0.302	0.801	13.34	56.46
	(1.24)	(80.716)	(40.334)		(.0001)	(.0001)
ARFIMA-LJ	−0.013	0.653	0.306	0.801	13.28	57.15
	(1.034)	(81.49)	(41.434)		(.0001)	(.0001)
Panel B. ATM: $0.97 < X/S < 1.03$ ($N = 713$)						
IV	−0.178	0.879		0.796		26.48
	(14.977)	(110.224)				(.0001)
GARCH-L	−0.116	0.837	0.079	0.798	1.23	2.72
	(6.559)	(70.814)	(7.055)		(.0001)	(.0001)
RV	−0.044	0.806	0.135	0.803	24.83	28.07
	(2.811)	(77.425)	(15.672)		(.00)	(.00)
ARFIMA	0.044	0.709	0.273	0.816	22.82	46.47
	(3.272)	(79.052)	(35.547)		(.0001)	(.0001)
ARFIMA-L	0.057	0.702	0.286	0.817	22.81	48.59
	(4.219)	(78.180)	(37.167)		(.0001)	(.0001)
ARFIMA-LJ	0.059	0.700	0.289	0.818	22.86	49.15
	(4.409)	(77.907)	(37.599)		(.0001)	(.0001)

Note: The sample consists of out-of-the-money (OTM) call options defined as $X/S \geq 1.03$ and at-the-money (ATM) call options defined as $0.97 < X/S < 1.03$ with maturity ranging from 6 to 60 days. For each subsample, historical volatility (HV) forecasts for five models are obtained as described in Section III. To obtain implied volatility (IV), we first apply option filters as described in Table 1. Then for each subsample, we identify the last option trading during the 2 p.m.–3 p.m. window each day and use the parameter vector for the IV function estimated on day *t*–1 to forecast *IV* for next day, that is, day *t,* where *t* represents each day in the sample (details of IV estimation are described in Appendix A). The encompassing regression is defined as:

$$\ln\left(\sigma_{t,T}^{IV}\right) = \alpha + \beta_1 \ln\left(\sigma_{t,T}^{FIV}\right) + \beta_2 \ln\left(\sigma_{t,T}^{FHV}\right) + e_{t,T}.$$

Results are reported for the joint null hypothesis: $\alpha = 0,\ \beta_1 = 1,\ \beta_2 = 0$. *FHV* and *FIV* refer to the volatility forecasts generated, respectively, from the historical volatility and implied volatility function approaches, conditional in the information set at time *t*–1. The HV forecasts are from the following models and are described in Section III: generalized autoregressive heteroskedasticity model with leverage (GARCH-L), realized volatility (RV), autoregressive fractionally integrated moving average (ARFIMA), ARFIMA model with leverage effect (ARFIMA-L), and the ARFIMA model with leverage and jump (ARFIMA-LJ). The volatility forecasts are for the maturity date, T using information from day t. The dependent variable $(\sigma_{t,T}^{IV})$ represents the actual *IV* calculated from the option on day t. The independent variable $(\sigma_{t,T}^{FIV})$ is the forecast of *IV* on day t using the parameter vector for the *IV* function estimated on day *t*–1. The independent variable $(\sigma_{t,T}^{FHV})$ is the forecast of *HV* obtained by averaging the daily (annualized) volatility forecasts from a given historical volatility model over the life of the option. The encompassing regressions are estimated using Hansen's generalized method of moments procedure and is robust to overlapping observations. For each regression, we report Hansen's (1982) *t*-statistics for individual coefficients (in parentheses) and the adjusted R^2. The F-statistic (p-values in parentheses) tests the joint null hypothesis: that $\alpha = 0, \beta_1 = 1, \beta_2 = 0$. We also report the Diebold–Mariano test statistic for each encompassing regression (p-values in parentheses) against the IV benchmark. We report the results for OTM call options in Panel A and for ATM call options in Panel B.

information content for puts relative to calls. In untabulated results, we also conduct the encompassing regression tests correcting for possible error-in-variable bias using the instrumental variable approach as in Christensen and Nielsen (2006) and find that the results hold.

Overall, we find that volatility forecasting for option pricing should employ information in both IV and RV; in particular, the IV and ARIMA-LJ combination has the highest explanatory power and F-test value.

V. Out-of-Sample Option Pricing Tests

We next use the Black–Scholes model to convert our out-of-sample IV forecasts into predicted options prices. The Black–Scholes model is merely used as a tool to provide one-to-one mapping between option prices and IV forecasts and by no means implies that it is the true underlying model (see Jiang and Tian 2005; Christensen and Nielsen 2006 for similar arguments). We employ four metrics to compare the relative pricing performance (see Dumas, Fleming, and Whaley 1998): mean squared error (MSE), the mean of squared price differences; mean absolute error (MAE), the mean of absolute price differences; absolute log error (ALE), the mean of absolute log price differences; and fequency (FREQ), the ratio of the total number of trading days on which a particular model has a lower daily MSE than the benchmark IV model. Following Christoffersen and Jacobs (2004), we expect our inferences based on the MSE to be least biased as the same MSE criterion is employed in both IV function estimation and out-of-sample option price evaluation.

Table 4 presents the out-of-sample pricing results. Panel A reports the summary statistics of one-day-ahead option price forecasts based on competing volatility models. For both OTM and ATM options, prices from the combined volatility models (from IV and intraday forecasts) are closest (in terms of means and medians) to those observed in the market. Similarly, the option prices generated by the combination models have standard deviations nearest to those observed in the market.

Panel B presents the comparative metrics for pricing errors. In general, OTM call options have lower pricing errors than ATM calls. Prices based on combined forecasts have the lowest pricing errors compared to stand-alone models. For OTM call options, the combination of the IV and ARFIMA forecasts outperform all other models based on the MSE, MAE, ALE, and FREQ metrics. The IV+ ARFIMA-LJ model combination has the best performance based on all four metrics. This implies that the IV+ARFIMA-LJ combination not only has the lowest MSE but also, on most days, has lower MSE than the benchmark IV model. For ATM call options, the IV and daily GARCH combinations have the lowest MSE and MAE, followed by the IV+ARFIMA-LJ combination. The FREQ measure indicates that IV+ARFIMA combinations have the best performance.

TABLE 4. Comparison of Alternative Volatility Models Based on Daily, Weekly, and Monthly Price Forecasts (S&P 500 Index Call Options: January 2000–December 2002.

Panel A. Summary Statistics for One-Day-Ahead Call Option Price Forecasts

	Market Price	IV	GARCH-L	RV	ARFIMA	ARFIMA-L	ARFIMA-LJ	IV+ GARCH-L	IV+ RV	IV+ ARFIMA	IV+ ARFIMA-L	IV+ ARFIMA-LJ
OTM												
Mean	5.97	6.07	6.59	3.62	4.06	4.06	4.07	5.99	6.00	5.98	5.96	5.96
Median	4.30	4.65	5.23	2.19	2.95	3.00	3.00	4.63	4.56	4.71	4.65	4.64
Standard dev.	5.12	5.13	5.39	4.01	3.71	3.71	3.71	4.86	4.86	4.85	4.84	4.84
Maximum	30.50	30.51	44.47	30.32	21.14	21.40	21.48	27.55	25.51	27.13	26.84	26.76
Minimum	0.40	0.04	0.02	0.01	0.02	0.02	0.02	0.22	0.20	0.21	0.20	0.20
ATM												
Mean	23.39	23.43	21.31	16.29	17.40	17.41	17.42	23.35	23.51	23.58	23.56	23.56
Median	22.00	21.37	19.67	14.21	15.69	15.71	15.78	22.02	22.20	22.25	22.10	22.07
Standard dev.	12.65	13.00	11.62	10.47	10.37	10.34	10.34	12.03	12.12	12.23	12.21	12.21
Maximum	78.00	78.90	91.50	72.98	68.77	68.70	68.53	74.11	73.26	74.24	74.19	74.20
Minimum	2.50	2.59	1.44	0.90	1.18	1.16	1.17	2.99	3.43	3.23	3.27	3.27

(Continued)

TABLE 4. Continued.

Panel B. Pricing Performance for Alternate Volatility Models Based on Daily Forecasts

	IV	GARCH-L	RV	ARFIMA	ARFIMA-L	ARFIMA-LJ	IV+ GARCH-L	IV+ RV	IV+ ARFIMA	IV+ ARFIMA-L	IV+ ARFIMA-LJ
OTM											
MSE	2.693	8.829	13.623	9.004	8.867	8.813	2.521	2.571	2.379	2.383	2.379
MAE	1.004	2.023	2.706	2.061	2.053	2.048	1.018	1.029	0.985	0.982	0.981
ALE	0.238	0.439	0.895	0.571	0.566	0.563	0.216	0.217	0.21	0.211	0.211
FREQ	0	0.277	0.169	0.228	0.217	0.22	0.5	0.492	0.505	0.515	0.517
ATM											
MSE	9.65	30.93	75.30	53.93	53.68	53.47	6.79	7.68	7.58	7.51	7.47
MAE	2.17	3.97	7.27	6.04	6.03	6.02	1.78	1.88	1.86	1.85	1.85
ALE	0.10	0.20	0.43	0.33	0.33	0.33	0.09	0.09	0.09	0.09	0.09
FREQ	0.00	0.32	0.12	0.14	0.14	0.14	0.58	0.58	0.59	0.59	0.59

Note: The sample consists of out-of-the-money (OTM) call options defined as $X/S \geq 1.03$ and at-the-money (ATM) call options defined as $0.97 < X/S < 1.03$ with maturity ranging from 6 to 60 days. In Panel A, we report summary statistics for one-day-ahead call option prices based on volatility forecasts from different models. We also consider combination forecasts where we combine implied volatility (IV) with volatilities from alternate models: generalized autoregressive heteroskedasticity model with leverage (GARCH-L), realized volatility (RV), autoregressive fractionally integrated moving average (ARFIMA), ARFIMA model with leverage effect (ARFIMA-L), and the ARFIMA model with leverage and jump (ARFIMA-LJ). The combination forecasts are obtained using the encompassing regression (3), that is,

$$\left(\hat{\sigma}_{t,T}^{IV}\right) = \hat{\delta}_0 + \hat{\delta}_1 \left(\sigma_{t,T|\Omega_{t-k}.}^{FIV}\right) + \hat{\delta}_2 \left(\sigma_{t,T|\Omega_{t-k}.}^{FHV}\right),$$

where the delta coefficients are calibrated to the previous 20 trading days of data, as described in Section III. In Panel B, we report out-of sample performance metrics for one-day-ahead call option price forecasts. MSE is the mean of squared price differences; MAE is the mean of absolute price differences; ALE is the mean of absolute log price differences; FREQ is ratio of total number of days on which a particular model has a lower daily MSE than the benchmark, that is, the IV model.

In untabulated results, we find that for ATM puts, the combination of IV+ARFIMA models outperforms all others on all metrics. For OTM puts, IV has the most information content as OTM puts are used for crash insurance, and their prices are hard to justify solely based on historical volatilities. We also find qualitatively similar results for call and put options prices based on weekly and monthly volatility forecasts, with forecast errors monotonically increasing at these intervals.

Overall, the out-of-sample pricing results indicate that the combination of IV and ARFIMA models has the lowest pricing errors on average and on most days. The pricing results are supportive of combination forecasts and are consistent with the encompassing regression test findings that volatility forecasting for option pricing should incorporate the information in both IV and RV measures.

VI. Volatility Timing and Dynamic Trading in Straddles

We now use competing out-of-sample volatility forecasts to trade in short-term ATM delta-neutral straddles. Straddle trading represents exposure to both call and put options on an underlying with identical strike price and maturity.[7] Straddle positions help traders exploit significant differences between their private and market estimates of spot volatility. We provide the details of the straddle trading strategy in Appendix B. As in Noh, Engle, and Kane (1994) and Jha and Kalimipalli (Forthcoming), we form delta-neutral ATM straddles each day, price them based on alternate volatility forecasts, and then buy (or sell) them on each day, depending on whether they are under- (or over-) priced. When the straddle is purchased (sold), funds are borrowed (invested) at a risk-free rate.

We consider both the underlying bid–ask spread and commission costs. We trade only when the absolute percentage price difference between the model and market price (as a proportion of market price) is expected to exceed the straddle bid–ask spread of 5.70% (where 5.70% is obtained as the average of the bid–ask spread of 5.5% for calls and 5.9% for puts; see Santa-Clara and Saretto 2005, Table 7). We apply the commission costs of 0.5% toward the transaction costs. The 0.5% commission cost corresponds to a trader with a $48,000 investment who pays a commission amounting to $120 plus 0.25% of the dollar amount, based on a standard commission schedule (see Hull 2000, p. 160).

For daily trading, we trade every day t conditional on the day $t-1$ information set and hold that position until the next trading day, when the position is reversed. Similarly, for weekly (monthly) trading, we trade every Wednesday (fifth Wednesday) conditional on the previous week's (month's) information set. We compare the profits net of transaction costs.

---

[7] Other straddle variants with different strike and maturity horizons may also be used but are not considered here.

Besides the summary statistics for profits from straddle trading, we present the significance levels based on p-values for the mean obtained using the modified t-statistic $t_j = (RET + \frac{\sigma S}{6n} + \frac{RET^2 S}{3\sigma})(\sigma^2 n)^{-0.5}$, where RET is the mean return, σ is the standard deviation, S is the skewness, and n is the number of observations. The modified t-statistic accounts for the asymmetric distributional property of the returns from option trading strategies (see Bollen and Whaley 2004). We also report the ratio of positive return versus negative return days. Finally, we present the Sharpe ratio and the corresponding significance levels, robust to non-iid (independent and identically distributed) returns that include autocorrelation in returns and are estimated using a GMM framework following Lo (2002).[8]

Table 5 reports the performance of delta-neutral ATM straddles for daily, weekly, and monthly return horizons, considering only the commission costs (0.5%) and ignoring the bid–ask spreads. We notice that the IV and IV-ARFIMA combination models in general have higher average returns, much higher volatility, (positive) skewness, kurtosis, and higher positive return days compared to the other models. The Sharpe ratios are positive and significant for trading based on these models and improve as the trading horizon increases. The stand-alone IV (GARCH) has the best performance at the daily (monthly) horizon with a Sharpe ratio of 0.120 (0.511). The IV+ARFIMA-LJ combination generates the best Sharpe ratio (0.283) only at the weekly interval.

Table 6 reports the straddle results with both bid–ask spreads and commission costs. The results in Panel A show that all the stand-alone historical volatility models now have significant negative mean returns and negative Sharpe ratios. For example, the Sharpe ratio from the straddle trades based on IV forecasts is 0.03, whereas the corresponding values for ARFIMA and ARFIMA-L forecasts are, respectively, −0.266, and −0.264, significant at the 1% level. For daily returns, the stand-alone IV model performs best with minimal losses. In comparison, the daily market Sharpe ratio was −0.0489 for exposure to the S&P 500 market portfolio during the same period, that is, 2000–2002. The returns from neither IV nor IV combinations are significantly different from zero. This implies that once both bid–ask and commission costs are factored, no significant excess returns can be earned from trading straddles using any of the volatility measures.

From Panels B and C in Table 6, we find that the Sharpe ratios improve and turn positive (though not significant) as the trading horizon increases. The stand-alone IV is once again the best model for weekly and monthly trades in terms of the Sharpe ratios.

In summary, although the IV+ARFIMA-LJ combination is most profitable for weekly straddle trades in the presence of only commission costs, it loses its vi-

[8] Lo (2002) shows that the standard error of the Sharpe ratio for a sample size T is asymptotically given as $\sqrt{\hat{V}_{GMM}/T}$, where $\hat{V}_{GMM}$ is the variance of Sharpe ratio that can be estimated using GMM.

TABLE 5. Daily, Weekly, and Monthly Trading in ATM Delta-Neutral Short-Term Straddles: Zero Bid–Ask Costs (January 2000–December 2002).

	IV	GARCH-L	RV	ARFIMA	ARFIMA-L	ARFIMA-LJ	IV+ GARCH-L	IV+RV	IV + ARFIMA	IV+ ARFIMA-L	IV+ ARFIMA-LJ
Panel A. Percentage Daily Returns from Straddle Trades											
Maximum	98.079	87.939	88.069	88.069	88.069	88.069	97.306	98.079	98.079	98.079	97.306
Minimum	−57.784	−57.682	−57.682	−57.682	−57.682	−57.682	−55.407	−55.407	−55.407	−55.407	−55.407
Mean	2.64***	0.219	−0.087	−0.045	−0.045	0.1	2.209**	1.924**	2.352**	2.228**	2.265**
Skewness	1.356	0.488	0.512	0.516	0.516	0.643	1.097	1.171	1.282	1.293	1.012
Kurtosis	8.019	6.532	5.345	5.346	5.346	5.278	7.697	7.817	7.926	7.950	7.609
Positive vs. negative returns	1.164	1.013	1.448	1.461	1.461	1.587	1.194	1.124	1.169	1.139	1.210
Sharpe ratio	0.12***	0.01	−0.005	−0.003	−0.003	0.004	0.104**	0.089*	0.109**	0.103**	0.107**
Panel B. Percentage Weekly Returns from Straddle Trades											
Maximum	99.649	92.899	92.899	92.899	92.899	76.027	99.649	99.649	99.649	99.649	99.649
Minimum	−57.323	−59.921	−59.921	−59.921	−59.921	−59.921	−57.323	−57.323	−57.323	−57.323	−46.838
Mean	8.665***	5.247*	3.546	3.546	3.546	3.153	7.866**	8.66***	8.513***	8.513***	9.903***
Skewness	0.684	0.432	0.153	0.153	0.153	0.077	0.685	0.699	0.708	0.708	0.667
Kurtosis	2.594	2.614	2.626	2.626	2.626	2.401	2.657	2.543	2.545	2.545	2.533
Positive vs. negative returns	0.908	1.033	1.404	1.404	1.404	1.268	0.824	0.879	0.851	0.851	0.984
Sharpe ratio	0.243***	0.158***	0.111***	0.111***	0.111***	0.098***	0.222***	0.242***	0.238***	0.238***	0.283***
Panel C. Percentage Monthly Returns from Straddle Trades											
Maximum	81.994	81.994	81.994	81.994	81.994	58.103	81.994	81.994	81.994	81.994	81.994
Minimum	−32.823	−24.277	−44.612	−44.612	−44.612	−46.175	−32.823	−32.823	−32.823	−32.823	−32.823
Mean	13.106***	14.218***	4.031	4.031	4.031	−1.611	11.976**	10.179**	10.179**	10.179**	11.494***
Skewness	0.758	0.786	0.775	0.775	0.775	0.136	0.795	0.865	0.865	0.865	0.813
Kurtosis	3.153	2.912	3.517	3.517	3.517	2.671	3.125	3.126	3.126	3.126	3.128
Positive vs. negative returns	2.000	2.000	1.143	1.143	1.143	1.143	1.727	1.308	1.308	1.308	1.727
Sharpe ratio	0.471***	0.511***	0.146***	0.146***	0.146***	−0.067***	0.424***	0.353***	0.353***	0.353***	0.404***

Note: The sample consists of at-the-money (ATM) call (put) options defined as $X/S \geq 1.03$ ($X/S \leq 0.97$) and out-of-the-money (OTM) options defined as $0.97 < X/S < 1.03$ with maturity ranging from 6 to 60 days. Daily straddle trading is based on 1-day-ahead volatility forecasts from competing models: implied volatility (IV), generalized autoregressive heteroskedasticity model with leverage (GARCH-L), realized volatility (RV), autoregressive fractionally integrated moving average (ARFIMA), ARFIMA model with leverage effect (ARFIMA-L), and the ARFIMA model with leverage and jump (ARFIMA-LJ). Weekly (monthly) straddle trading is based on 5-day- (20-day-) ahead volatility forecasts from competing models. The average maturity of the options used in the analysis was 20.40 days and the average moneyness was 1.005. We present the summary statistics from profits from option trading, in addition to ratios of positive returns versus negative return days and the Sharpe ratios. Significance levels are based on robust p-values for means (calculated using Johnson's non-normal t-statistics) and Sharpe ratios (Lo 2002). Only commission costs of 0.5% are considered, and bid–ask costs are ignored. Details of the straddle trading are described in Appendix B.

***Significant at the 1% level.

**Significant at the 5% level.

*Significant at the 10% level.

TABLE 6. Daily, Weekly, and Monthly Trading in ATM Delta-Neutral Short-Term Straddles: Bid–Ask Costs as in Santa-Clara and Saretto (2005) (January 2000–December 2002).

	IV	GARCH-L	RV	ARFIMA	ARFIMA-L	ARFIMA-LJ	IV+ GARCH-L	IV+RV	IV+ ARFIMA	IV+ ARFIMA-L	IV+ ARFIMA-LJ
Panel A. Percentage Daily Returns from Straddle Trades											
Maximum	97.903	82.239	82.369	82.369	82.369	82.369	97.903	97.903	97.903	97.903	97.903
Minimum	−61.107	−63.382	−63.382	−63.382	−63.382	−63.382	−61.107	−61.107	−61.107	−61.107	−61.107
Mean	0.749	−4.119***	−5.586***	−5.637***	−5.584***	−5.258***	−0.335	−0.47	−0.011	−0.131	−0.83
Standard deviation	25.636	21.239	20.935	21.199	21.199	20.539	22.899	23.327	22.638	22.731	21.707
Skewness	1.463	0.407	−0.508	−0.531	−0.538	−0.613	1.59	1.653	1.682	1.68	1.207
Kurtosis	6.969	6.302	5.193	5.098	5.114	5.207	7.633	7.983	8.279	8.219	8.102
Positive vs. negative returns	0.598	0.554	0.608	0.621	0.624	0.654	0.545	0.555	0.615	0.608	0.62
Sharpe ratio	0.029	−0.194***	−0.267***	−0.266***	−0.264***	−0.257***	−0.015	−0.021	−0.001	−0.006	−0.039
Panel B. Percentage Weekly Returns from Straddle Trades											
Maximum	91.606	87.199	70.327	70.327	70.327	70.327	78.426	91.606	91.606	91.606	87.199
Minimum	−43.779	−65.621	−65.621	−65.621	−60.966	−63.689	−43.779	−49.464	−49.464	−49.464	−43.779
Mean	5.464	−0.347	−2.862	−2.689	−2.127	−1.239	0.384	3.236	3.787	2.619	3.37
Standard deviation	36.279	33.836	31.577	31.541	31.109	32.223	33.892	35.785	36.073	35.703	34.204
Skewness	0.792	0.401	0.036	0.021	0.043	0.076	0.843	0.807	0.751	0.821	0.832
Kurtosis	2.522	2.501	2.313	2.323	2.322	2.347	2.546	2.547	2.317	2.462	2.62
Positive vs. negative returns	0.781	0.804	1.018	1.055	1.074	1.017	0.676	0.698	0.683	0.643	0.8
Sharpe ratio	0.15	−0.011	−0.091	−0.086	−0.069	−0.039	0.011	0.09	0.105	0.073	0.098
Panel C. Percentage Monthly Returns from Straddle Trades											
Maximum	76.294	76.294	76.294	76.294	76.294	52.403	76.294	76.294	76.294	76.294	76.294
Minimum	−38.523	−27.683	−50.312	−50.312	−50.312	−51.875	−19.469	−27.683	−26.582	−26.582	−38.523
Mean	11.922	9.054	−1.532	−1.532	−1.532	−6.065	8.524	6.398	7.959	7.959	5.794
Standard deviation	31.995	30.162	27.723	27.723	27.723	24.143	26.434	30.076	28.945	28.945	28.409
Skewness	0.672	0.874	0.767	0.767	0.767	0.07	1.341	0.856	0.958	0.958	0.813
Kurtosis	2.398	2.72	3.698	3.698	3.698	2.685	3.803	2.567	2.904	2.904	3.128
Positive vs. negative returns	1.429	1.333	0.857	0.857	0.857	0.75	1.333	0.778	1	1	1.143
Sharpe ratio	0.372	0.3	−0.056	−0.056	−0.056	−0.252	0.322	0.212	0.275	0.275	0.204

Note: The sample consists of at-the-money (ATM) call (put) options defined as $X/S \geq 1.03$ ($X/S \leq 0.97$) and out-of-the-money (OTM) options defined as $0.97 < X/S < 1.03$ with maturity ranging from 6 to 60 days. Daily straddle trading is based on 1-day-ahead volatility forecasts from competing models: implied volatility (IV), generalized autoregressive heteroskedasticity model with leverage (GARCH-L), realized volatility (RV), autoregressive fractionally integrated moving average (ARFIMA), ARFIMA model with leverage effect (ARFIMA-L), and the ARFIMA model with leverage and jump (ARFIMA-LJ). Weekly (monthly) straddle trading is based on 5-day- (20-day-) ahead volatility forecasts from competing models. The average maturity of the options used in the analysis was 20.40 days and the average moneyness was 1.005. We present the summary statistics from profits from option trading, in addition to ratios of positive returns versus negative return days and the Sharpe ratios. Significance levels are based on robust p-values for means (calculated using Johnson's non-normal t-statistics) and Sharpe ratios (Lo 2002). In addition to commission costs of 0.5%, we consider bid–ask costs of 5.7% based on Santa-Clara and Saretto (2005). Details of the straddle trading are described in Appendix B.

***Significant at the 1% level.

**Significant at the 5% level.

*Significant at the 10% level.

ability once the bid–ask costs are included. The stand-alone IVs, however, seem to generate the best overall Sharpe ratios. Our tests indicate that the statistical superiority of intraday volatility combination forecasts demonstrated in the encompassing regressions and out-of-sample pricing errors may not be reflected in the option trading in the presence of transaction costs. Our results are consistent with Santa-Clara and Saretto (2005), who show that high returns in options strategies documented in studies such as Coval and Shumway (2001) may not translate into profitable strategies after considering trading costs.[9]

VII. Risk Exposure for Option Writers

Finally, we examine the performance of naked and delta-hedged option positions held by financial institutions. Naked calls (puts) involve short positions on calls (puts), whereas delta-hedged calls (puts) involve shorting calls (puts) and delta hedging using the underlying. Short call positions are delta-hedged by buying (short-selling) the underlying when the spot price goes up (down). Similarly, short put positions are delta-hedged by short-selling (buying) the underlying when the spot price goes down (up). As in Green and Figlewski (1999), we further assume that (1) options constitute risky assets in the overall diversified portfolio of financial institutions, and (2) financial institutions use the Black–Scholes model to price and delta hedge based on competing volatility models. We implement the strategies for each moneyness subgroup at daily, weekly (5-day, i.e., every Wednesday) or monthly (20-day, i.e., every fifth Wednesday) trading horizons, using the corresponding horizon volatility forecasts. For weekly (monthly) horizons, we exclude options with maturity less than 5 (20) trading days.[10] All options are sold at model prices (and delta-hedged) based on the competing conditional volatility forecasts.

As in Bollen and Whaley (2004), the abnormal returns to delta hedging for calls over the life of the option are computed as:

$$\frac{\Delta_0\left(S_T+\sum_{t=0}^{T} D_t e^{r(T-t)}-S_0 e^{rT}\right)-\left(\mathrm{Max}\,(0, S_T-K)-c_0 e^{rT}\right)+\sum_{t=0}^{T-1}\Delta_t\left(S_{t+1}+D_t-S_t\right)e^{r(T-t)}}{\Delta_0 S_0-c_0}, \tag{4}$$

[9]Santa-Clara and Saretto (2005) note that margin requirements imposed by option exchanges on short-sale positions can further affect the implementation and profitability of option strategies, in addition to the effect of transaction costs (see Goyal and Saretto Forthcoming for similar arguments).

[10]Following Bakshi, Cao, and Chen (1997), we use the S&P 500 spot index rather than the S&P 500 futures for our delta-hedging exercises. As Bakshi, Cao, and Chen (p. 2033) argue, the spot index and near maturity futures have a near-perfect correlation, and hence the hedging results are likely to be unaffected. Bollen and Whaley (2004) find that use of index rather than futures for delta-hedged positions leaves the average abnormal returns unchanged (p. 750).

where Δ_t is the delta value on day t computed from updated volatility estimates, S_T is the closing value of the index on day t, D_t is the dividend at time t, c_t is the call option value on day t, K is the strike price, r is the risk-free rate, and subscripts 0 and T refer to the opening and closing days. The analogous formula is used for puts. The first term refers to the income from holding the underlying asset until maturity, net of its financing costs. The second term is the gain or loss in the financial position, and the third term represents cash inflows from marked to market adjustments on a daily basis. The denominator is the cost of the option position at inception.

Table 7 reports the Sharpe ratios along with their robust p-values (following Lo 2002) for daily, weekly, and monthly return horizons for OTM and ATM options. In Table 7, only commission costs of 0.5% are employed, and the implied bid–ask costs are ignored. We notice that naked short-call positions generate significant Sharpe ratios at all trading horizons, particularly for OTM calls. For ATM calls, IV+GARCH combination (IV+ARFIMA) models outperform others at daily (weekly) intervals with a Sharpe ratio of 0.281 (0.422). The IV+RV combinations deliver the highest Sharpe ratios at monthly horizons. Delta-hedged call positions have much lower Sharpe ratios compared to naked calls. In summary, although for naked calls, IV+ARFIMA or IV+RV combinations generate marginally significant Sharpe ratios at weekly and monthly horizons, there is no clear winner for delta-hedged calls.

Next, we examine the profitability of short-put positions. Naked puts yield lower Sharpe ratios compared to naked calls, as can be expected in the bear market phase of 2000–2002. IV+ARFIMA-L combinations deliver the highest Sharpe ratios for naked daily and weekly OTM puts (0.133 and 0.569, respectively) and delta-hedged monthly puts (0.342 and 0.538, respectively, for OTM and ATM puts).

We next consider the implications of bid–ask costs. The returns for calls (puts) are adjusted for half the bid-ask spread of 5.5% (5.9%), based on the estimates in Santa-Clara and Saretto (2005), and a commission cost of 0.5%. Only one-way bid–ask spreads are considered because we are assuming that the FI holds the call or put option position to maturity. Table 8 reports the Sharpe ratios. The intraday combination models continue to be profitable for naked calls (at weekly and monthly horizons) and puts (at daily and weekly horizons), and results are similar to Table 7. All of the delta-hedged strategies are, however, unprofitable.

Overall, we find that shorting calls or puts based on IV+ARFIMA-L volatility forecasts can yield significant Sharpe ratios at selective horizons, and the results are robust to transaction costs. Though delta-hedged monthly puts can be profitable using IV+ARFIMA-L combinations, the profits disappear once bid–ask costs are factored in. No significant Sharpe ratios are found for delta-hedged call positions using any of the volatility models. As in the case of straddle trading, our results indicate that the usefulness of intraday volatility measures for delta-hedged option positions are dampened in the presence of transaction costs.

TABLE 7. Sharpe Ratios for Naked and Delta-Hedged Option Strategies: Zero Bid–Ask Costs (S&P 500 Index Options: January 2000–December 2002).

		IV	GARCH-L	RV	ARFIMA	ARFIMA-L	ARFIMA-LJ	IV+GARCH-L	IV+RV	IV+ARFIMA	IV+ARFIMA-L	IV+ARFIMA-LJ
OTM: Naked calls	Daily	0.493***	0.552***	0.218***	0.209***	0.197***	0.197***	0.468***	0.46***	0.45***	0.446***	0.444***
	Weekly	1.31***	1.392***	1.102***	1.148***	1.146***	1.148***	1.298***	1.287***	1.294***	1.293***	1.293***
	Monthly	2.965***	3.503***	2.47***	2.833***	2.875***	2.882***	3.729***	3.756***	3.738***	3.741***	3.741***
ATM: Naked calls	Daily	0.27***	0.19***	0.006	0.059*	0.056*	0.056*	0.281***	0.267***	0.263***	0.267***	0.267***
	Weekly	0.417***	0.386***	0.101	0.174	0.178	0.178	0.324**	0.417***	0.42***	0.422***	0.422***
	Monthly	1.123***	1.138***	0.659**	0.769***	0.773***	0.776***	1.229***	1.275***	1.263***	1.266***	1.267***
OTM: Delta-hedged calls	Daily	0.044	0.113	−0.056	−0.019	−0.018	0.106	0.073	0.034	0.067	0.068	0.017
	Weekly	0.095	0.672	0.151	0.435	0.445	−0.034	0.135	0.265	0.618	0.646*	−0.062
	Monthly	0.043	0.223	0.046	0.108	0.103	0.211*	−0.063	0.083	0.004	0.000	−0.058
ATM: Delta-hedged calls	Daily	−0.070	−0.126	−0.231	−0.200	−0.200	−0.182	−0.066	−0.068	−0.060	−0.058	−0.055
	Weekly	0.318	0.253	0.071	0.115	0.114	0.155	0.302	0.312	0.308	0.308	0.274
	Monthly	0.052	0.020	−0.181*	−0.134	−0.133	−0.070	0.073	0.075	0.067	0.070	0.032
OTM: Naked puts	Daily	0.103***	−0.097***	−0.077***	−0.103***	−0.103***	−0.102***	0.101***	0.118***	0.119***	0.133***	0.132***
	Weekly	0.448***	0.293**	0.239*	0.246*	0.269*	0.28**	0.502***	0.536***	0.568***	0.569***	0.569***
	Monthly	0.456	1.121**	0.646*	0.808*	0.819*	0.818*	0.591*	0.588*	0.552*	0.565*	0.57*
ATM: Naked puts	Daily	0.029	−0.007	−0.156***	−0.121***	−0.115***	−0.115***	0.023	0.019	0.024	0.022	0.023
	Weekly	0.149	0.095	−0.028	0.030	0.033	0.034	0.145	0.147	0.141	0.142	0.142
	Monthly	−0.016	0.022	−0.175	−0.143	−0.143	−0.142	−0.008	0.017	0.001	0.002	0.001
OTM: Delta-hedged puts	Daily	−0.042	−0.064	−0.074	−0.072	−0.072	−0.072	−0.039	−0.035	−0.036	−0.037	−0.038
	Weekly	0.031	0.005	−0.007	−0.005	−0.005	−0.006	0.042	0.043	0.043	0.044	0.039
	Monthly	0.321*	1.121	0.239	0.244	0.245	0.245	0.338**	0.344**	0.331*	0.342**	0.303*
ATM: Delta-hedged puts	Daily	−0.069	−0.100	−0.174	−0.152	−0.152	−0.156	−0.075	−0.073	−0.071	−0.071	−0.083
	Weekly	0.112	0.151	0.048	0.072	0.072	0.065	0.206	0.177	0.169	0.168	0.114
	Monthly	0.355***	0.332**	0.167	0.193	0.193	0.193	0.498***	0.378***	0.526***	0.53***	0.538***

Note: The sample consists of out-of-the-money (OTM) options defined as $0.97 < X/S < 1.03$ and at-the-money (ATM) options defined as $X/S \geq 1.03$, with maturity ranging from 6 to 60 days. We present both naked and delta-hedged positions for calls and puts at daily (1-day), weekly (5-day) and monthly (20-day) horizons. All options are sold based on specific model prices, which are obtained using 1-, 5-, or 20-day-ahead volatility forecasts, respectively, from competing models: implied volatility (IV), generalized autoregressive heteroskedasticity model with leverage (GARCH-L), realized volatility (RV), autoregressive fractionally integrated moving average (ARFIMA), ARFIMA model with leverage effect (ARFIMA-L), and the ARFIMA model with leverage and jump (ARFIMA-LJ). Delta-hedged positions are based on 1-, 5-, or 20-day-ahead delta forecasts from the same competing volatility models. Significance levels for Sharpe ratios are based on robust p-values (Lo 2002). Only commission costs of 0.5% are employed and no bid–ask costs are assumed.

***Significant at the 1% level.

**Significant at the 5% level.

*Significant at the 10% level.

TABLE 8. Sharpe Ratios for Naked and Delta-Hedged Option Strategies: Bid–Ask Costs as in Santa-Clara and Saretto (2005) (S&P 500 Index Options: January 2000–December 2002).

		IV	GARCH-L	RV	ARFIMA	ARFIMA-L	ARFIMA-LJ	IV+GARCH-L	IV+RV	IV+ARFIMA	IV+ARFIMA-L	IV+ARFIMA-LJ
OTM: Naked calls	Daily	0.475***	0.533***	0.206***	0.197***	0.186***	0.186***	0.45***	0.443***	0.433***	0.43***	0.428***
	Weekly	1.305***	1.39***	1.089***	1.138***	1.135***	1.137***	1.292***	1.281***	1.288***	1.286***	1.286***
	Monthly	2.926***	3.454***	2.416***	2.779***	2.82***	2.827***	3.684***	3.711***	3.692***	3.696***	3.695***
ATM: Naked calls	Daily	0.251***	0.173***	−0.006	0.045	0.042	0.042	0.261***	0.248***	0.244***	0.248***	0.248***
	Weekly	0.397***	0.367***	0.089	0.159	0.163	0.163	0.307**	0.397***	0.4***	0.402***	0.402***
	Monthly	1.083***	1.097***	0.629**	0.736***	0.74***	0.743***	1.186***	1.232***	1.22***	1.223***	1.223***
OTM: Delta-hedged calls	Daily	−0.387	−0.339	−0.404	−0.378	−0.378	−0.077	−0.372	−0.388	−0.382	−0.381	−0.341
	Weekly	−0.143	0.050	0.038	−0.051	−0.062	−0.091*	−0.130	−0.171	−0.097	−0.142	−0.134
	Monthly	−0.197	−0.036	−0.216	−0.164	−0.169	0.070	−0.223	−0.150	−0.164	−0.169	−0.208**
ATM: Delta-hedged calls	Daily	−0.487	−0.513	−0.591	−0.565	−0.564	−0.548	−0.483	−0.484	−0.475	−0.474	−0.471
	Weekly	−0.173	−0.227	−0.353	−0.318	−0.319	−0.281	−0.185	−0.181	−0.185	−0.186	−0.224
	Monthly	−0.29**	−0.338***	−0.552***	−0.501***	−0.499***	−0.422***	−0.265**	−0.267**	−0.277***	−0.275**	−0.305***
OTM: Naked puts	Daily	0.092***	−0.101***	−0.079***	−0.105***	−0.106***	−0.105***	0.09***	0.107***	0.109***	0.122***	0.121***
	Weekly	0.432***	0.28**	0.228*	0.234*	0.257*	0.268*	0.486***	0.52***	0.551***	0.552***	0.553***
	Monthly	0.437	1.087**	0.623	0.782*	0.792*	0.791*	0.567*	0.564*	0.53*	0.542*	0.547*
ATM: Naked puts	Daily	0.011	−0.024	−0.168***	−0.134***	−0.129***	−0.128***	0.005	0.001	0.006	0.004	0.006
	Weekly	0.130	0.078	−0.043	0.014	0.017	0.018	0.127	0.129	0.122	0.123	0.123
	Monthly	−0.036	0.001	−0.190	−0.159	−0.159	−0.159	−0.028	−0.004	−0.020	−0.018	−0.019
OTM: Delta-hedged puts	Daily	−0.326	−0.346	−0.358	−0.356	−0.356	−0.355	−0.323	−0.319	−0.320	−0.320	−0.321
	Weekly	−0.281*	−0.305*	−0.313**	−0.312**	−0.312**	−0.314**	−0.273*	−0.274*	−0.272*	−0.272*	−0.276*
	Monthly	−0.032	−0.080	−0.108	−0.103	−0.103	−0.102	−0.011	−0.015	−0.021	−0.017	−0.043
ATM: Delta-hedged puts	Daily	−0.457	−0.480	−0.536	−0.517	−0.517	−0.516	−0.463	−0.460	−0.460	−0.459	−0.471
	Weekly	−0.290	−0.301	−0.380	−0.359	−0.359	−0.346	−0.248	−0.282	−0.294	−0.296	−0.340
	Monthly	−0.105	−0.134	−0.297*	−0.27*	−0.27*	−0.269*	−0.020	−0.093	−0.009	−0.006	0.003

Note: The sample consists of out-of-the-money (OTM) options defined as $0.97 < X/S < 1.03$ and at-the-money (ATM) options defined as $X/S \geq 1.03$, with maturity ranging from 6 to 60 days. We present both naked and delta-hedged positions for calls and puts at daily (1-day), weekly (5-day), and monthly (20-day) horizons. All options are sold based on specific model prices, which are obtained using 1-, 5-, or 20-day-ahead volatility forecasts, respectively, from competing models: implied volatility (IV), generalized autoregressive heteroskedasticity model with leverage (GARCH-L), realized volatility (RV), autoregressive fractionally integrated moving average (ARFIMA), ARFIMA model with leverage effect (ARFIMA-L), and the ARFIMA model with leverage and jump (ARFIMA-LJ). Delta-hedged positions are based on 1-, 5-, or 20-day-ahead delta forecasts from the same competing volatility models. Significance levels for Sharpe ratios are based on robust p-values (Lo 2002). In addition to commission costs of 0.5%, we consider bid–ask costs of 5.5% (5.9%) for calls (puts) based on Santa-Clara and Saretto (2005).

***Significant at the 1% level.

**Significant at the 5% level.

*Significant at the 10% level.

VIII. Conclusion

In this article we examine the incremental economic value of using RV in combination with option IV to forecast future IV for out-of-sample option pricing, trading, and hedging in the SPX options market. Based on a comprehensive out-of-sample analysis, we find that historical RV forecasts retain statistical superiority in the encompassing regressions and out-of-sample pricing tests, but do not have incremental economic value in option trading and hedging. This suggests that RV, which has significant economic value in volatility timing in asset-allocation decisions in the equity and bond markets, may not yield similar economic benefits in the context of SPX options markets in the presence of transaction costs.

Empirical evidence in this paper substantiates the need for future research on closed-form pricing of derivatives written on realized quadratic variation measures (see, e.g., Carr et al. 2005 for work in this direction). This follows the recent developments in the over-the-counter variance option and swap markets, where participants can speculate and hedge in the underlying quadratic variation measures.

Appendix A: Implied Volatility Function Approach

We first calculate implied volatility (IV) using the Black–Scholes model for all observed calls (or puts) on a given day t after applying the filters to the option data and dropping options that meet the following criteria: (1) options outside the upper and lower boundaries, (2) options below the minimum tick size of 3/8, (3) options below six-day maturities, and (d) options with IV higher than 2 are dropped.

Next, we define log volatility on day t as a polynomial function of that day's moneyness and maturity as below:

$$\begin{aligned}\ln\sigma_{t,i}^{IV} &= \alpha_0+\alpha_1\left(\frac{X_{t,i}}{S_t}\right)+\alpha_2\left(\frac{X_{t,i}}{S_t}\right)^2+\alpha_3(mat_{t,i})\\ &\quad+\alpha_4(mat_{t,i})^2+\alpha_5\left(\frac{X_{t,i}}{S_t}\right)(mat_{t,i}),\\ &\forall \quad i=1,2,\ldots N \text{ on day } t \end{aligned}\tag{A1}$$

where $\ln\sigma_{t,i}^{IV}$ is the day t log implied volatility for option i, $X_{t,i}$ is strike price for option i on day t, S_t is the underlying spot value on day t, $mat_{t,i}$ is the annualized maturity of option i on day t, and N refers to the total number of call (or put) options available on day t. The log function ensures nonnegative volatility values.

We obtain the parameter vector α, where $\alpha=\{\alpha_0,\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5\}$, on a given day t for the polynomial (A1) by minimizing the mean square error of the market and Black–Scholes model prices for all N number of observed calls (or puts) on that day, that is,

$$\operatorname*{Min}_{\alpha}\left\{\frac{1}{N}\sum_{i=1}^{N}\left[mktpr_{t,i} - BSpr_{t,i}(S_t, X_{t,i}, r_t, mat_{t,i}, \sigma_{t,i}^{iv}\left(S_t/X_{t,i}, mat_i, \alpha\right)\right]^2\right\}, \tag{A2}$$

where *Mkt* and *BS* refer to observed and Black–Scholes model option prices, respectively. The model price for option i on a given day t is a function of corresponding strike price, annualized maturity and that day's spot price (S_t), risk-free rate (r_t), and latent volatility, which is in turn a function of moneyness, maturity, and parameter vector of the polynomial.

Next, suppose that we have an option trading on day t with a given moneyness m and maturity of $T-t$ days. We then condition on day $t-k$ parameter vector α (where α is obtained from step 3 using specific moneyness m and maturity $T-t$ options trading on day $t-k$), and forecast IV on day t for the $T-t$ day horizon; k here refers to the forecasting horizon equal to 1, 5, or 20 days corresponding to one-day-, one-week-, or one-month-ahead forecasts.

Appendix B: Straddle Trading Strategy

We consider an option trading with $T-t$ days to maturity on day t. We obtain the volatility forecast for $T-t$ horizon on day t for each volatility model, as described in Section III. We use this measure as a volatility proxy to price an option on day t using the Black–Scholes pricing formula.

From the out-of-sample daily volatility forecast for day t from step 1, we obtain the delta-neutral straddle for that day. In particular, given that delta of the call and put are, respectively, $N(d_1)$ and $N(d_1) - 1$, the strike price x that gives a delta-neutral straddle on day t can be solved as:

$$N(d_1) + N(d_1) - 1 = 0; \quad x = \frac{S}{\exp\left(\left(-r_f - \frac{\sigma^2}{2}\right)(T-t)/365\right)},$$

where S, r_f, σ, *and* $(T-t)/365$ refer to the spot price, risk-free rate, volatility, and annualized option maturity, respectively. We next determine that traded straddle combination whose exercise price is closest to the x described above. We use only at-the-money (ATM) calls and puts in this exercise as it enables use of the Black–Scholes option-pricing model. We then obtain the price of the ATM delta-neutral straddle using the Black–Scholes model.

Next, we buy or sell the ATM delta-neutral straddle on day t depending on whether the straddle model price is under- or overpriced relative to the closing market straddle price on that day.

The rate of return is calculated as follows:

$$\text{Return on buying a straddle} = \frac{C_t + P_t - C_{t-1} - P_{t-1}}{C_{t-1} + P_{t-1}} - r_f$$

$$\text{Return on selling a straddle} = \frac{-(C_t + P_t - C_{t-1} - P_{t-1})}{C_{t-1} + P_{t-1}} + r_f,$$

where C_t and P_t refer to the call and puts prices of S&P 500 index options. When the straddle is purchased, we assume that the agent can borrow at a risk-free rate. Similarly, when the straddle is sold, we let the agent invest the proceeds in a risk-free rate asset.

We implement steps 1–4 for each trading day in the sample.

For all trades, we apply the bid–ask spread filter. As a result, the agent trades only when the percentage absolute price difference between the model and market price as a proportion of model price, that is, 100x|model price – market price|/model price is expected to exceed the straddle bid–ask spread of 5.70%, (where 5.70% is obtained as the average of the bid–ask spread of 5.5% for calls and 5.9% for puts; see Santa-Clara and Saretto 2005, Table 7). The bid–ask filter accounts for the bid–ask bounce effects and implies that percentage pricing errors $\leq$ 5.70 do not constitute any arbitrage opportunities.

In addition, we apply the bid–ask spread of 5.70% plus commission cost of 0.5% toward the transaction costs while calculating the net returns for the strategies. The 0.5% trading cost corresponds to a trader with a \$48,000 investment who pays a commission amounting to \$120 + 0.0025 of the dollar amount as per a standard commission schedule (see Hull 2000, p. 160).

References

Aït-Sahalia, Y., 2002, Telling from discrete data whether the underlying continuous time model is a diffusion, *Journal of Finance* 57, 2075–2112.

Aït-Sahalia, Y., and J. Jacod, 2009, Testing for jumps in a discretely observed process, *Annals of Statistics* 37, 184–222.

Aït-Sahalia, Y., P. Mykland, and L. Zhang, 2005, How often to sample a continuous-time process in the presence of market microstructure noise, *Review of Financial Studies* 18, 351–416.

Andersen, T. G., and L. Benzoni, 2009, Realized volatility, in *Handbook of Financial Time Series* (Springer Verlag, New York).

Andersen, T. G., T. Bollerslev, and F. X. Diebold, 2007, Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility, *Review of Economics and Statistics* 89, 701–20.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens, 2001, The distribution of realized stock return volatility, *Journal of Financial Economics* 61, 43–76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, 2001, The distribution of realized exchange rate volatility, *Journal of the American Statistical Association* 96, 42–55.

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

Andersen, T. G., T. Bollerslev, and D. Dobrev, 2007, No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications, *Journal of Econometrics* 138, 125–80.

Bakshi, G., C. Cao, and Z. Chen, 1997, Empirical performance of alternative option pricing models, *Journal of Finance* 52, 2003–49.

Bakshi, G., and N. Kapadia, 2003, Delta hedged gains and the negative volatility risk premium, *Review of Financial Studies* 16, 527–66.

Barndorff-Nielsen, O. E., and N. Shephard, 2003, Realized power variation and stochastic volatility, *Bernoulli* 9, 243–65.

Barndorff-Nielsen, O. E., and N. Shephard, 2004, Power and bipower variation with stochastic volatility and jumps, *Journal of Financial Econometrics* 2, 1–48.

Barndorff-Nielsen, O. E., and N. Shephard, 2006, Econometrics of testing for jumps in financial economics using bipower variation, *Journal of Financial Econometrics* 4, 1–30.

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

Bollen, N. P. B., and R. E. Whaley, 2004, Does net buying pressure affect the shape of implied volatility functions? *Journal of Finance* 59, 711–54.

Carr, P., H. Geman, D. B. Madan, and M. Yor, 2005, Pricing options on realized variance, *Finance and Stochastics* 9, 453–75.

Christensen, B. J., C. S. Hansen, and N. R. Prabhala, 2002, The telescoping error problem in options data, FEN Working Paper.

Christensen, B. J., and M. O. Nielsen, 2006, The implied realized-volatility relation with jumps in underlying asset prices, FEN Working Paper.

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

Christoffersen, P., 2003, *Elements of Financial Risk Management* (Academic Press, San Diego, CA).

Christoffersen, P., and K. Jacobs, 2004, The importance of the loss function in option valuation, *Journal of Financial Economics* 72, 291–318.

Clemen, R. T., 1989, Combining forecasts: A review and an annotated bibliography, *International Journal of Forecasting* 5, 559–83.

Coval, J. D., and T. Shumway, 2001, Expected option returns, *Journal of Finance* 56, 983–1009.

Diebold, F. X., and J. A. Lopez, 1996, Forecast evaluation and combination, in G. Maddala and R. Rao, eds.: *The Handbook of Statistics: Statistical Methods in Finance* (North-Holland, Amsterdam), 241–68.

Donaldson, R. G., and M. J. Kamstra, 2005, Volatility forecasts, trading volume, and the ARCH versus option-implied volatility trade-off, *Journal of Financial Research* 28, 519–38.

Dumas, B., J. Fleming, and R. E. Whaley, 1998, Implied volatility functions: Empirical tests, *Journal of Finance* 53, 2059–2106.

Engle, R. F., and G. M. Gallo, 2006, A multiple indicators model for volatility using intra-daily data, *Journal of Econometrics* 131, 3–27.

Fleming, J., 1998, The quality of market volatility forecasts implied by S&P100 index option prices, *Journal of Empirical Finance* 5, 317–45.

Fleming, J., C. Kirby, and B. Ostdiek, 2003, The economic value of volatility timing using realized volatility, *Journal of Financial Economics* 67, 473–509.

Ghysels, E., P. Santa-Clara, and R. Valkanov, 2006, Predicting volatility: How to get the most out of returns data sampled at different frequencies? *Journal of Econometrics* 131, 59–95.

Giot, P., and S. Laurent, 2004, Modeling daily value-at-risk using realized volatility and ARCH type models, *Journal of Empirical Finance* 11, 379–98.

Glosten, L., R. Jagannathan, and D. Runkle, 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* 48, 1779–1801.

Goyal A., and A. Saretto, Forthcoming, Cross-section of option returns and volatility, *Journal of Financial Economics*.

Granger, C. W. J., and R. Ramanathan, 1984, Improved methods of combining forecasting, *Journal of Forecasting* 3, 197–204.

Green, T. C., and S. Figlewski, 1999, Market risk and model risk for a financial institution writing options, *Journal of Finance* 54, 1465–99.

Hansen, L. P., 1982, Large sample properties of generalized method of moments estimation, *Econometrica* 50, 1029–54.

Hull, J., 2000, *Options, Futures, and Other Derivatives* (Prentice Hall, Englewood Cliffs, NJ).

Jha, R., and M. Kalimipalli, Forthcoming, The economic significance of conditional skewness in option markets, *Journal of Futures Markets*.

Jiang, G. J., and R. C. A. Oomen, 2008, Testing for jumps when asset prices are observed with noise—A "swap variance" approach, *Journal of Econometrics* 144, 352–70.

Jiang, G. J., and Y. S. Tian, 2005, Model-free implied volatility and its information content, *Review of Financial Studies* 18, 1305–42.

Jiang, G. J., and Y. S. Tian, 2007, Extracting model-free volatility from option prices: An examination of the VIX index, *Journal of Derivatives* 14, 1–26.

Koopman, S. J., B. Jungbacker, and E. Hol, 2005, Forecasting daily variability of the S&P 100 stock index using historical, realized and implied volatility measurements, *Journal of Empirical Finance* 12, 445–75.

Lee, S. S., and P. A. Mykland, 2008, Jumps in financial markets: A new nonparametric test and jump dynamics, *Review of Financial Studies* 21, 2535–63.

Lo, A., 2002, Statistics of Sharpe ratios, *Journal of Portfolio Measurement*, 36–52.

Martens, M., and J. Zein, 2004, Predicting financial volatility: High-frequency time-series forecasts vis-à-vis implied volatility, *Journal of Futures Markets* 24, 1005–28.

Nandi, S., 1998, How important is the correlation between returns and volatility in a stochastic volatility model? Empirical evidence from pricing and hedging in the S&P 500 Index option market, *Journal of Banking and Finance* 22, 589–610.

Neely, C. J., 2009, Forecasting foreign exchange volatility: Why is implied volatility biased and inefficient? And does it matter? *Journal of International Financial Markets, Institutions and Money* 19, 188–205.

Noh, J., R. Engle, and A. Kane, 1994, Forecasting volatility and option prices of the S&P index, *Journal of Derivatives* Fall, 17–30.

Pong, S., M. B. Shackleton, S. J. Taylor, and X. Xu, 2004, Forecasting currency volatility: A comparison of implied volatilities and AR(FI)MA models, *Journal of Banking and Finance* 28, 2541–63.

Santa-Clara, P., and A. Saretto, 2005, Option strategies: Good deals and margin calls, Working Paper, UCLA.