Model-Free Implied Volatility and Its Information Content[1]

George J. Jiang
University of Arizona
and
York University

Yisong S. Tian
York University

March, 2003

[1]Address correspondence to George J. Jiang, Finance Department, University of Arizona, Tucson, Arizona 85721-0108, phone: 520-621-3373, email: gjiang@eller.arizona.edu or Yisong S. Tian, Schulich School of Business, York University, 4700 Keele Street, Toronto, ON M3J 1P3, phone: 416-736-2100 ext. 77943, email: ytian@schulich.yorku.ca. We thank the Social Sciences and Humanity Research Council of Canada for financial support.

Abstract

We implement an estimator of the model-free implied volatility derived by Britten-Jones and Neuberger (2000) and investigate its information content in the S&P 500 index options. In contrast to the commonly used Black-Scholes implied volatility, the model-free implied volatility is not based on any specific option pricing model and thus provides a direct test of the informational efficiency of the option market. Our results suggest that the model-free implied volatility is an efficient forecast for future realized volatility and subsumes all information contained in the Black-Scholes implied volatility and past realized volatility. We also find that the model-free implied volatility is, under the log specification, an unbiased forecast for future realized volatility after a constant adjustment. These results are shown to be robust to alternative estimation methods and volatility series over different horizons.

There has been considerable debate on the use of the Black-Scholes (B-S, hereafter) implied volatility as an unbiased forecast for future volatility of the underlying asset. Since option prices reflect market participants' expectations of future movements of the underlying asset, the volatility implied from option prices is widely believed to be informationally superior to historical volatility of the underlying asset. If option markets are informationally efficient and the option pricing model (used to imply the volatility) is correct, implied volatility should subsume the information contained in other variables in explaining future volatility. In other words, implied volatility should be an efficient forecast of future volatility over the horizon spanning the remaining life of the option. This testable hypothesis has been the subject of many empirical studies.

Early studies find that implied volatility is a biased and an inefficient forecast of future volatility and it contains little or no incremental information beyond that of historical volatility. For example, Canina and Figlewski (1993) study the daily closing prices of call options on the S&P 100 index from March 15, 1983 through March 28, 1987 and conclude that implied volatility is a poor forecast of subsequent realized volatility on the underlying index. Based on an encompassing regression analysis, they find that implied volatility has virtually no correlation with future return volatility and does not appear to incorporate information contained in historical return volatility.

In contrast, several other studies including Day and Lewis (1992), Lamoureux and Lastrapes (1993), Jorion (1995), and Fleming (1998) report evidence supporting the hypothesis that implied volatility has predictive power for future volatility. Like Canina and Figlewski (1993), Day and Lewis (1992) also analyze S&P 100 options but over a longer period from 1983 to 1989. They find that the implied volatility has significant information content for weekly volatility. However, the information content of implied volatility is not necessarily higher than that of standard time series models such as GARCH/EGARCH. Lamoureux and Lastrapes (1993) reach a similar conclusion by examining implied volatility from options on ten individual stocks over the period between April 19, 1982 and March 31, 1984. Using data on currency options, Jorion (1995) find that implied volatility dominates the moving average and GARCH models in forecasting future volatility, although the implied volatility does appear to be a biased volatility forecast.

More recent research on the information content of implied volatility attempts to correct various data and methodological problems ignored in earlier studies. These newer studies consider longer time series, possible regime shift around the October 1987 crash and the use of non-overlapping samples. For example, Christensen and Prabhala (1998) find that implied volatility of S&P 100 options is an unbiased estimator of future volatility and it subsumes all information contained in historical volatility. Unlike previous studies, they use a much longer time series (from November 1983 to May 1995) to take into account possible regime shift around the October 1987 crash, and more importantly a monthly non-overlapping sample to ensure the validity of statistical tests. Instrumental variables (IVs, hereafter) are also used to correct for potential errors-in-variable problem in the B-S implied volatility. Blair, Poon and Taylor (2000) reach a similar conclusion by examining the forecast performance of the VIX volatility index. An innovation in their study is that they use high frequency instead of daily index returns to calculate realized volatility. Ederinton and Guan (2000) study implied volatility of futures options on the S&P 500 index and find that the apparent bias and inefficiency of implied volatility from early studies are due to measurement errors. After correcting for these measurement errors, they find that implied volatility of S&P 500 futures options is an efficient forecast for future volatility. Christensen, Hansen and Prabhala (2001) investigate the telescoping overlap problem in options data and conclude that statistical tests are no longer meaningful when such a problem is present. Using monthly non-overlapping data, they find support for the null hypothesis that implied volatility of S&P 100 options is an efficient forecast of future volatility and subsumes all information in historical volatility, both before and after the October 1987 crash. Using daily overlapping data with telescoping maturities, they indeed find that standard statistical tests reject the above null hypothesis. Since the daily sample is highly overlapped and telescoping, standard statistical inferences are no longer valid. Their new insight provides a convincing argument against the early evidence on forecast inefficiency of implied volatility.

However, nearly all previous research on the information content of implied volatility focuses on the at-the-money B-S implied volatility. Granted, at-the-money options are generally more actively traded than other options and are certainly a good starting point. By concentrating on at-the-money options alone, however, these studies fail to incorporate information contained in other options. In addition, tests based on the B-S implied volatility are joint tests of market efficiency and the B-S model. The results are thus potentially contaminated with additional measurement errors due to model misspecification.

Beginning with Breeden and Litzenberger (1978), there has been considerable effort in deriving probability densities or stochastic processes that are consistent with all option prices observed at the same point in time. Derman and Kani (1994), Dupire (1994, 1997), Rubinstein (1994, 1998), and Derman, Kani and Chriss (1996) are among the first to derive binomial or trinomial tree models that are consistent with a subset or the complete set of observed option prices. The early work, though path breaking, falls into a class of deterministic volatility processes and is criticized by Dumas, Fleming and Whaley (1998) for their poor out-of-sample pricing performance. Buraschi and Jackwerth (2001) present empirical evidence that volatility is not deterministic.

Built on these early models, several recent studies such as Ledoit and Santa-Clara (1998), Derman and Kani (1998) and Britten-Jones and Neuberger (2000) generalize implied tree models to incorporate stochastic volatility and/or jumps. In particular, Britten-Jones and Neuberger

(2000) derive a model-free implied volatility that is the expected sum of squared returns under the risk-neutral measure. Unlike the traditional concept of implied volatility, their model-free implied volatility is not based on any particular option pricing model. In deriving this new implied volatility, no assumption is made regarding the underlying stochastic process except that the underlying asset price and volatility do not have jumps. No-arbitrage conditions are invoked to extract common features of all stochastic processes that are consistent with observed option prices. In particular, the set of prices of options expiring on the same date is used to derive the risk-neutral expected sum of squared returns between the current date and option expiry date (Proposition 2). In other words, the risk-neutral variance of the underlying asset is fully specified by market prices of options that span the corresponding horizon. Following Britten-Jones and Neuberger (2000), this risk-neutral variance will be referred to as the model-free implied variance and its square root the model-free implied volatility. For clarity, we will refer to the conventional implied volatility as the B-S implied volatility or simply implied volatility.

In this study, we implement the model-free implied volatility and investigate its information content and forecast efficiency. In contrast to the commonly used B-S implied volatility, the model-free implied volatility has the advantage that it is not based on any particular option pricing model and extracts information from all relevant observed option prices. It provides a direct test of the informational efficiency of the option market, rather than a joint test of market efficiency and the assumed option pricing model. We compare and contrast the efficiency of the model-free implied volatility as a volatility forecast against two other commonly used volatility forecast – the B-S implied volatility and past realized volatility. We choose to conduct our empirical tests using the S&P 500 options traded on the Chicago Board Options Exchange (CBOE). As the calculation of the model-free implied volatility requires European options, the more actively traded S&P 100 options are not suited for our study. Following prior research, we make every effort to eliminate any potential measurement errors in option prices. We use tick-by-tick data rather than daily closing data in order to eliminate non-synchronous trading problems. We apply commonly used data filters to eliminate observations that are most likely to be contaminated with measurement errors. We use mid bid-ask quotes instead of transaction prices to avoid any potential bias from the bid-ask bounce. We use monthly non-overlapping samples to avoid the telescoping overlap problem described by Christensen, Hansen and Prabhala (2001). We also use high-frequency index returns to obtain a more accurate estimate of realized volatility. Following prior literature, we employ several univariate and encompassing regressions to analyze the forecast efficiency and information content of different volatility measures. In particular, we place greater emphasis on the log volatility specification which avoids the positivity constraint when volatility is directly used in the analysis. This setup also has the advantage of being more conformable with the normality assumption required by standard statistical tests.

Consistent with many early studies in the literature, our OLS results show that the B-S implied volatility contains more information than past realized volatility in forecasting future volatility and is a biased and inefficient estimator of future realized volatility. We also analyze IV regressions that are designed to filter out measurement errors in implied volatility and find evidence that the B-S implied volatility subsumes all information contained in past realized volatility and can be regarded as an unbiased forecast for future volatility with a constant adjustment. These results confirm that tests based on the B-S implied volatility are subject to model misspecification errors due to the well-documented pricing biases of the B-S model from numerous empirical studies. Our results are consistent with findings from more recent research such as Christensen and Prabhala (1998) and Ederington and Guan (2000).

In constrast, the model-free implied volatility used in our study is independent of any option pricing model and provides a vehicle for a direct test of market efficiency. We find that it is an efficient forecast for future realized volatility and, under the log specification, an unbiased forecast for future realized volatility after a constant adjustment. In particular, we find that the model-free implied volatility subsumes all information contained in the B-S implied volatility and past realized volatility and can be regarded as an unbiased estimator of future realized volatility with a constant adjustment. Unlike the B-S implied volatility, however, measurement errors are not apparent in the model-free implied volatility since both the OLS and IV regressions provide nearly identical estimates for coefficients and test statistics. These results support the informational efficiency of the option market. Information in historical returns are correctly incorporated in option prices.

The rest of the paper proceeds as follows. The next section describes the model-free implied volatility and a simple method to calculate it using observed option prices. We use simulation experiments to demonstrate the validity and accuracy of this method. In Section 2, we discuss the data used in this study and the empirical method for implementing the model-free and B-S implied volatility. In Section 3, the relationship between the model-free and B-S implied volatilities is examined. The information content and forecast efficiency of the two implied volatility measures are investigated in Section 4 using monthly non-overlapping samples. Section 5 conducts robustness tests to ensure the generality of our results. Conclusions and discussions for future research directions are in the final section.

1 Model-Free Implied Volatility

In this section, we first describe the model-free implied volatility derived by Britten-Jones and Neuberger (2000). Several modifications are suggested for its empirical implementation. Simulations are then conducted to verify the validity of the model-free implied volatility.

1.1 Model-Free Implied Volatility

Suppose a complete set of call options with a continuum of strike prices (K) and a continuum of maturities (T) are traded on an underlying asset. The current price of the option is denoted by $C(T, K)$ while the current and future price of the underlying asset are S_0 and S_t, respectively. The price of the underlying asset follows a diffusion with time-varying (deterministic or stochastic) volatility. For simplicity, it is further assumed (and relaxed subsequently) that the underlying asset does not make any interim payments such as dividends or interest and the risk-free rate of interest is zero. Under this quite general setting, Britten-Jones and Neuberger (2000) show that the risk-neutral expected sum of squared returns between two arbitrary dates (T_1 and T_2) is completely specified by the set of prices of options expiring on the two dates [see Proposition 2 in Britten-Jones and Neuberger (2000)]:[1]

$$E_0^Q\left[\int_{T_1}^{T_2}\left(\frac{dS_t}{S_t}\right)^2\right] = 2\int_0^\infty \frac{C(T_2, K) - C(T_1, K)}{K^2}dK, \tag{1}$$

where the expectation is taken under the risk-neutral probability measure. This relationship means that the asset return variance (or squared volatility) under the risk-neutral measure can be obtained from the set of option prices observed at a single point in time. Because it is based on observed option prices and derived without any specific assumption about the underlying stochastic process (such as the geometric Brownian motion assumed in the B-S model), it is model free. Following Britten-Jones and Neuberger (2000), the right-hand-side (RHS) of Equation (1) will be referred to as the model-free implied variance (or squared volatility) and its square root the model-free implied volatility.

To implement and calculate the model-free implied volatility, several modifications to Equation (1) are required. First, in order to apply Equation (1) to cases where the underlying asset pays

[1] Derman and Kani (1998) derive a similar result in a more general setting.

interim income, we need to reduce the observed price of the underlying asset by the PV of all expected future income to be paid prior to the option maturity. For simplicity, we abuse the notation by continuing to use S_t to denote the adjusted price of the underlying asset on date t. Second, the zero interest rate assumption is relaxed by utilizing the relationship between options on the underlying asset and options on a related asset with zero drift. It is straightforward to show that Equation (1) changes to the following equation when interest rate is positive:

$$E_0^Q\left[\int_{T_1}^{T_2}\left(\frac{dS_t}{S_t}\right)^2\right] = 2\int_0^{\infty}\frac{C(T_2, K\,e^{rT_2}) - C(T_1, K\,e^{rT_1})}{K^2}dK, \tag{2}$$

where r is the risk-free rate. This is done by considering a hypothetical security defined by:

$$F_t = S_t e^{-rt},$$

which has zero drift under the risk-neutral measure. From the following relationship:

$$E_0^Q\left[\exp(-rT)\,\max\{S_T - K\exp(rT), 0\}\right] = E_0^Q\left[\max(F_T - K, 0)\right],$$

it is clear that an option on F_t with strike price K is equivalent to a corresponding option on S_t with strike price $K\,e^{rT}$. Equation (2) follows right away.

Third, since we are mostly interested in forecasting volatility over a period between the present and some future date T, the following special case of Equation (2) is more useful in empirical research:

$$E_0^Q\left[\int_0^{T}\left(\frac{dS_t}{S_t}\right)^2\right] = 2\int_0^{\infty}\frac{C(T, K\,e^{rT}) - \max(S_0 - K, 0)}{K^2}dK. \tag{3}$$

In other words, only the set of call options maturing on date T are needed to forecast volatility between date 0 and date T.

The final modification is due to the discrete nature of strike prices for traded options. Options with strike prices in discrete increments are typically traded in the marketplace and usually in a range not too far away from the price of the underlying asset. As a result, the required option prices with a continuum of strike prices in (3) are not available in empirical applications. A discretized version of Equation (3) is thus required and can be stated as follows:

$$E_0^Q\left[\sum_{i=0}^{n-1}\left(\frac{S_{t_{i+1}}-S_{t_i}}{S_{t_i}}\right)^2\right]=\left(u-\frac{1}{u}\right)\sum_{i=-m}^{m}\frac{C(T,K_i\,e^{rT})-\max(S_0-K_i,0)}{K_i}, \tag{4}$$

where

$$t_i = ih, \quad \text{for } i=0,1,2,\cdots,n,$$

$$h=\frac{T}{n},$$

$$K_i = S_0 u^i, \quad \text{for } i=0,\pm1,\pm2,\cdots,\pm m,$$

$$u=(1+k)^{1/m},$$

k is a positive constant, and n and m are positive integers. In this discretization scheme, the time interval $[0,T]$ is divided into n equal subintervals. The underlying asset can take on any of the $2m+1$ values covered by K_i's over the range $[S_0/(1+k),\ S_0(1+k)]$. Approximation errors may occur if either k, m or n is not sufficiently large.

1.2 Simulation I: Discretization and Truncation Errors

Before the model-free implied volatility is used in our empirical test, we carry out several simulation experiments to verify the validity and accuracy of the volatility calculation. Two types of approximation errors are of particular concern in Equation (4): truncation and discretization. Truncation errors occur when k is not sufficiently large. The two tails of the distribution beyond $[S_0/(1+k),\ S_0(1+k)]$ are ignored in the calculation, resulting in truncation errors. On the other hand, discretization errors occur when m is not sufficiently large. A continuum of strike prices are sampled by a discrete set of strike prices, introducing discretization errors.

To examine the size and significance of both types of approximation errors, we conduct a simulation experiment using Heston's (1993) stochastic volatility model where the price of the underlying asset and its volatility are specified by the following stochastic process:[2]

$$\begin{aligned} dS_t/S_t &= rdt + V_t^{1/2} dW_t, \\ dV_t &= (\theta_v - \kappa_v V_t)dt + \sigma_v V_t^{1/2} dW_t^v, \\ dW_t dW_t^v &= \rho dt, \end{aligned}$$

where V_t is the return variance of the underlying asset and W_t and W_t^v are standard Wiener processes. Model parameters are adopted from Bakshi, Cao and Chen (1997): $\theta_v = 0.08$, $\kappa_v = 2$, $\sigma_v = 0.225$, and $\rho = -0.5$. The initial instantaneous volatility (V_0) is set equal to the unconditional mean of the volatility process at 0.2. The expected integrated volatility of the asset return under the risk-neutral measure is also 0.2. The initial asset price (S_0) is 100 and the forecast horizon (T) is three months. The risk-free rate (r) is assumed to be zero. The results are robust to the choice of parameters and we merely use these parameters to illustrate typical approximation errors.

To implement the model-free implied volatility, we first calculate option prices using Heston's closed-form solution and the above known model parameters. For a given truncation parameter k and discretization parameter m, we need to calculate prices of $2m+1$ call options with strike prices specified in Equation (4). The right-hand-side (RHS) of Equation (4) is then used to calculate the model-free implied volatility. The approximation error is the difference between the calculated volatility and the known volatility of 0.2.

To visualize the approximation error, we plot the approximation error against the discretization parameter m for a given truncation parameter k. As m increases from 1 to a sufficiently large

[2] Note that the B-S model is a special case of Heston's (1993) stochastic volatility model. A separate simulation experiment shows that the approximation error is even smaller if the simple B-S model is used as the true model for option prices. The results are not reported here for brevity.

number (e.g., 40), the discretization error is expected to diminish but the truncation error may remain depending on the size of k used. Several plots are presented in Figure 1 for various truncation parameters (ranging from $k = 0.1$ to $k = 0.4$) to illustrate the approximation error due to truncation and discretization. As shown in Figure 1, the approximation error converges quickly as m increases. However, the approximation error does not seem to diminish (to zero) if the truncation parameter is too small (e.g., $k = 0.1$), indicating the presence of truncation errors. This is not surprising because a relatively small range of strike prices are used in this case. When $k = 0.1$, for example, only strike prices in the narrow range from [90.91, 110] are used in the calculation, covering approximately 10% of the asset price on either side. When a larger k is used, the truncation error diminishes rapidly. In fact, both types of approximation errors are negligible when $k \geq 0.2$ and $m \geq 20$.

1.3 Simulation II: Limited Availability of Strike Prices

In the previous simulation experiment, we made the assumption that options with any strike prices are available for the computation of the model-free implied volatility. We showed how the model-free implied volatility can be calculated with negligible approximation error. This assumption is, however, not realistic because not all strike prices are listed for trading in the marketplace. How well does the approximation work in this case? In a new simulation experiment, we make the alternative assumption that only a limited set of strike prices are available. To imitate actual market conditions, we adopt the strike price structure of S&P 500 options on September 23, 1988 as a prototype.[3] On this date, the closing index value is approximately 270 (S_0) and the listed strike prices (K) are 200, 220 to 310 with a 5-point increment, 325 and 350.

We continue to use Heston's (1993) stochastic volatility model and the same parameter values

[3]The choice of the sampling date is not important because we don't use observed option prices in our simulation. We only need the set of strike prices listed on that date to provide a sense of availability of strike prices in the marketplace.

used in the previous simulation experiment. We calculate the model prices of call options with the listed strike prices and a specified maturity. These model prices (instead of market prices) are used as inputs in our simulation experiment. Using model prices has the advantage that the true volatility is known and the approximation error due to our numerical implementation can be accurately measured.

To calculate the model-free implied volatility, we need to evaluate the RHS of Equation (4). This requires prices of call options with strike prices K_i for $i = 0, \pm 1, \pm 2, \cdots, \pm m$. Because these options are not listed on our sample date, their prices are unavailable (even though we could calculate their prices using the specified stochastic volatility model). Instead, their prices must be inferred from the prices of listed options. Among the approaches attempted in previous studies, the curve-fitting method is most practical and effective. Although some studies apply the curve-fitting method directly to option prices [e.g., Bates (1991)], the severely non-linear relationship between option price and strike price often leads to numerical difficulties. Following Shimko (1993) and Aït-Sahalia and Lo (1998), we apply the curve-fitting method to implied volatilities instead of option prices. Prices of listed calls are first translated into implied volatilities using the B-S model. A smooth function is then fitted to these implied volatilities. We then extract implied volatilities at strike prices K_i from the fitted function. The B-S model is used once again to translate the extracted implied volatilities into call prices. With the extracted call prices, the model-free implied volatility is calculated using the RHS of Equation (4). Following Bates (1991) and Campa, Chang and Reider (1998), we use cubic splines in the curve-fitting of implied volatilities. Using cubic splines has the advantage that the obtained volatility function is smooth everywhere (with a continuous second-order derivative) and provides an exact fit to the known implied volatilities.[4]

[4]For a more detailed review of curve-fitting and other methods, see the survey paper by Jackwerth (1999).

Note that this curve-fitting procedure does not make the assumption that the B-S model is the true model underlying option prices. It is merely used as a tool to provide a one-to-one mapping between option prices and implied volatilities. It is also important to note that cubic splines are effective only for extracting option prices between the maximum and minimum available strike prices. For options with strike prices beyond the maximum and minimum, we make the convenient assumption that implied volatility is identical to that of the option with the nearest available strike price (either the maximum or minimum). In other words, we assume that the volatility function is flat beyond the range of available strike prices.

To demonstrate the accuracy of this procedure, we examine the model-free implied volatility using call prices calculated using Heston's stochastic volatility model with the same parameter values as in the previous simulation experiment. The discretization and truncation parameters used are $k = 0.4$ and $m = 35$, respectively. The model-free implied volatility is calculated to be 0.2016, 0.2009, 0.2003, 0.2001 and 0.2001 over time horizon of 30, 45, 60, 75, and 90 days, respectively. Comparing with the true volatility of 0.2, the approximation error is negligible in all cases. The accurate results are obtained despite of the fairly coarse discretization scheme used. These simulation experiments verify the validity of the model-free implied volatility and our implementation method.

2 Data and Summary Statistics

Data used in this study are from several sources. Intraday data on the S&P 500 (SPX) options are obtained from the CBOE. Daily cash dividends are obtained from the Standard and Poors' DRI database. High frequency data at 5-minute intervals for the S&P 500 index are extracted from the contemporaneous index levels recorded with the quotes of S&P 500 index options. In addition, daily Treasury bill yields, used as the risk-free interest rate, are obtained from the Federal Reserve

Bulletin.

Our sample period is from June 1988 to December 1994. This is chosen for two reasons. First, daily cash dividend series from Standard and Poors begins in June 1988. Following Harvey and Whaley (1991, 1992a, 1992b), we use daily cash dividends instead of constant dividend yield. This means that the earliest possible date in our sample is June 1988. Second, the calculation of model-free implied volatility requires option prices from a sufficient range of option strike prices (and maturities). Prior to 1988, trading volume and market depth in SPX options are frequently insufficient and there are not enough quoted prices from multiple strike prices to reliably calculate the model-free implied volatility.

Consistent with previous research [e.g., Bakshi, Cao and Chen (1997, 2000)], we use mid bid-ask quotes instead of actual transaction prices in order to avoid the bid-ask bounce problem. Standard data filters are applied to SPX option data. First, option quotes less than $3/8 are excluded from the sample. These prices may not reflect true option value due to proximity to tick size. Second, options with less than a week remaining to maturity are excluded from the sample. These options may have liquidity and microstructure concerns. Third, option quotes that are time-stamped later than 3:00 PM Central Standard Time (CST) are excluded. The stock market closes at 3:00 PM but the options market continues to trade for another 15 minutes. Nonsynchronous trading problem is avoided as index and option data are matched simultaneously. Next, our sample includes both call and put options. Following Aït-Sahalia and Lo (1998), we exclude in-the-money options from the sample. In-the-money options are more expensive and often less liquid than at-the-money or out-of-the-money options. We define in-the-money options as call options with strike prices less than 97% of the asset price or put options with strike prices more than 103% of the asset price. In addition, options violating the boundary conditions are eliminated from the sample. These options are significantly undervalued and the B-S implied volatility is in fact negative for these options. Finally, only option quotes from 2:00 to 3:00 PM CST are included in the sample. To reduce computational burden, we construct daily series for implied volatility surface. As option quotes over a range of strike prices and maturities are needed, we use all option quotes in the final trading hour before the stock market closes. The last trading hour is used because trading volume is typically higher than any other time of the day.

In addition, we use 5-minute high frequency index returns to estimate realized return volatility of the S&P 500 index. Several recent studies, including Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold and Ebens (2001), Andersen, Bollerslev, Diebold and Labys (2001) and Barndorff-Nielsen and Shephard (2001), argue that there is considerable advantage in using high frequency return data over daily return data in estimating realized return volatility. In particular, Andersen and Bollerslev (1998) show that the typical squared returns method for calculating realized volatility produces inaccurate forecasts for correctly specified volatility models if daily return data are used. The inaccuracy is a result of noise in the return data. They further show that the noisy component is diminished if high frequency return data (e.g., 5-minute returns) are used.

Our empirical tests require the construction of implied volatility surface from available option quotes. In theory, this is straightforward to do if the set of traded options covers a sufficient range of strike prices and maturities. For each available option, we first use the B-S model to back out the unobservable volatility from option quote. As SPX options are European, the observed index value is adjusted by subtracting the PV of all cash dividends expected to be paid before option maturity. Treasury-bill yield closest to matching the option maturity is used as the risk-free rate. If options are available for all strike prices and maturities, the calculated implied volatilities naturally form a surface.

In reality, however, only a limited number of exercise prices and maturities are available. A curve-fitting method is needed to extract the implied volatility surface from available option prices. Consistent with our simulation experiment and prior research [e.g., Bates (1991) and Campa, Chang and Reider (1998)], we use cubic splines to fit a smooth surface to available B-S implied volatilities. The curve-fitting method is implemented in two stages. During the first stage, available options are divided into non-overlapping groups of identical expiry months. For options in the same maturity group, a set of cubic splines are used in the strike price dimension to provide a smooth fit to B-S implied volatilities. Because option quotes during a one-hour window are used, these quotes are not observed simultaneously and index value may vary from one quote to the next. For this reason, we use option moneyness instead of strike price as the fitting variable (i.e., assuming sticky delta). Consistent with common practice in the literature, we define option moneyness as the ratio of strike price over adjusted index value (with the PV of expected cash dividends subtracted from the actual index value):

$$x = \frac{K}{S},$$

where K is the strike price and S is the index value adjusted for the PV of expected dividends to be paid prior to expiry date. We then partition the moneyness dimension by placing equally spaced points $x_i = 1 + i\delta$ where $i = 0, \pm 1, \pm 2, \cdots$ and δ is a small positive constant. These partition points form moneyness bins represented by non-overlapping intervals $[x_{i-1}, x_i)$. In the empirical implementation, we set $\delta = 0.01$ which is sufficiently small for our analysis. Options with moneyness falling into the same bin are considered to be equivalent for the purpose of curve fitting. If more than one option fall into the same bin, we calculate the average moneyness and implied volatility of all options in the bin. Each bin is then represented by the average moneyness ($\overline{x}_i$) and average implied volatility ($\overline{\sigma}_i$) of options in the bin. Of course, some bins may not have any options in them at all and these empty bins are not useful for curve fitting. Denote $\overline{x}_{\min}$ and $\overline{x}_{\max}$ as the minimum and maximum moneyness, respectively, of all non-empty bins. Cubic splines are then applied to fit a smooth curve to all non-empty bins based on average implied volatility and moneyness ($\overline{\sigma}_i$ and $\overline{x}_i$). Implied volatility at any moneyness point x_i between $\overline{x}_{\min}$ and $\overline{x}_{\max}$ can be extracted from the fitted curve. Beyond the maximum and minimum moneyness (i.e., $\overline{x}_{\min}$ and $\overline{x}_{\max}$), the volatility function is assumed to be flat which is shown to be a reasonable assumption in our simulation experiment. The cubic-spline fitting procedure is then repeated for all other maturity groups.

During the second stage of the curve-fitting method, cubic splines are used in the maturity dimension to complete the construction of the implied volatility surface. For a given level of moneyness (x_i), collect implied volatilities across all available maturity months obtained from the first stage. Fit a set of cubic splines in the maturity dimension to match all available implied volatilities at the given moneyness level. The same procedure is then repeated for all other moneyness levels, which completes the construction of the implied volatility surface.

To avoid the telescoping overlap problem described by Christensen, Hansen and Prabhala (2001), we extract implied volatilities from the implied volatility surface at predetermined fixed maturity intervals. The model-free implied volatility is then calculated using Equation (4) for fixed maturities of 30, 60, 120, and 180 (calendar) days. The B-S implied volatilty is extracted directly from the implied volatility surface for the same fixed maturities and moneyness levels from 0.94 to 1.06. We then calculate realized volatility for the matching time periods. While it is usually calculated from the price of a single option, the B-S implied volatility used in our empirical analysis is not. Due to curve-fitting using cubic splines, it should be interpreted as a weighted average of implied volatilities from all available options, with the options with nearby maturity and moneyness receiving a larger weight. Our approach is similar to the existing literature that uses a simple average of B-S implied volatilities from options at different levels of moneyness as an estimate of asset return volatility [e.g., Blair, Poon, and Taylor (2000)].

Table 1 provides summary statistics for the model-free implied volatility, B-S implied volatility and realized volatility. All volatility measures are estimated monthly on the Wednesday immediately following the expiry date of the month. Option maturity or forecast horizon ranges from 30 days to 180 days. For the B-S model, implied volatility is reported for at-the-money options only. For other moneyness levels, the results are similar but exhibit the commonly known "smile" pattern. As shown in Table 1, the realized volatility is substantially lower than both the B-S and model-free implied volatility over all horizons. While the implied volatility is often viewed as an aggregation of market information over the maturing period, the B-S and model-free implied volatility are both surprisingly more volatile than the realized volatility series. Based on the reported skewness and excess kurtosis, the log volatility is more conformable with the normal distribution. Finally, for both the B-S and model-free implied volatility, as the horizon increases the annualized volatility has an overall upward slope albeit not monotonically. However, the standard deviation decreases monotonically. In other words, the model-free implied volatility over longer time horizon fluctuates less.

Table 2 reports the correlation matrix of monthly 30-day volatility for the B-S implied volatility (from options with different degrees of moneyness), the model-free implied volatility, and the realized volatility. Overall, all B-S implied volatility series are highly correlated with the model-free implied volatility. As expected, the B-S implied volatility from at-the-money options has the highest correlation with the model-free implied volatility at 94.1%. While all implied volatility series are highly correlated with the future realized volatility, the model-free implied volatility has the highest correlation at 85.4%, followed by the at-the-money B-S implied volatility at 81.1%.

3 Relationship between the Model-Free and B-S Implied Volatility

Conceptually, the model-free implied volatility is quite different from the B-S implied volatility. The latter is derived from a single option by backing out the volatility from option price using the B-S formula while the former is independent of any option pricing model and derived directly from prices of all options maturing on relevant future dates. Whether these differences are measurable is, however, an empirical question. In this section, we examine the relationship between the model-free and B-S implied volatility and see if they are statistically different time series. Specifically, we test whether the B-S implied volatility is an unbiased estimator of the model-free implied volatility.

It is important to note that while there is a single point estimate for the model-free implied volatility for a given maturity, the B-S implied volatility can be calculated for options with different strike prices. Although a single point estimate of the B-S implied volatility from short-term, near-the-money option is often used in the empirical literature to represent the entire volatility smile, it is prudent to examine the relationship between the model-free implied volatility and the B-S implied volatility at different moneyness levels and maturities. We test the null hypothesis that the B-S implied volatility for a given moneyness level and maturity is an unbiased estimator for the corresponding model-free implied volatility, i.e.,

$$E[V_{t,\tau}^{MF}] = V_{t,x,\tau}^{BS},$$

where t is the time of observation, τ is option maturity, x is option moneyness, and V^{MF} and V^{BS} are model-free and B-S implied variance, respectively.

Consistent with prior literature, we estimate three different specifications of the null hypothesis, based on the volatility, variance and log variance:

$$\sigma_{t,\tau}^{MF} = \alpha_{x,\tau} + \beta_{x,\tau}^{BS} \sigma_{t,x,\tau}^{BS} + \epsilon_{t,x,\tau}, \tag{5}$$

$$V_{t,\tau}^{MF} = \alpha_{x,\tau} + \beta_{x,\tau}^{BS} V_{t,x,\tau}^{BS} + \epsilon_{t,x,\tau}, \tag{6}$$

$$\ln V_{t,\tau}^{MF} = \alpha_{x,\tau} + \beta_{x,\tau}^{BS} \ln V_{t,x,\tau}^{BS} + \epsilon_{t,x,\tau}, \tag{7}$$

where $E[\epsilon_{t,x,\tau}] = 0$. With these specifications, the null hypothesis can be stated as $H_0 : \alpha = 0$ and $\beta^{BS} = 1$ where the subscript for moneyness and maturity is dropped for simplicity.

Strictly speaking, the three specifications are not exactly equivalent. If one is deemed the theoretically correct specification, the other two are necessarily misspecified. However, as pointed out by Canina and Figlewski (1993), the bias induced by the nonlinear relationship between the regression specifications is likely to be small. The volatility specification in Equation (5) has been the most commonly adopted in previous research, testing the null hypothesis based on implied volatility or standard deviation of asset returns. It has the advantage of being nicely scaled and widely used in previous research.

Table 3 reports the regression results of the model-free implied volatility on the B-S implied volatility based on the three different specifications for the sample period from June 1988 to December 1994. Option maturities range from 30 days to 180 days and moneyness for the B-S implied volatility varies from 0.94 to 1.06. Results for subperiods in the sample are not materially different and are thus not reported. The numbers in the bracket are the standard errors of the parameter estimates, which are estimated following a robust procedure taking into account of the heteroscedastic and autocorrelated error structure [Newey and West (1987)].

As shown in the table, the three regressions produce similar results. Regression R^2s are generally quite high, ranging from 53% to 91%. The B-S implied volatility thus explains much of the variation in the model-free implied volatility. However, the null hypothesis that the B-S implied volatility is an unbiased estimator for the model-free implied volatility ($H_0 : \alpha = 0$ and $\beta^{BS} = 1$) is unanimously rejected at the 1% significance level in all regressions. These results imply that while the two implied volatilities may have some common information content, they do contain sufficiently different information.

In addition, the regression results do exhibit some notable variations across the three specifications, option maturity and moneyness, although most of the differences are relatively small. For example, regressions based on variance tend to have the highest R^2 while those based on log variance the lowest R^2. The main reason for such differences in R^2 is that taking the square root or logarithm of variance reduces the relative differences among observations in the sample thus diminishing the linear pattern (or visually clouding the scatter diagram) and lowering R^2. Furthermore, both the regression R^2 and the estimate of β^{BS} are decreasing functions of option maturity. As option maturity increases from 30 days to 180 days, the regression R^2 drops from the range between 79% and 91% to between 56% and 72% while the estimate of β declines from the range between 0.74 and 1.54 to between 0.50 and 0.86. Option market liquidity tends to worsen as maturity increases, due to reduced trading activities for longer-term options. Prices of longer-term options are thus less likely to be informationally efficient, increasing the difference between model-free and B-S implied volatilities. Finally, the regression R^2 and estimates of coefficients are generally not significantly different across moneyness levels.

4 The Information Content of Implied Volatility

Prior research has extensively examined the information content of the B-S implied volatility. While the empirical results have been mixed, most recent empirical studies seem to agree that the B-S implied volatility is not informationally efficient. It does not subsume all information contained in historical volatility and is a biased (upward) forecast for future realized volatility. However, these studies typically adopt a point estimate of the B-S implied volatility obtained from the price of a single option or prices of several options. Information contained in other options is discarded altogether, which might have led to the inefficiency conclusion. In this study, we examine, for the first time, the information content of a volatility forecast that is derived from the entire set of all available option prices – the model-free implied volatility. Because all available options are used to extract the risk-neutral volatility, it is more likely to be informationally efficient. To nest most previous research within our framework, we compare and contrast three different volatility forecasts – the model-free implied volatility, the B-S implied volatility and historical volatility. As mentioned previously, the model-free and B-S implied volatilities are calculated off the constructed implied volatility surface while the realized volatility is calculated using high-frequency asset returns (i.e., 5-minute S&P 500 index returns).

4.1 Regression Setup and Testable Hypotheses

Following prior research [e.g., Canina and Figlewski (1993) and Christensen and Prabhala (1998)], we employ both univariate and encompassing regressions to analyze the information content of volatility forecast. In a univariate regression, realized volatility is regressed against one volatility forecast. In comparison, realized volatility is regressed against two or more different volatility forecasts in encompassing regressions. While the univariate regression focuses on the information content of one volatility forecast, the encompassing regression addresses the relative importance of different volatility forecasts and whether one volatility forecast subsumes all information contained in other volatility forecast(s).

The univariate regression for the B-S implied volatility is specified in three different forms:

$$\sigma_{t,\tau}^{RE} = \alpha_{x,\tau} + \beta_{x,\tau}^{BS}\sigma_{t,x,\tau}^{BS} + \epsilon_{t,x,\tau}, \tag{8}$$

$$V_{t,\tau}^{RE} = \alpha_{x,\tau} + \beta_{x,\tau}^{BS}V_{t,x,\tau}^{BS} + \epsilon_{t,x,\tau}, \tag{9}$$

$$\ln V_{t,\tau}^{RE} = \alpha_{x,\tau} + \beta_{x,\tau}^{BS}\ln V_{t,x,\tau}^{BS} + \epsilon_{t,x,\tau}. \tag{10}$$

Similar to the previous section, σ and V are asset return volatility and variance, respectively. The superscripts RE and BS stand for REalized and Black-Scholes, respectively.

Similarly, the information content of the model-free implied volatility can be investigated using the following univariate regression specifications:

$$\sigma_{t,\tau}^{RE} = \alpha_{\tau} + \beta_{\tau}^{MF}\sigma_{t,\tau}^{MF} + \epsilon_{t,\tau}, \tag{11}$$

$$V_{t,\tau}^{RE} = \alpha_{\tau} + \beta_{\tau}^{MF}V_{t,\tau}^{MF} + \epsilon_{t,\tau}, \tag{12}$$

$$\ln V_{t,\tau}^{RE} = \alpha_{\tau} + \beta_{\tau}^{MF}\ln V_{t,\tau}^{MF} + \epsilon_{t,\tau}, \tag{13}$$

where superscript MF stands for Model-Free.

The univariate regression for the lagged realized volatility is also specified in three different forms:

$$\sigma_{t,\tau}^{RE} = \alpha_{\tau} + \beta_{\tau}^{LRE}\sigma_{t,\tau}^{LRE} + \epsilon_{t,\tau}, \tag{14}$$

$$V_{t,\tau}^{RE} = \alpha_{\tau} + \beta_{\tau}^{LRE}V_{t,\tau}^{LRE} + \epsilon_{t,\tau}, \tag{15}$$

$$\ln V_{t,\tau}^{RE} = \alpha_{\tau} + \beta_{\tau}^{LRE}\ln V_{t,\tau}^{LRE} + \epsilon_{t,\tau}, \tag{16}$$

where the superscript LRE stands for Lagged REalized and $V_{t,\tau}^{LRE}$ is the realized variance of asset returns over the τ time period prior to time t, i.e. $[t-\tau, t)$. We refer to $\sigma_{t,\tau}^{LRE}$ and $V_{t,\tau}^{LRE}$ as lagged realized volatility and variance, respectively.

Similar to Christensen and Prabhala (1998), we formulate three testable hypotheses associated with the information content of volatility measures. First of all, if a given volatility forecast contains no information about future realized volatility, the slope coefficient (β) should be zero. This leads to our first testable hypothesis $H_0 : \beta = 0$. The subscript and superscript are dropped here and subsequently for brevity whenever there is no apparent confusion. Secondly, if a given volatility forecast is an unbiased estimator of future realized volatility, the slope coefficient (β) should be one and the intercept (α) should be zero. This testable hypothesis is formulated as $H_0 : \alpha = 0$ and $\beta = 1$. Finally, if a given volatility forecast is informationally efficient, the residuals in the corresponding regression should be white noise and orthogonal to the conditional information set of the market.

In addition to these univariate regressions, we also employ encompassing regressions to analyze the information content of volatility forecast. Consistent with prior research, we use the following encompassing regressions to investigate the relative information content among different volatility forecasts:

$$\sigma_{t,\tau}^{RE} = \alpha_{x,\tau} + \beta_{x,\tau}^{MF}\sigma_{t,\tau}^{MF} + \beta_{x,\tau}^{BS}\sigma_{t,x,\tau}^{BS} + \beta_{x,\tau}^{LRE}\sigma_{t,\tau}^{LRE} + \epsilon_{t,x,\tau}, \tag{17}$$

$$V_{t,\tau}^{RE} = \alpha_{x,\tau} + \beta_{x,\tau}^{MF}V_{t,\tau}^{MF} + \beta_{x,\tau}^{BS}\sigma_{t,x,\tau}^{BS} + \beta_{x,\tau}^{LRE}\sigma_{t,\tau}^{LRE} + \epsilon_{t,x,\tau}, \tag{18}$$

$$\ln V_{t,\tau}^{RE} = \alpha_{x,\tau} + \beta_{x,\tau}^{MF}\ln V_{t,\tau}^{MF} + \beta_{x,\tau}^{BS}\sigma_{t,x,\tau}^{BS} + \beta_{x,\tau}^{LRE}\sigma_{t,\tau}^{LRE} + \epsilon_{t,x,\tau}. \tag{19}$$

The univariate regressions discussed previously can be regarded as a restricted form of the encompassing regressions. In particular, the encompassing regressions adopted by prior research to examine the information content of the B-S implied volatility are also restricted forms of our encompassing regressions, with the coefficient associated with the model-free implied volatility restricted to zero:

$$\sigma_{t,\tau}^{RE} = \alpha_{x,\tau} + \beta_{x,\tau}^{BS}\sigma_{t,x,\tau}^{BS} + \beta_{x,\tau}^{LRE}\sigma_{t,\tau}^{LRE} + \epsilon_{t,x,\tau}, \tag{20}$$

$$V_{t,\tau}^{RE} = \alpha_{x,\tau} + \beta_{x,\tau}^{BS}V_{t,x,\tau_i}^{BS} + \beta_{x,\tau}^{LRE}V_{t,\tau}^{LRE} + \epsilon_{t,x,\tau}, \tag{21}$$

$$\ln V_{t,\tau}^{RE} = \alpha_{x,\tau} + \beta_{x,\tau}^{BS}\ln V_{t,x,\tau}^{BS} + \beta_{x,\tau}^{LRE}\ln V_{t,\tau}^{LRE} + \epsilon_{t,x,\tau}. \tag{22}$$

Based on the encompassing regressions, several additional testable hypotheses are formulated. First of all, if either the B-S or model-free implied volatility is informationally efficient relative to past realized volatility, we should expect the lagged realized volatility to be statistically insignificant in all encompassing regressions. This leads to the following null hypothesis $H_0 : \beta^{LRE} = 0$. If this hypothesis is supported, past realized volatility is redundant and its information content has been subsumed in implied volatility. A second hypothesis is formed as a joint test of information content and forecast efficiency. For the B-S implied volatility, this joint test is stated as $H_0 : \beta^{BS} = 1$ and $\beta^{LRE} = 0$. It states that the B-S implied volatility subsumes all information contained in lagged realized volatility and is an unbiased forecast for future volatility. The intercept term is ignored in this null hypothesis. If supported, the B-S implied volatility can be interpreted as an unbiased forecast for future volatility after a constant adjustment. A similar null hypothesis is formulated for the model-free implied volatility $H_0 : \beta^{MF} = 1$ and $\beta^{LRE} = 0$. Finally, we test whether the model-free implied volatility subsumes all information contained in both the B-S implied volatility and the lagged realized volatility and is an unbiased estimator for future realized volatility. The null hypothesis is formally stated as $H_0 : \beta^{MF} = 1$ and $\beta^{BS} = \beta^{LRE} = 0$.

It is likely that regression results from the three different specifications may not agree with each other. However, it should be noted that these specifications have different economic implications and suggest different relationships between implied volatility and expected future volatility. The specification in (11) suggests a linear relationship between implied volatility and expected future volatility. Since the difference between implied volatility and expected future volatility is determined by the risk premium, an additive risk premium of stochastic volatility is thus assumed under this specification. On the other hand, the specification in (13) suggests a log linear relationship between implied volatility and expected future volatility. In other words, the risk premium of stochastic volatility is assumed to be multiplicative. Based on empirical evidence and from a statistical point of view, there are several important advantages for the log linear specification. First, the positivity restriction on both the variables and the random error term is lifted. This is not true for either the volatility or variance specification [e.g., (5) or (6)] as neither volatility nor variance can take on negative values. Second, the log linear relationship is invariant to the use of either variance or volatility as $\ln V_t = 2 \ln \sigma_t$. Third, as documented in the literature, the log volatility or variance is more conformable with the normality assumption [e.g. Andersen, Bollerslev, Diebold and Labys (2001)]. Evidence on skewness and excess kurtosis provided in Table 1 also supports this argument. It is thus more reasonable to assume $\ln V_t$ rather than either V_t or σ_t to be normally distributed. Finally, the log specification is the least likely to be adversely affected by the presence of outliers in the regression analysis. These advantages ensure that the estimation procedure based on the log linear model is more robust and the test statistics have more power.

4.2 Results

Early empirical research in option pricing typically involves severely overlapped daily samples. As options expire on a fixed calendar date, implied volatilities calculated from the same option over two consecutive business days are likely to be highly correlated because the time horizons differ by just one day or at most several days (over the weekend or holidays). As demonstrated by Christensen, Hansen and Prabhala (2001), such overlapped samples may lead to the so-called "telescoping overlap problem" and render the t-statistics and other diagnostic statistics in the linear regression invalid.

Following their lead, we test our null hypotheses using monthly non-overlapping samples. We choose the Wednesday immediately following the expiration date (the Saturday following the third Friday) of the month. If it happens to be a holiday, we go to the following Thursday then the proceeding Tuesday. We select the monthly sample this way because option trading seems to be more active during the week following the expiration date and Wednesday has the fewest holidays among all weekdays. When calculating the monthly volatility measures, we fix the option maturity at 30 calendar days. Time series obtained this way are neither overlapping nor telescoping.

Table 4 reports the OLS regression results from the univariate and encompassing regressions using the monthly non-overlapping sample. While the estimate for the other two volatility measures is only a function of maturity, the estimate for the B-S implied volatility is a function of both option maturity and option moneyness. To keep it manageable, we only present the results for at-the-money options ($x = 1$) since B-S implied volatility from these options has the highest correlation with the realized volatility as shown in Table 2. Numbers in brackets underneath the parameter estimates are the standard errors, which are estimated following a robust procedure taking into account of heteroscedasticity [White (1980)]. The Durbin-Watson (DW, hereafter) statistic is not significantly different from two in any of the regressions, indicating that the regression residuals are not autocorrelated. The column with the heading "χ^2 test(a)" presents the test statistics from a χ^2-test in univariate regressions for the null hypothesis $H_0 : \alpha = 0$ and $\beta^j = 1$ where $j = MF$, BS, or LRE. Numbers in brackets below the test statistics are p-values. This χ^2-test is conducted to examine whether each volatility forecast is by itself an unbiased estimator for future realized volatility. Likewise, the column with the heading "χ^2test(b)" presents the test statistics from a χ^2-test in encompassing regressions for the null hypothesis that the coefficient is one for the implied volatility (either the model-free or the B-S) and zero for all other volatility measures. This second χ^2-test is conducted to see if one volatility forecast (the model-free or B-S implied volatility) subsumes all information contained in other volatility forecasts and at the same time is an unbiased estimator (with possible constant adjustment) for future realized volatility. The superscripts ***, ** and * indicate that the slope coefficient (β) for implied volatility is insignificantly different from one at the 10%, 5% and 1% critical level, respectively. Similarly, the superscripts $^{+++}$, $^{++}$ and $^{+}$ indicate that the slope coefficient (β) is insignificantly different from zero at the 10%, 5% and 1% critical level, respectively.

Consider first the results from univariate regressions. Several common features are observed from all univariate regressions. First of all, the slope coefficient (β) is positive and significantly different from zero at any conventional significance level, implying that all three volatility measures contain substantial information for future volatility. Secondly, the null hypothesis that the volatility forecast is an unbiased estimator of expected volatility ($H_0 : \alpha = 0$ and $\beta = 1$) is strongly rejected in all cases. This is true across all regression specifications and volatility forecasts. This result is not surprising because summary statistics in Table 1 indicate that both the model-free and B-S implied volatilities are on average much greater than realized volatility. In addition, the realized volatility declines gradually in our sample period after the October 1987 crash. The evidence is consistent with the existing option pricing literature which documents that the stochastic volatility is priced with a negative market price of risk (or equivalently a positive risk premium). Thus, the volatility implied from option price(s) is higher than their counterpart under the objective measure due to investors' risk aversion.

Some notable differences also exist in univariate regressions across model specifications and volatility measures. For instance, the regression R^2 is the highest for the model-free implied volatility regression ranging from 75% to 77% while it is the lowest for the lagged realized volatility regressions ranging from 27% to 43%. This evidence suggests that, among the three volatility measures, the model-free implied volatility explains the most of the variations in future realized volatility. In other words, the model-free implied volatility contains the most information among the three volatility measures while the lagged realized volatility the least. It also implies that previous empirical findings based on the B-S implied volatility have substantially underestimated the information content of volatilities implied by option prices. By aggregating the information contained in all options, the model-free implied volatility provides an improved test for the information efficiency of the option market. In addition, the regression R^2 in the univariate regression is substantially higher than those obtained in most previous studies using the B-S implied volatility. The most likely explanation for this improvement is the use of high frequency returns (i.e., the 5-minute returns of the S&P 500 index) in the calculation of realized volatility in our study. In most existing studies, the estimation of realized volatility is based on daily returns due to the unavailability of high frequency data. As demonstrated in several recent studies such as Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold and Labys (2001), Andersen, Bollerslev, Diebold and Ebens (2001), and Barndorff-Nielsen and Shephard (2001), the realized volatility estimator based on high frequency returns is more robust than that based on low frequency returns (e.g., daily returns). The realized volatility estimator based on high frequency returns can also reduce the bias due to the sample autocorrelation problem in low frequency returns [see for example, French, Schwert and Stambaugh (1987)].

Furthermore, t-statistics from univariate regressions show that the null hypothesis that the slope coefficient (β) is one cannot be rejected for the model-free implied volatility but is strongly rejected for the B-S implied volatility. From the log regression, the slope coefficient (β) for the model-free implied volatility is 0.93, insignificantly different from 1 at the 10% significance level.

It is thus not unreasonable to consider the model-free implied volatility as an unbiased forecast for future volatility after a constant adjustment (related to the intercept). We recognize that the slope coefficient for the model-free implied volatility is significantly different from 1 in the other two specifications. However, as suggested by the adjusted R^2, the log regression involving the model-free implied volatility is much better specified than the other two regressions. The evidence thus strongly favors a log linear relationship between the model-free implied volatility and expected future volatility. In contrast, the slope coefficient for the B-S implied volatility is statistically different from 1 at the 1% significance level in all three regression specifications. As a result, it cannot be regarded as an unbiased forecast for future realized volatility even if a constant adjustment is allowed. These results suggest that the B-S implied volatility is neither an efficient forecast nor an unbiased estimator for future realized volatility, consistent with previous research in the existing literature such as Jorion (1995) on foreign currency options and Lamoureux and Lastrapes (1993) on individual stock options.

Consider next the results from encompassing regressions. In these regressions, we compare the information content of one volatility forecast with other volatility forecast(s). In encompassing regressions involving the B-S implied volatility and lagged realized volatility [Equations (20) – (22)], the null hypothesis $H_{0:}$ $\beta^{BS} = 1$ and $\beta^{LRE} = 0$ is strongly rejected by the χ^2-test in all three specifications. This result suggests that the B-S implied volatility is not informationally efficient and does not subsume all information contained in past realized volatility. Additional support is provided by standard t-test that the coefficient of the lagged realized volatility (β^{LRE}) is significantly different from zero in two out of three regression specifications. The lagged realized volatility is insignificantly different from zero only in the variance regression (21). These findings are consistent with previous research [e.g., Christensen and Prabhala (1998)] and our results from univariate regressions.

Finally, we examine the encompassing regressions involving the model-free implied volatility. Depending on the model specification and volatility measures included in the regression, a total of nine encompassing regressions involving the model-free implied volatility are run. Combining the results from all nine regressions, we find strong evidence in support of the hypothesis that the model-free implied volatility is informationally efficient and subsumes all information contained in both the B-S implied volatility and the lagged realized volatility. To begin with, once the model-free implied volatility is included in the regression, the addition of either the B-S implied volatility or the lagged realized volatility does not improve the regression goodness-of-fit (adjusted R^2) at all. This is clear from a direct comparison between the univariate regression and the related encompassing regressions involving the model-free implied volatility. Secondly, in all nine encompassing regressions involving the model-free implied volatility, the t-test statistics suggest that the slope coefficients for the B-S implied volatility and the lagged realized volatility are both insignificantly different from zero at the 10% level. In fact, the estimated coefficient of the lagged realized volatility is virtually zero in all encompassing regressions involving the model-free implied volatility. The estimated coefficient for the B-S implied volatility is either nearly zero or much reduced from the corresponding estimate in the univariate regression. In addition, the t-test statistics in the log encompassing regressions suggest that the slope coefficient for the model-free implied volatility is insignificantly different from one at the 10% significance level. Furthermore, the χ^2-test in the log encompassing regression does not reject the null hypothesis that the slope coefficient is one for the model-free implied volatility and zero for all other slope coefficients. The p-value for the χ^2-test ranges from 0.353 to 0.397, far greater than any commonly used critical values and thus providing strong support for the null hypothesis. These findings provide strong support for the informational efficiency of the model-free implied volatility. It subsumes all information contained in the B-S implied volatility and the lagged realized volatility and is an unbiased forecast for future realized volatility after a constant adjustment.

5 Robustness Tests

Results from the previous section provide strong support for the information efficiency of the model-free implied volatility. Although the option market is found to be informationally efficient, it is important to extract information from all available options. The price of a single option is not sufficient to convey all useful information absorbed in the set of all available options. Volatility forecast extracted from one or a small number of options can be biased and informationally inefficient. These findings are striking and have the potential to shed new light on volatility forecasting and market efficiency. It is thus important to conduct robustness tests to ensure the generality of our results.

5.1 IV Regressions

Christensen and Prabhala (1998) employ IV regressions to correct for potential error-in-variable (EIV) problems. They recognize that the B-S implied volatility (their volatility forecast) may contain significant measurement errors due to either the early exercise premium in the American style S&P 100 index options, the possible non-synchronous observations of option quotes and index levels in their data set, or the misspecification error of the B-S option pricing model. It is well known that the error-in-variable problem tends to drive the slope coefficient downwards (biased toward zero). This may explain why the coefficient of the B-S implied volatility or variance (β^{BS}) in both the univariate and encompassing regressions is below one. Christensen and Prabhala (1998) find substantial differences between the estimation results from the OLS and IV regressions, supporting the existence of measurement errors in their volatility forecast (i.e., the B-S implied volatility).

Compared to Christensen and Prabhala (1998) and other related studies, our data samples are less prone to measurement errors for reasons discussed previously. In fact, our regressions involving the B-S implied volatility, both univariate and encompassing, have much higher R^2 than corresponding figures in existing studies. Nevertheless, it is still important to investigate whether there is any error-in-variable problem in the B-S and model-free implied volatility in our sample and whether the IV estimation procedure can improve the estimation results over and above those from the OLS regressions.

Following Christensen and Prabhala (1998), we apply a two-stage least squares regression to implement the IV estimation procedure. For the B-S implied volatility, we use the lagged realized volatility and lagged B-S implied volatility as IVs. Similarly, we use the lagged realized volatility, the lagged model-free implied volatility and the B-S implied volatility as IVs for the model-free implied volatility. Details of the two-stage least squares implementation can be found in Christensen and Prabhala (1998).

Table 5 reports the results from the IV regressions for the log specification for the monthly 30-day maturity sample examined in the previous section. Results from the other two specifications are omitted for two reasons. As explained previously, the log specification is the most appropriate for our analysis. In addition, whereas the IV results for the volatility and variance regressions are not qualitatively different from the OLS results, they are very much so for the log regression. Thus, we report and concentrate on the log regressions. Results from the OLS regressions in the first stage are reported in Panel A while those from the second stage are reported in Panel B.

Consider first the information content and forecast efficiency of the B-S implied volatility. The IV regression results are found in the top half of Panel B in Table 5. From the univariate log regression, the joint hypothesis that $H_0 : \alpha = 0$ and $\beta^{BS} = 1$ is again strongly rejected by the χ^2-test, consistent with the result from the OLS regression reported in Table 4. More interestingly, the β^{BS} coefficient increases from 0.74 in the OLS regression (Table 4) to 1.09 in the IV regression which is not statistically different from one at the 5% critical level based on the t-test. This result suggests that the B-S implied volatility, though not an entirely unbiased forecast for future volatility, can be regarded as an unbiased forecast for future volatility after a constant adjustment. This finding is both quantitatively and qualitatively different from the OLS regression results in Table 4, which completely rejects the unbiasedness of the B-S implied volatility as a forecast for future volatility.

Additional support for the forecast efficiency of the B-S implied volatility is found in the encompassing regression involving the lagged realized volatility. The coefficient for the B-S implied volatility (β^{BS}) increases from 0.58 in the OLS regression to 1.29 in the IV regression while the corresponding change in the coefficient for the lagged realized volatility (β^{LRE}) is from 0.27 to -0.20. However, estimates of both coefficients have noticeably large standard errors. Partly due to the substantially increased standard errors, the coefficient for the B-S implied volatility (β^{BS}) is now not statistically different from one and the coefficient for the lagged realized volatility (β^{LRE}) is not statistically different from zero, both at the 10% critical level. The χ^2-statistic for the joint hypothesis $H_0 : \beta^{BS} = 1$ and $\beta^{LRE} = 0$ has a p-value of 0.295, thus cannot reject the null hypothesis at any conventional significance level. This finding suggests that the B-S implied volatility subsumes all information contained in lagged realized volatility and is an unbiased forecast for future volatility (after a constant adjustment). The Hausman (1978) χ^2 statistics (one degree of freedom) for testing the EIV problem are 6.61 and 6.72 with p-values 0.0101 and 0.0095 for the two regressions, respectively. The test strongly suggests the presence of measurement error in the B-S implied volatility. The larger standard errors also imply that the IV regressions are less efficient than the OLS regressions, which weakens the result somewhat. Overall, these results are consistent with the findings in Christensen and Prabhala (1998) on the B-S implied volatility from S&P 100 index options. Christensen and Prabhala (1998) attribute the differences in statistical inference between the OLS and IV regressions to measurement errors in the B-S implied volatility. Although our sample may contain less measurement errors compared to their samples, the IV regression exhibits similar patterns and inferences.

Next, we examine the information content and forecast efficiency of the model-free implied volatility. The IV regression results are found in the bottom half of Panel B in Table 5. In both univariate and encompassing regressions, we find little change in the estimated coefficients either in sign or magnitude. What is changed is the increased standard errors and reduced regression R^2. However, the magnitude of the change is relatively small compared to that in corresponding IV regressions associated with the B-S implied volatility. Not surprisingly, all inferences remain unchanged for the model-free implied volatility. The Hausman (1978) χ^2 test also indicates that there is essentially no error-in-variable problem in the model-free implied volatility and the IV estimates are less efficient than the OLS estimates. As a result, the IV regressions continue to support the hypothesis that the model-free implied volatility subsumes all information contained in both the B-S implied volatility and lagged realized volatility and is an unbiased forecast for future volatility (after a constant adjustment). Our findings for the model-free implied volatility are thus robust to the estimation method used.

Because the B-S and model-free implied volatility are extracted from the same implied volatility surface constructed using cubic splines, the same type of measurement errors is likely to be present in both of them. Consequently, estimation error in the implied volatility surface is unlikely to be solely responsible for the differences between the OLS and IV regressions for the B-S implied volatility. This suggests that measurement errors in the B-S implied volatility are largely due to model misspecification. Another possibility is that measurement errors in individual option prices are largely offset when the information from a large number of options is aggregated to obtain the model-free implied volatility.

5.2 Monthly Overlapping Samples

In addition to the monthly non-overlapping samples of 30-day options, we also extract monthly samples of volatility measures for options with a fixed maturity varying from 60 to 180 days. We use the same monthly series of implied volatility surface constructed in Section 2 to extract the required implied volatility series. The realized volatility and lagged realized volatility series are calculated using 5-minute high-frequency index returns over matching time horizons. The monthly samples obtained this way exhibit some degree of overlapping. The longer the option maturity is, the more overlapping the sample becomes. We are interested in finding out whether or not regression results change if the monthly overlapping samples are used in the analysis instead of the non-overlapping samples used previously. Overlapping samples may lead to severe serial correlation and render the OLS test statistics invalid. As pointed out by Richardson and Smith (1991), however, correct test statistics can be computed using the GMM approach. It is also important to note that our monthly overlapping samples do not have the "telescoping overlap" problem described by Christensen, Hansen and Prabhala (2001). Although these monthly samples are overlapping, they are not "telescoping" because the option maturity is fixed rather than declining for each maturity series in our sample.

Table 6 reports the OLS results from the univariate and encompassing regressions for the log volatility measures over the 60-day, 120-day and 180-day maturity horizons. Regression results from the three model specifications are more or less the same. Because the log regression is econometrically better specified, we again concentrate on the log regressions in the analysis. We also repeat the analysis using IV regressions and find no material change in statistical inferences.

The most striking finding from Table 6 is that parameter estimates and statistical inferences from the overlapping samples are very similar to those in the monthly non-overlapping samples reported in Table 4. First of all, the slope coefficient from univariate regressions is consistently positive and statistically significant, supporting the hypothesis that all three volatility measures contain information regarding future realized volatility. In addition, while both the t-test and χ^2-test strongly reject the null hypothesis that the B-S implied volatility or lagged realized volatility is an unbiased forecast for future realized volatility, the results are mixed for the model-free implied volatility. The null hypothesis that the model-free implied volatility is an unbiased forecast for future realized volatility is rejected by the χ^2-test [$\chi^2(a)$]. However, the t-test indicates that the model-free implied volatility can be regarded as an unbiased estimator for future volatility with a constant adjustment (related to the intercept). These results are consistent with the findings from the monthly non-overlapping samples. Furthermore, the encompassing regressions using overlapping samples continue to show that the B-S implied volatility does not subsume all information contained in lagged realized volatility. The lagged realized volatility is consistently positive and statistically significant in the encompassing regression involving both the B-S implied volatility and lagged realized volatility. In contrast, the model-free implied volatility is found to subsume all information contained in the B-S implied volatility and lagged realized volatility even in overlapping samples. As long as the model-free implied volatility is included in the regression, neither the B-S implied volatility nor the lagged realized volatility is statistically significant any more. These inferences regarding the information content of volatility measures are again consistent with those from the non-overlapping samples.

There are also some notable but relatively minor differences between the results from overlapping and non-overlapping samples. As indicated by the DW statistics, there is an increased autocorrelation in the regression errors from the overlapping samples. In contrast, autocorrelation is nearly non-existent in the non-overlapping samples. Not surprisingly, the DW statistic declines as the forecast horizon increases from 60 to 180 days, indicating that the degree of autocorrelation increases as the sample becomes more overlapped. The standard errors in all regressions are thus estimated following a robust procedure taking into account of both heteroscedasticity and autocorrelation [Newey and West (1987)]. The number of lags used in the estimation is set equal to the number of overlapping periods in the regression. In addition, the overlapping samples produce slightly different test results for the null hypothesis that the volatility measure is an unbiased forecast for future volatility than the non-overlapping samples do. While the χ^2-test [$\chi^2(b)$] from encompassing regressions continues to reject the null hypothesis that the B-S implied volatility or lagged realized volatility is an unbiased forecast for future realized volatility, the results for the model-free implied volatility are not as clear cut as in the non-overlapping case. Although the same χ^2-test does not reject the null hypothesis that the model-free implied volatility is an unbiased forecast for future realized volatility for the 180-day horizon, it does so for the 60-day horizon at the 5% level and the 120-day horizon at the 10% level. In comparison, the null hypothesis is overwhelmingly supported by the non-overlapping samples with the p-value from the χ^2-test exceeding 0.30. The support for the null hypothesis is weaker in the overlapping samples.

6 Conclusions

In this paper, we empirically test the information content and forecast efficiency of implied volatility. Unlike the B-S implied volatility used in previous research, we implement and estimate the model-free implied volatility derived by Britten-Jones and Neuberger (2000). This new implied volatility has several advantages over its predecessor. First, it is independent of any option pricing model whereas the commonly used B-S implied volatility is based on the B-S model. The calculation of the model-free implied volatility requires option prices only, nothing else. Second, the model-free implied volatility extracts information from all relevant option prices instead of a single option as in the case of the B-S implied volatility. By aggregating information across options, the model-free implied volatility is more likely to be informationally efficient than the B-S implied volatility. Third, tests based on the model-free implied volatility is a direct test of market efficiency instead of a joint test of market efficiency and the assumed option pricing model. Evidence on market efficiency from the model-free implied volatility is thus not contaminated by model misspecification errors.

Our research makes several contributions to the related literature and enhances our understanding of the informational efficiency of the option market. First of all, we develop a simple method for the implementation of the model-free implied volatility. A curve-fitting technique based on cubic splines is used to construct the implied volatility surface, which is then used to calculate the model-free implied volatility. We design and conduct several simulation experiments to demonstrate the validity and accuracy of this method. Secondly, we use univariate and encompassing regressions to investigate the information content and forecast efficiency of the model-free implied volatility. Because our empirical tests are independent of any option pricing model, our evidence provides a more direct test of market efficiency than previous studies. Our findings support the informational efficiency of the option market. In particular, we find that the model-free implied volatility is an efficient forecast for future realized volatility and subsumes all information contained in the B-S implied volatility and past realized volatility. It can also be regarded, under the log specification, as an unbiased forecast for future realized volatility after a constant adjustment. Moreover, our results do not seem to be contaminated with measurement errors found in previous studies. Both OLS and IV regressions provide similar parameter estimates and statistical inferences. Likewise, the results do not vary significantly between overlapping and non-overlapping samples. We attribute the near absence of measurement errors to several factors. The most important factor is the use of the model-free implied volatility which eliminates model misspecification errors. In addition, we use high-frequency asset returns to estimate realized volatility and intra-day prices from European-style options. Finally, we avoid the telescoping overlap problem by extracting monthly, fixed-maturity implied volatility series from option prices.

There are still some unresolved issues that deserve attention in future research. To begin with, a stronger informational efficiency test can be formulated and used to examine market efficiency. In our current study, a single observation on past realized volatility is used as a proxy for historical information contained in asset returns. A better test can be implemented by incorporating a greater number of observations from past realized volatility. One possibility is to adopt a GARCH model based volatility forecast, which extracts information from all historical asset returns. Secondly, the telescoping overlap problem needs to be examined further in future research. Our evidence shows that it is likely the telescoping rather than the overlapping feature that is the root cause of the problem. However, this problem should be examined in greater details using daily samples as the telescoping overlap problem is expected to be more severe in daily samples. Finally, it may be interesting to examine the term structure of model-free implied volatilities. In order to focus on the information content and forecast efficiency, we do not address this issue in this study. Since the B-S implied volatilities exhibit well-known term structure patterns, it is interesting to find out if this is also true for the model-free implied volatility.

References

Aït-Sahalia, Y., and A.W. Lo, 1998, "Nonparametric estimation of state-price densities implicit in financial asset prices," *Journal of Finance*, 53, 499–547.

Andersen, T.G., and T. Bollerslev, 1998, "Answering the Critics: Yes, ARCH Models Do Provide Good Volatility Forecasts," *International Economic Review*, 39, 885-905.

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.

Barndorff-Nielsen, O.E. and N. Shephard, 2001, "Realized Power Variation and Stochastic Volatility Models", *Manuscript Oxford University.*

Bakshi, G., C. Cao, and Z. Chen, 1997, "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52, 2003-2049.

Bakshi, G., C. Cao, and Z. Chen, 2000, "Pricing and Hedging Long-Term Options," *Journal of Econometrics*, 94, 277-318.

Bates, D., 1991, "The Crash of '87: Was it Expected? The Evidence from Options Markets," *Journal of Finance*, 46, 1009-1044.

Blair, B., S.H. Poon, and S.J. Taylor, 2001, "Forecasting S&P 100 Volatility: The Incremental Information Content of Implied Volatility and High Frequency Index Returns," *Journal of Econometrics*, 105, 5-26.

Breeden, D.T., and R.H. Litzenberger, 1978, "Prices of State-Contingent Claims Implicit in Option Prices," *Journal of Business*, 51, 621-651.

Britten-Jones, M., and A. Neuberger, 2000, "Option Prices, Implied Price Processes, and Stochastic Volatility," *Journal of Finance*, 55, 839-866.

Buraschi, A., and J. Jackwerth, 2001, "The Price of a Smile: Hedging and Spanning in Option Markets," *Review of Financial Studies*, 14, 495-527.

Campa, J.M., K.P. Chang, and R.L. Reider, 1998, "Implied Exchange Rate Distributions: Evidence from OTC Option Markets," *Journal of International Money and Finance*, 17, 117-160.

Canina, L. and S. Figlewski, 1993, "The Informational Content of Implied Volatility," *Review of Financial Studies*, 6, 659-681.

Christensen, B.J., C.S. Hansen, and N.R. Prabhala, 2001, "The Telescoping Overlap Problem in Options Data," Working paper, University of Aarhus and University of Maryland.

Christensen, B.J., and N.R. Prabhala, 1998, "The Relation between Implied and Realized Volatility," *Journal of Financial Economics*, 50,125-150.

Day, T.E., and C.M. Lewis, 1992, "Stock Market Volatility and the Information Content of Stock Index Options," *Journal of Econometrics*, 52, 267-287.

Derman, E., and I. Kani, 1994, "Riding on a Smile," *Risk*, 7, 32-39.

Derman, E., and I. Kani, 1998, "Stochastic Implied Trees: Arbitrage Pricing with Stochastic Term and Strike Structure of Volatility," *International Journal of Theoretical and Applied Finance*, 1, 61-110.

Derman, E., I. Kani, and N. Chriss, 1996, "Implied Trinomial Trees of the Volatility Smile," *Journal of Derivatives*, 3, 7-22.

Dumas, B., J. Fleming, and R.E. Whaley, 1998, "Implied Volatility Functions: Empirical Tests," *Journal of Finance*, 53, 2059-2106.

Dupire, B., 1994, "Pricing with a Smile," *Risk*, 8, 76-81.

Dupire, B., 1997, "Pricing and Hedging with Smiles," In Dempster, M.A.H., and S.R. Pliska, eds.: *Mathematics of Derivative Securities*, Cambridge University Press, Cambridge, U.K.

Ederington, L.H., and W. Guan, 2000, "Is Implied Volatility an Informationally Efficient and Effective Predictor of Future Volatility?" NBER working paper, University of Oklahoma and Delaware State University.

Fleming, J., 1998, "The Quality of Market Volatility Forecast Implied by S&P 100 Index Option Prices," *Journal of Empirical Finance*, 5, 317-345.

French, K.R., G.W. Schwert, and R.F. Stambaugh, 1987, "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19, 3-30.

Harvey, C.R., and R.E. Whaley, 1991, "S&P 100 Index Option Volatility," *Journal of Finance*, 46, 1551-1561.

Harvey, C.R., and R.E. Whaley, 1992a, "Market Volatility Prediction and the Efficiency of the S&P 100 Index Option Market," *Journal of Financial Economics*, 31, 43-73.

Harvey, C.R., and R.E. Whaley, 1992b, "Dividends and S&P 100 Index Option Valuation," *Journal of Futures Markets*, 12, 123-137.

Harvey, C.R., and R.E. Whaley, 1992b, "Dividends and S&P 100 Index Option Valuation," *Journal of Futures Markets*, 12, 123-137.

Hausman, J., 1978, "Specification Tests in Econometrics," *Econometrica*, 46, 1251-1271.

Jackwerth, J.C., 1999, "Option-Implied Risk-Neutral Distributions and Implied Binomial Trees: A Literature Review," Journal of Derivatives, 6, 1-17.

Jorion, P., 1995, “Predicting Volatility in the Foreign Exchange Market,” *Journal of Finance*, 50, 507-528.

Lamoureux, C.G. and W.D. Lastrapes, 1993, “Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities,” *Review of Financial Studies*, 6, 293-326.

Ledoit, O., and P. Santa-Clara, 1998, “Relative Pricing of Options with Stochastic Volatility,” Working paper, University of California at Los Angeles.

Newey, W.K., and K.D. West, 1987, “A simple positive definite heteroskedasticity and autocorrelation consistent covariance matrix,” *Econometrica*, **55**, 703–708.

Richardson, M., and T. Smith, 1991, “Tests of Financial Models in the Presence of Overlapping Observations,” *Review of Financial Studies*, 4, 227-254.

Rubinstein, M., 1994, “Implied Binomial Trees,” *Journal of Finance*, 49, 771-818.

Rubinstein, M., 1998, “Edgeworth Binomial Trees,” *Journal of Derivatives*, 5, 20-27.

Shimko, D., 1993, “Bounds of Probability,” *Risk*, 6, 33-37.

White, H., 1980, “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity,” *Econometrica*, 48, 817-838.

Table 1
Summary statistics of monthly volatility series

τ	N	Mean	Std Dev	Skewness	Excess Kurtosis	Minimum	Maximum
Panel A: σ^{BS}							
30	78	14.78	4.396	1.1258	2.0636	6.912	31.21
60	79	15.71	4.399	1.0210	0.9198	8.325	29.90
120	78	16.82	4.254	0.8038	0.6587	8.761	30.81
180	78	17.31	3.844	0.7213	0.3137	10.62	29.21
Panel B: σ^{RE}							
30	79	12.51	3.693	1.4435	2.8899	6.812	25.88
60	79	12.69	3.400	1.0437	1.1558	6.985	23.09
120	79	12.87	3.194	1.0192	1.0739	8.038	22.56
180	79	13.04	3.132	0.8722	0.3246	8.254	21.57
Panel C: σ^{MF}							
30	78	15.68	4.450	1.5966	3.8305	8.589	34.61
60	79	16.11	3.964	1.1685	1.3999	10.16	28.72
120	78	17.51	3.976	0.7028	0.1716	10.06	28.24
180	78	17.33	3.566	0.4522	-0.1985	9.666	25.80
Panel D: $\ln(\sigma^{BS})$							
30	78	2.653	0.284	0.1563	0.2628	1.933	3.441
60	79	2.719	0.264	0.3500	-0.1444	2.119	3.397
120	78	2.792	0.246	0.1050	0.0401	2.170	3.427
180	78	2.828	0.216	0.1901	-0.3050	2.363	3.374
Panel E: $\ln(\sigma^{RE})$							
30	79	2.488	0.270	0.4957	0.5317	1.918	3.253
60	79	2.508	0.253	0.3368	-0.0222	1.939	3.139
120	79	2.527	0.235	0.3856	-0.0890	2.084	3.116
180	79	2.542	0.229	0.3646	-0.3687	2.110	3.071
Panel F: $\ln(\sigma^{MF})$							
30	78	2.718	0.255	0.6607	0.6865	2.150	3.544
60	79	2.752	0.229	0.5321	0.1377	2.318	3.357
120	78	2.838	0.221	0.1528	-0.2150	2.308	3.341
180	78	2.832	0.205	-0.0519	-0.2103	2.368	3.250

Note: σ^{BS} is the at-the-money B-S implied volatility.

Table 2
Correlation matrix of monthly 30-day volatility series ($N = 60$)

Panel A: Correlation matrix of volatility

	σ^{BS} (0.94)	σ^{BS} (0.97)	σ^{BS} (1.00)	σ^{BS} (1.03)	σ^{MF}	σ^{RE}
σ^{BS} (0.94)	1.000					
σ^{BS} (0.97)	0.879	1.000				
σ^{BS} (1.00)	0.861	0.887	1.000			
σ^{BS} (1.03)	0.864	0.872	0.898	1.000		
σ^{MF}	0.922	0.927	0.941	0.936	1.000	
σ^{RE}	0.755	0.761	0.811	0.784	0.854	1.000

Panel B: Correlation matrix of log volatility

	$\ln(\sigma^{BS})$ (0.94)	$\ln(\sigma^{BS})$ (0.97)	$\ln(\sigma^{BS})$ (1.00)	$\ln(\sigma^{BS})$ (1.03)	$\ln(\sigma^{MF})$	$\ln(\sigma^{RE})$
$\ln(\sigma^{BS})$ (0.94)	1.000					
$\ln(\sigma^{BS})$ (0.97)	0.847	1.000				
$\ln(\sigma^{BS})$ (1.00)	0.813	0.852	1.000			
$\ln(\sigma^{BS})$ (1.03)	0.828	0.829	0.872	1.000		
$\ln(\sigma^{MF})$	0.889	0.903	0.918	0.916	1.000	
$\ln(\sigma^{RE})$	0.705	0.745	0.767	0.778	0.850	1.000

Note: The σ^{BS} (1.06) is excluded in the calculation of correlation matrix as it has very few observations.

Table 3
Regression of model-free implied volatility on B-S implied volatility (OLS)

		σ_t			V_t			$\ln V_t$		
K/S	N	α	β^{BS}	R^2	α	β^{BS}	R^2	α	β^{BS}	R^2
Panel A: $\tau = 30$										
94%	65	-1.47	0.90	0.85	-22.1	0.74	0.89	-0.45	1.08	0.80
		(0.93)	(0.047)		(18.5)	(0.049)		(0.16)	(0.055)	
97%	77	0.20	0.94	0.87	-7.56	0.94	0.90	0.11	0.94	0.84
		(0.77)	(0.047)		(13.7)	(0.053)		(0.13)	(0.047)	
100%	78	1.67	0.95	0.88	9.92	1.07	0.91	0.54	0.82	0.82
		(0.81)	(0.055)		(16.4)	(0.076)		(0.15)	(0.057)	
103%	67	2.18	1.05	0.85	23.3	1.33	0.89	0.67	0.81	0.80
		(0.83)	(0.062)		(13.2)	(0.071)		(0.13)	(0.052)	
106%	15	0.86	1.22	0.84	-54.2	1.54	0.85	0.20	0.98	0.79
		(3.08)	(0.182)		(78.0)	(0.264)		(0.38)	(0.133)	
Panel B: $\tau = 60$										
94%	70	1.62	0.78	0.72	20.5	0.70	0.76	0.25	0.86	0.67
		(1.08)	(0.058)		(19.5)	(0.057)		(0.20)	(0.067)	
97%	78	2.56	0.82	0.75	37.8	0.81	0.74	0.54	0.80	0.73
		(1.01)	(0.064)		(20.5)	(0.083)		(0.17)	(0.060)	
100%	79	3.61	0.80	0.78	50.3	0.85	0.84	0.79	0.72	0.69
		(0.80)	(0.048)		(11.9)	(0.041)		(0.17)	(0.060)	
103%	77	4.58	0.84	0.74	68.2	1.01	0.78	1.04	0.66	0.68
		(0.87)	(0.063)		(16.0)	(0.087)		(0.15)	(0.058)	
106%	45	7.61	0.68	0.60	117.4	0.87	0.65	1.55	0.49	0.53
		(1.41)	(0.095)		(28.4)	(0.123)		(0.20)	(0.075)	
Panel C: $\tau = 120$										
94%	68	4.32	0.73	0.69	83.0	0.69	0.72	0.74	0.74	0.64
		(0.96)	(0.051)		(22.1)	(0.062)		(0.14)	(0.050)	
97%	77	3.61	0.82	0.68	62.9	0.85	0.70	0.71	0.76	0.64
		(1.20)	(0.070)		(25.6)	(0.090)		(0.20)	(0.071)	
100%	78	4.65	0.76	0.67	78.5	0.81	0.72	0.93	0.68	0.57
		(0.97)	(0.052)		(17.4)	(0.052)		(0.19)	(0.066)	
103%	77	4.95	0.84	0.69	90.4	0.96	0.71	1.00	0.69	0.63
		(1.06)	(0.067)		(21.3)	(0.092)		(0.19)	(0.069)	
106%	51	9.10	0.62	0.60	171.3	0.72	0.62	1.57	0.50	0.56
		(1.12)	(0.068)		(22.5)	(0.081)		(0.19)	(0.069)	
Panel D: $\tau = 180$										
94%	67	3.90	0.73	0.69	76.3	0.66	0.72	0.57	0.78	0.65
		(1.06)	(0.053)		(22.0)	(0.058)		(0.18)	(0.060)	
97%	72	3.55	0.80	0.66	68.1	0.78	0.66	0.61	0.79	0.63
		(1.07)	(0.058)		(18.5)	(0.056)		(0.21)	(0.070)	
100%	78	4.45	0.74	0.64	77.2	0.75	0.69	0.80	0.72	0.56
		(1.02)	(0.054)		(18.3)	(0.053)		(0.19)	(0.065)	
103%	75	4.68	0.80	0.64	88.2	0.86	0.66	0.85	0.73	0.60
		(1.07)	(0.066)		(20.7)	(0.080)		(0.20)	(0.071)	
106%	57	8.80	0.58	0.60	160.8	0.64	0.63	1.50	0.50	0.56
		(1.00)	(0.059)		(19.2)	(0.065)		(0.17)	(0.061)	

Note: The numbers in brackets are the standard errors of the parameter estimates, which are computed following a robust procedure taking into account of the heteroscedastic and autocorrelated error structure [see Newey and West (1987)]. The p-values of the χ^2 test for the joint hypothesis $H_0 : \alpha = 0, \beta = 1$ are all below 1%.

Table 4
Univariate and encompassing regressions of 30-day volatility (OLS)

N	α	β^{MF}	β^{BS}	β^{LRE}	adjusted R^2	Durbin-Watson	χ^2 test(a)	χ^2 test(b)
PANEL A: σ_t $(\tau = 30)$								
78	2.30		0.69		0.64	1.97	100.6	
	(0.94)		(0.066)				(0.000)	
78	1.51		0.60	0.20	0.68	2.11		37.61
	(1.07)		(0.082)	(0.063)				(0.000)
78	5.00			0.72^{**}	0.37	2.36	79.93	
	(1.45)			(0.153)			(0.000)	
78	1.11	0.73			0.76	1.97	257.8	
	(0.91)	(0.063)					(0.000)	
78	1.10	0.71	0.02^{+++}		0.76	1.96		34.87
	(0.92)	(0.109)	(0.113)					(0.000)
78	1.06	0.72		0.03^{+++}	0.76	2.00		34.91
	(0.98)	(0.078)		(0.093)				(0.000)
78	1.05	0.70^{*}	0.02^{+++}	0.02^{+++}	0.76	1.99		22.16
	(0.99)	(0.124)	(0.114)	(0.094)				(0.000)
PANEL B: V_t $(\tau = 30)$								
78	20.7		0.63		0.66	2.08	103.2	
	(14.5)		(0.071)				(0.000)	
78	17.2		0.61	0.08^{+++}	0.66	2.17		71.08
	(16.5)		(0.090)	(0.156)				(0.000)
78	82.6			0.73^{***}	0.27	2.47	51.62	
	(23.1)			(0.241)			(0.000)	
78	19.0	0.57			0.75	1.94	244.9	
	(16.7)	(0.076)					(0.000)	
78	16.6	0.41	0.19^{+++}		0.75	2.05		138.1
	(16.5)	(0.135)	(0.165)					(0.000)
78	21.1	0.58		-0.05^{+++}	0.75	2.07		133.6
	(17.9)	(0.093)		(0.154)				(0.000)
78	18.1	0.42	0.18^{+++}	-0.03^{+++}	0.75	2.01		60.00
	(17.9)	(0.142)	(0.165)	(0.156)				(0.000)
PANEL C: $\ln V_t$ $(\tau = 30)$								
78	0.52		0.74		0.61	1.85	96.32	
	(0.19)		(0.072)				(0.000)	
78	0.33		0.58	0.27	0.65	2.06		26.04
	(0.19)		(0.082)	(0.067)				(0.000)
78	0.99			0.65	0.43	2.19	102.0	
	(0.21)			(0.094)			(0.000)	
78	-0.03	0.93^{***}			0.77	1.95	249.6	
	(0.15)	(0.057)					(0.000)	
78	-0.04	1.02^{***}	-0.09^{+++}		0.77	1.92		2.078
	(0.16)	(0.112)	(0.101)					(0.353)
78	-0.04	0.90^{***}		0.05^{+++}	0.77	2.03		1.866
	(0.16)	(0.076)		(0.062)				(0.393)
78	-0.05	0.98^{***}	-0.08^{+++}	0.04^{+++}	0.76	2.01		2.961
	(0.16)	(0.136)	(0.103)	(0.063)				(0.397)

Note: The at-the-money B-S implied volatility and past τ-day realized volatility are used in the regressions. The numbers in brackets are the standard errors of the parameter estimates, which are computed following a robust procedure taking into account of the heteroscedasticity [see White (1980)]. The χ^2 test (a) is for the joint hypothesis $H_0 : \alpha = 0$ & $\beta^j = 1$ $(j = MF, BS, LRE)$ in univariate regressions, and the χ^2 test (b) is for the joint hypothesis $H_0 : \beta^{BS} = 1$ & $\beta^{LRE} = 0$ or $H_0 : \beta^{MF} = 1$ & $\beta^{BS} = \beta^{LRE} = 0$ in encompassing regressions. The test statistics are reported with the p-values in the brackets beneath. The ***, ** and * indicate β insignificantly different from one at the 10%, 5% and 1% level, and $^{+++}$, $^{++}$ and $^{+}$ indicate β insignificantly different from zero at the 10%, 5% and 1% level, respectively.

Table 5
Univariate and encompassing regressions of 30-day log volatility (IV)

Panel A: First stage regression

N	α	β^{LBS} or β^{LMF}	β^{LRE}	β^{BS}	adjusted R^2	Durbin-Watson
Dependent variable: $\ln V_t^{BS}$ ($\tau = 30$)						
77	1.35	0.49			0.23	2.34
	(0.22)	(0.080)				
77	0.86	0.21	0.53		0.41	2.61
	(0.21)	(0.111)	(0.156)			
Dependent variable: $\ln V_t^{MF}$ ($\tau = 30$)						
77	0.69	0.75			0.55	2.25
	(0.16)	(0.058)				
77	0.43	0.50	0.41		0.68	2.74
	(0.18)	(0.079)	(0.105)			
77	0.16	0.23	0.17	0.58	0.89	2.06
	(0.12)	(0.060)	(0.051)	(0.076)		

Panel B: Second stage IV estimates

N	α	β^{MF}	β^{BS}	β^{LRE}	adj. R^2	DW	χ^2 test(a)	χ^2 test(b)
Instrument variable for the B-S implied volatility ($\tau = 30$)								
77	-0.39		1.09^{***}		0.30	1.90	42.55	
	(0.48)		(0.178)				(0.000)	
77	-0.49		1.29^{***}	-0.20^{+++}	0.48	2.44		2.440
	(0.62)		(0.296)	(0.237)				(0.295)
Instrument variables for both model-free and B-S implied volatility ($\tau = 30$)								
77	-0.09	0.95^{***}			0.63	2.11	330.2	
	(0.31)	(0.111)					(0.000)	
77	-0.12	0.87^{***}	0.10^{+++}		0.69	2.11		1.449
	(0.20)	(0.112)	(0.129)					(0.484)
77	-0.09	0.93^{***}		0.02^{+++}	0.67	2.12		0.234
	(0.24)	(0.160)		(0.114)				(0.889)
77	-0.08	0.87^{***}	0.06^{+++}	0.02^{+++}	0.69	2.11		1.322
	(0.47)	(0.118)	(0.341)	(0.254)				(0.724)

Note: The at-the-money B-S implied volatility and past τ-day realized volatility are used in the regressions. The standard errors of the IV estimates are computed following a robust procedure taking into account of the heteroscedastic and autocorrelated error structure [see Newey and West (1987)]. The χ^2 test (a) is for the joint hypothesis $H_0 : \alpha = 0$ & $\beta^j = 1$ ($j = MF, BS, LRE$) in univariate regressions, and the χ^2 test (b) is for the joint hypothesis $H_0 : \beta^{BS} = 1$ & $\beta^{LRE} = 0$ or $H_0 : \beta^{MF} = 1$ & $\beta^{BS} = \beta^{LRE} = 0$ in encompassing regressions. The test statistics are reported with the p-values in the brackets beneath. The ***, ** and * indicate β insignificantly different from one at the 10%, 5% and 1% level, and $^{+++}$, $^{++}$ and $^{+}$ indicate β insignificantly different from zero at the 10%, 5% and 1% level, respectively.

Table 6
Univariate and encompassing regressions of 60-, 120- and 180-day log volatility (OLS)

N	α	β^{MF}	β^{BS}	β^{LRE}	adjusted R^2	Durbin-Watson	χ^2 test(a)	χ^2 test(b)
Panel A: $\ln V_t$ ($\tau = 60$)								
79	0.74		0.71		0.54	1.74	30.78	
	(0.23)		(0.080)				(0.000)	
79	0.47		0.52	0.33	0.61	1.73		34.21
	(0.18)		(0.073)	(0.083)				(0.000)
79	1.18			0.64	0.42	1.34	337.2	
	(0.22)			(0.101)			(0.000)	
79	-0.01	0.90^{***}			0.72	1.73	33.21	
	(0.22)	(0.085)					(0.000)	
79	0.01	0.78^{***}	0.13^{+++}		0.73	1.81		4.428
	(0.22)	(0.142)	(0.106)					(0.109)
79	-0.01	0.87^{**}		0.06^{+++}	0.73	1.69		5.287
	(0.20)	(0.078)		(0.071)				(0.071)
79	-0.02	0.75^{**}	0.13^{+++}	0.07^{+++}	0.72	1.76		7.964
	(0.20)	(0.132)	(0.103)	(0.069)				(0.047)
Panel B: $\ln V_t$ ($\tau = 120$)								
78	0.62		0.68		0.51	1.40	254.5	
	(0.29)		(0.104)				(0.000)	
78	0.30		0.49	0.37	0.60	1.18		40.03
	(0.22)		(0.101)	(0.122)				(0.000)
78	1.01			0.65^{*}	0.42	0.74	186.1	
	(0.30)			(0.138)			(0.000)	
78	-0.05	0.91^{***}			0.67	1.19	514.1	
	(0.24)	(0.089)					(0.000)	
78	-0.10	0.79^{**}	0.14^{++}		0.68	1.24		5.117
	(0.23)	(0.119)	(0.084)					(0.077)
78	-0.07	0.83^{**}		0.09^{+++}	0.68	1.10		4.042
	(0.24)	(0.101)		(0.082)				(0.132)
78	-0.12	0.76^{*}	0.14^{++}	0.08^{+++}	0.68	1.16		7.211
	(0.23)	(0.121)	(0.082)	(0.083)				(0.065)
Panel B: $\ln V_t$ ($\tau = 180$)								
78	0.48		0.72		0.46	1.17	277.2	
	(0.27)		(0.095)				(0.000)	
78	0.23		0.50	0.38	0.55	0.87		28.00
	(0.27)		(0.081)	(0.113)				(0.000)
78	1.02			0.65^{*}	0.41	0.52	219.3	
	(0.33)			(0.148)			(0.000)	
78	-0.13	0.93^{***}			0.69	1.05	437.8	
	(0.22)	(0.082)					(0.000)	
78	-0.20	0.85^{***}	0.12^{++}		0.69	1.02		2.929
	(0.22)	(0.110)	(0.072)					(0.231)
78	-0.14	0.87^{***}		0.09^{+++}	0.69	0.97		2.794
	(0.24)	(0.080)		(0.077)				(0.247)
78	-0.20	0.78^{*}	0.11^{+++}	0.09^{+++}	0.69	0.95		5.920
	(0.23)	(0.104)	(0.071)	(0.075)				(0.115)

Note: The numbers in brackets are the standard errors of the parameter estimates, which are computed following a robust procedure taking into account of the heteroscedastic and autocorrelated error structure [see Newey and West (1987)]. The χ^2 test (a) is for the joint hypothesis $H_0 : \alpha = 0$ & $\beta^j = 1$ ($j = MF, BS, LRE$) in univariate regressions, and the χ^2 test (b) is for the joint hypothesis $H_0 : \beta^{BS} = 1$ & $\beta^{LRE} = 0$ or $H_0 : \beta^{MF} = 1$ & $\beta^{BS} = \beta^{LRE} = 0$ in encompassing regressions. The test statistics are reported with the p-values in the brackets beneath. The ***, ** and * indicate β insignificantly different from one at the 10%, 5% and 1% level, and $^{+++}$, $^{++}$ and $^{+}$ indicate β insignificantly different from zero at the 10%, 5% and 1% level, respectively.

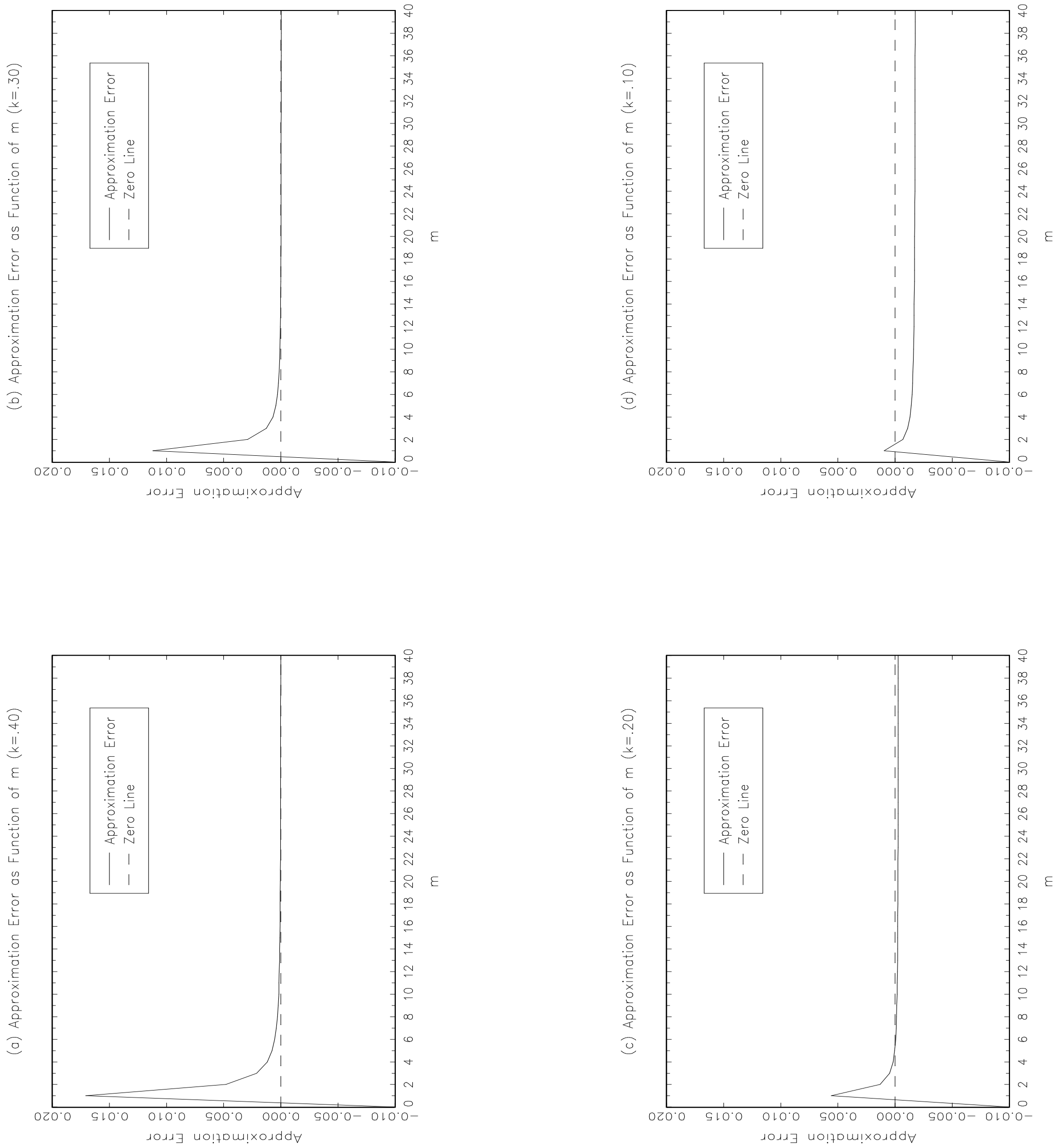

Figure 1
Approximation error of the model-free implied volatility from the SV option prices