Journal of Financial Economics 50 (1998) 125–150

The relation between implied and realized volatility[1]

B.J. Christensen[a], N.R. Prabhala[b,*]

[a] School of Economics and Management, University of Aarhus, DK-8000 Aarhus C, Denmark
[b] Yale School of Management, Yale University, New Haven, CT 06511, USA

Received 10 October 1996; received in revised form 29 December 1997

Abstract

Previous research finds the volatility implied by S&P 100 index option prices to be a biased and inefficient forecast of future volatility and to contain little or no incremental information beyond that in past realized volatility. In contrast, we find that implied volatility outperforms past volatility in forecasting future volatility and even subsumes the information content of past volatility in some of our specifications. Our results differ from previous studies because we use longer time series and nonoverlapping data. A regime shift around the October 1987 crash explains why implied volatility is more biased in previous work.

JEL classification: G13; G14; C53

Keywords: Options; Volatility; Stock market crash; Information; Market efficiency

1. Introduction

The volatility implied in an option's price is widely regarded as the option market's forecast of future return volatility over the remaining life of the relevant

* Corresponding author. Tel.: 203/432-6030; fax: 203/432-3003; e-mail: nagpurnanand.prabhala@yale.edu.

[1] We thank Interactive Data Corporation for providing the option price data used in this study, and David Bates, N.K. Chidambaran, Young–Ho Eom, Steven Feinstein, Stephen Figlewski, Ken French, Will Goetzmann, Bob Jarrow, Nikunj Kapadia, Cheng–Few Lee, K. Geert Rouwenhorst, Chris Sims, Suresh Sundaresan, Zhen Yu Wang, seminar participants at the American Finance Association, Boston University, European Finance Association, Financial Management Association, and the Cornell–Queens Derivative Securities Conference for useful comments. We are also grateful to G. William Schwert (the editor) and an anonymous referee for very extensive and helpful feedback.

PII: S0304-405X(97)00034-8

option. If option markets are efficient, implied volatility should be an efficient forecast of future volatility, i.e., implied volatility should subsume the information contained in all other variables in the market information set in explaining future volatility. Implied volatility is interpreted as an efficient volatility forecast in a wide range of settings (e.g., Day and Lewis, 1988; Harvey and Whaley, 1992; Poterba and Summers, 1986; Sheikh, 1989).

Whether implied volatility predicts future volatility and whether it does so efficiently are both empirically testable propositions. Such questions have been the subject of several papers. Early papers (e.g., Latane and Rendleman, 1976) focus on static cross-sectional tests. These papers essentially document that stocks with higher implied volatilities also have higher ex-post realized volatility. With the availability of sufficient time series data, however, more recent tests have focused on the information content of implied volatility in dynamic settings. Such studies examine whether the implied volatility of an option predicts the ex-post realized volatility over the remaining life of the option.

The time series literature has produced mixed results. At one end of the spectrum, Jorion (1995) reports that implied volatility is an efficient (though biased) predictor of future return volatility for foreign currency futures. In contrast, Day and Lewis (1992), who study S&P 100 index options with expiries from 1985–1989, and Lamoureux and Lastrapes (1993), who examine options on ten stocks with expiries from 1982 to 1984, conclude that implied volatility is biased and inefficient: past volatility contains predictive information about future volatility beyond that contained in implied volatility. Both studies use overlapping samples, and, additionally, are characterized by a 'maturity mismatch' problem, in that Lamoureux and Lastrapes examine one-day-ahead and Day and Lewis examine one-week-ahead predictive power of implied volatilities computed from options that have a much longer remaining life (up to 129 trading days in the former and 36 trading days in the latter). Thus, their results are hard to interpret. Most striking, however, is the evidence reported by Canina and Figlewski (1993) – henceforth CF – who find that '. . . implied volatility has virtually no correlation with future return volatility. . .' and does not incorporate information contained in recent observed volatility.' These results are surprising, in part because they are so extreme, and more so because the evidence pertains to the most active options market in the U.S. – the market for S&P 100 index (OEX) options.

One explanation for the CF results is that index option markets process volatility information inefficiently. However, this explanation is unlikely given the liquidity, depth, and trading activity in the OEX options market. A second possibility, suggested by CF, is that the Black and Scholes (1973) option pricing model, which is used to compute implied volatility, cannot be used to price index options because of prohibitive transaction costs associated with hedging of options in the cash index market. This explanation is also incomplete, though. The Black–Scholes formula does not necessarily require continuous trading in cash markets. One-period equilibrium representative agent models, of the sort analyzed in Brennan (1979) and Rubinstein (1976), also lead to the Black–Scholes formula. Further, Constantinides (1994) shows that transaction costs do not have a first-order effect on option prices. Hence, transaction costs could not, in isolation, explain the apparent failure of the Black–Scholes pricing model for the OEX options market.

Finally, while the Black–Scholes implied volatility can be thought of as a volatility forecast, it can also be interpreted as a measure of an option's *price* – one that controls for option-specific characteristics such as the moneyness of an option, time to expiry, etc. All option pricing theory of which we are aware implies that option prices should be positively correlated with the underlying asset's volatility; this is established, most recently, in Theorem 6 of Bergman et al. (1996). Thus, CF's finding of no significant relation between an option's price (implied volatility) and future realized volatility appears to refute this basic tenet of option pricing theory.

We reexamine the relation between implied volatility and subsequent realized volatility for the OEX options market. Our study differs from previous work along two primary dimensions. First, we use volatility data sampled over a longer period of time than in previous studies. This increases statistical power and allows for evolution in the efficiency of the market for OEX index options since their introduction in 1983. A second difference relative to previous work is that we sample the implied and realized volatility series at a *lower* (monthly) frequency. This enables us to construct volatility series with nonoverlapping data with exactly one implied and one realized volatility covering each time period in our sample. Our nonoverlapping sample yields more reliable regression estimates relative to less precise and potentially inconsistent estimates obtained from overlapping samples used in previous work.

We find that implied volatility is a less biased forecast of future volatility than reported in previous studies. This difference could be attributed to our use of longer volatility time series. Our data span the $11\frac{1}{2}$ year period between November 1983 and May 1995, while the CF findings are based on a shorter time period preceding the October 1987 crash. We document that there was a regime shift around the crash period. Implied volatility is more biased before the crash than after, at least partially due to a poor signal-to-noise ratio prevalent prior to the crash, and perhaps also due to learning by market participants in the wake of the crash.

We also find that past volatility has much less explanatory power than reported in CF. In fact, it has *no* incremental explanatory power over implied volatility in some of our specifications. We attribute this difference to differences in sampling procedures. While our sampling procedure yields nonoverlapping data, the CF procedure results in an extreme degree of overlap in consecutive observations in the time series of past and future volatility as well as highly autocorrelated errors with a low Durbin–Watson statistic (of the order of 0.2) and, hence, imprecise and potentially inconsistent regression estimates. In fact, applying our sampling procedure to a pre-crash subperiod similar to that used by CF, we find that implied volatility is an efficient, albeit biased, volatility forecast. Thus, the apparent inefficiency of implied volatility reported by CF seems to be an artifact of their overlapping sampling method.

The paper is organized as follows. Section 2 describes how our volatility series are constructed and provides descriptive statistics for these series. Section 3 presents the empirical results, and Section 4 discusses why our results are different from those reported in earlier literature. Section 5 concludes.

2. Data and sampling procedure

Our empirical analysis focuses on S&P 100 index options. Exchange-traded OEX options with one-month expiration cycles became available in November 1983. Our monthly implied and realized volatility series begin at that time and end in May 1995. Thus, the data span a time period of 139 months, or about $11\frac{1}{2}$ years.

2.1. Sampling procedure

By convention, OEX options expire on the third Saturday of every month. We move to the Wednesday that immediately follows the expiry, and record the OEX level – say, S_t – on this date. On the same date, we locate a call option expiring the next month that is closest to being at-the-money. We record the price, C_t, of this call, as well as its strike price, K_t. This option expires on the third Saturday of the following month $t+1$; the next $(t+1)$ call option is sampled on the Wednesday that immediately follows. An entire sequence of option prices is constructed in this manner. The key feature of this sampling procedure is that it results in nonoverlapping data, as time periods covered by successive options exhibit no overlap whatsoever.

2.2. Variable definitions

From each observed call price C_t, we compute implied volatility σ_{it} by numerically solving the Black–Scholes call option pricing formula, i.e.,

$$C_t = S_t N(d_t) - K_t e^{-r_{f,t}\tau_t} N(d_t - \sigma_{it}\sqrt{\tau_t}), \tag{1}$$

where $d_t = [\,\log(S_t/K_t) + (r_{f,t} + \sigma_{it}^2/2)\tau_t]/\sigma_{it}\sqrt{\tau_t}$, τ_t denotes the time to expiration, $r_{f,t}$ stands for the interest rate, and $N(\cdot)$ denotes the standard normal distribution function. For the interest rate, we use the one-month LIBOR (the inter-bank borrowing rate in the Eurodollar market) as this is probably close to the rate faced by option traders. The implied volatility estimate obtained by solving formula (1) is based on options with about 24 days (18 trading days) to expiration.

While implied volatility represents an ex-ante volatility forecast, we also compute the ex-post return volatility over each option's life. This is computed as the sample standard deviation of the daily index returns over the remaining life of the option. That is,

$$\sigma_{ht} = \sqrt{\frac{1}{\tau_t}\sum_{k=1}^{\tau t}(r_{t,k} - \bar{r}_t)^2}, \tag{2}$$

where τ_t is the number of days to expiration, $\bar{r}_t = \tau_t^{-1}\sum_{k=1}^{\tau_t} r_{t,k}$, and $r_{t,k}$ is the index return on day k of month t. Both volatility measures are expressed in annual terms to facilitate interpretation. Finally, much of our empirical work is based on the log-volatility series, which we denote by $i_t = \log \sigma_{it}$ and $h_t = \log \sigma_{ht}$ (we use log to denote natural logarithm throughout). All data are from the financial databases of Interactive Data Corporation.

2.3. Descriptive statistics

Table 1 presents descriptive statistics for the volatility and log-volatility series. Statistics for the entire sample period (November 1983 to May 1995) are presented in Panel A. Information for two subperiods surrounding the October 1987 stock market crash are provided in Panels B & C.

Starting with the means, we find that both average implied volatility and average log implied volatility exceed the means of the corresponding realized volatility series. The mean difference between the two log-volatility series is greater than the mean difference between the level series. For instance, in the post-crash period, mean implied volatility and mean realized volatility differ by 0.01, while mean log implied volatility and mean log realized volatility differ by 0.12. This difference can be reconciled by observing that if the means of two lognormal series are equal, then the more volatile series should have the lower mean after taking logarithms.[2]

[2] Specifically, the mean of log implied volatility should be $\log \mu_i - \frac{1}{2}\sigma_i^2$ and that of log realized volatility should be $\log \mu_h - \frac{1}{2}\sigma_h^2$, where σ_i^2 and σ_h^2 denote the variances of $\log \sigma_{it}$ and $\log \sigma_{ht}$ and the μ's are the means of the level series. For instance, in the post-crash subperiod, mean implied and realized volatility are 0.14 and 0.13, respectively, while the variances of the log-volatility series are $\sigma_i^2 = 0.0891$ and $\sigma_h^2 = 0.1514$. Thus, log implied volatility should have a mean of $\log(0.14) - 0.5 \times 0.0891 = -2.01$ and log realized volatility a mean of $\log(0.13) - 0.5 \times 0.1514 = -2.12$, and these are virtually identical to the observed means of the log-volatility series. Similar computations reconcile the difference in the log means for full period, and half the difference for the pre-crash subperiod.

Table 1
Descriptive statistics

Statistic	Implied volatility	Realized volatility	Log implied volatility	Log realized volatility
Panel A: Full period – 11/1983 to 05/1995				
Mean	0.15	0.14	– 1.98	– 2.07
100*Variance	0.28	0.68	10.34	15.21
Skewness	2.48	5.64	0.41	1.04
Kurtosis	16.45	47.61	3.71	7.17
Panel B: Subperiod 11/1983 to 09/1987				
Mean	0.15	0.14	– 1.95	– 2.02
100*Variance	0.18	0.10	9.49	5.11
Skewness	0.04	0.59	– 0.50	– 0.10
Kurtosis	2.17	3.43	2.40	3.06
Panel C: Subperiod 12/1987 to 05/1995				
Mean	0.14	0.13	– 2.00	– 2.12
100*Variance	0.19	0.35	8.91	15.14
Skewness	0.91	2.17	0.29	0.48
Kurtosis	3.34	9.33	2.42	3.88

Descriptive statistics for monthly time series of levels natural logarithms of implied volatility and realized volatility for the S&P 100 index. Here, implied volatility is computed each month using the Black–Scholes formula for an at-the-money OEX call option which expires on the third Saturday of the month and has 18 trading days to expiration. Realized volatility is the annualized ex-post daily return volatility (sample standard deviation) of the index over the life of the option. Statistics for the entire sample (Panel A) are based on 139 nonoverlapping monthly observations on each volatility series, covering the period November 1983 to May 1995. We also report statistics for two subperiods that exclude the October 1987 stock market crash – the 47 months from November 1983 to September 1987 (Panel B) and the 90 months from December 1987 to May 1995 (Panel C).

The data also reveal interesting patterns in the variances of the volatility series across the two subperiods. Realized volatility is more variable in the post-crash period relative to the pre-crash period, e.g., the variance of realized volatility in the pre-crash period is only about 29% of that in the post-crash period. More strikingly, implied volatility is *more* variable than realized volatility in the pre-crash subperiod. This is at odds with the notion that implied volatility is a smoothed expectation of realized volatility, in which case it should be *less* variable than realized volatility. We examine this issue further in Section 4 of the paper.

Finally, for the full period, the distributions of both implied and realized volatility are highly skewed and leptokurtic, whereas the distributions of the log-volatility series are less so. Skewness and kurtosis are considerably lower in the individual subperiods relative to the whole period. This feature reflects the rather extreme nature of the October 1987 stock market crash, which is omitted from the subperiods.

To assess the time series properties of the two series, we fit ARIMA(p, d, q) models of the form

$$\Phi(B)(\Delta^d x_t - \mu) = \Theta(B)\varepsilon_t \tag{3}$$

where x_t represents one of the two log-volatility series, as in French et al. (1987) or Schwert (1987). In Eq. (3), parameter μ is the mean, ε_t is white noise, Φ and Θ are polynomials of order p and q in B, the backshift operator defined by $Bx_t = x_{t-1}$, and $\Delta = 1 - B$ is the first-difference operator. The time series models are fit to the log-volatility series rather than the level series, since the former conform more closely to normality (see also French et al., 1987; Schwert, 1987, 1989).

Table 2 displays results for both series. For the nonintegrated AR(1), AR(2), and ARMA(1, 1) models, the Box–Pierce (1970) portmanteau statistics suggest that the nonintegrated ARMA(1, 1) process appears to best describe both volatility series. Since the first-order autocorrelation is close to one, we also fit an integrated ARIMA(1, 1, 1) model to the log-volatility series. The moving average coefficient is larger and the Box–Pierce statistic is actually worse for the integrated model than for the nonintegrated model. The results are similar to those obtained in previous work with monthly volatility series (e.g., French et al., 1987) and indicate that from a time series perspective, our volatility series do not appear to be unusual relative to similar series used in previous research.

2.4. Measurement error in implied volatility

Our estimate of implied volatility is afflicted with several sources of measurement error. First, the Black–Scholes formula in Eq. (1) applies to a European style call option on an asset that is known in advance to pay no dividends prior to expiration of the option. OEX options are American-style and thus can be exercised early, and the underlying asset – the S&P 100 index – pays dividends. Early exercise precipitated by dividend payments causes few problems, since OEX dividends are small, smooth, and not concentrated on any one day. Thus, the early exercise premium due to dividends is likely to be small for call options. However, dividends do reduce call values, so implied volatility computed via Eq. (1) understates true implied volatility. The difference should be roughly constant for all time periods since dividends are relatively uniform for the OEX. Thus, in regressions that use implied volatility as an independent variable, estimates of the intercept term should be biased (upwards, if the regression

Table 2
ARIMA(p, d, q) models for implied and realized volatility

Fitted model	μ	ϕ_1	ϕ_2	θ_1	Box–Pierce statistic Q_{12}	Degrees of freedom
Panel A: Implied volatility $\{i_t\}$						
AR(1)	− 1.99[a]	0.70[a]			14.69	11
	(− 33.40)	(11.37)				
ARMA(1, 1)	− 1.99[a]	0.88[a]		0.40[a]	4.19	10
	(− 25.08)	(16.60)		(3.25)		
AR(2)	− 2.00[a]	0.53[a]	0.23[a]		6.62	10
	(− 21.47)	(6.28)	(2.74)			
ARIMA(1, 1, 1)	− 0.00	0.36[a]		0.80[a]	5.77	10
	(− 0.16)	(2.81)		(9.66)		
Panel B: Realized volatility $\{h_t\}$						
AR(1)	− 2.07[a]	0.56[a]			27.95[a]	11
	(− 33.40)	(8.03)				
ARMA(1, 1)	− 2.07[a]	0.90[a]		0.56[a]	8.98	10
	(− 33.40)	(16.34)		(3.44)		
AR(2)	− 2.09[a]	0.40[a]	0.28[a]		13.73	10
	(− 25.40)	(4.91)	(3.44)			
ARIMA(1, 1, 1)	− 0.00	0.20		0.82[a]	11.86	10
	(− 0.49)	(1.72)		(12.11)		

[a] p-value < 0.01.
[b] p-value $= 0.05$.
Estimates of ARIMA(p, d, q) models from specification (3) in the paper of the form

$$\Phi(B)(\Delta^d x_t - \mu) = \Theta(B)\varepsilon_t,$$

fitted to the time series $\{x_t\}$, with $x_t = i_t$ (Panel A) or $x_t = h_t$ (Panel B), where i_t denotes the natural logarithm of the Black–Scholes implied volatility for at-the-money call options on the S&P 100 index, h_t denotes the natural logarithm of the ex-post daily return volatility of the index, ε_t is white noise, $\Phi(B)$ denotes the AR polynomial $1 - \phi_1 B - \phi_2 B^2$, $\Theta(B)$ denotes the MA polynomial $1 - \theta_1 B$, B denotes the backshift operator, and $\Delta = 1 - B$ denotes the first-difference operator. The data consist of 139 nonoverlapping monthly observations on each volatility series, covering the period November 1983 to May 1995. Numbers in parentheses denote asymptotic t-statistics.

coefficient for implied volatility is positive) but the slope coefficient should not be affected.[3]

Second, option prices and closing OEX index levels may be nonsynchronous, either because they are simply recorded at different times or because closing prices of at least some of the one hundred stocks that constitute the OEX are stale. Assuming such random errors in index prices to be 0.25%, Jorion (1995) estimates that the error in measured implied volatility is about 1.2%.

Additional measurement error is introduced by bid-ask spreads in option prices and the 'wild-card' option embedded in OEX options, particularly short-maturity options (Harvey and Whaley, 1992). To some extent, we minimize these sources of error, since we exclude deep out-of-the-money options and options with very short maturities. Nevertheless, some measurement error will remain and we account for this in our empirical analysis.

Finally, we note that the Black–Scholes formula in Eq. (1) assumes that index levels follow a log-normal diffusion process with deterministic volatility. If this assumption is not satisfied – for instance, because of jumps in index prices – several papers (e.g., Cox and Ross, 1976; Heston, 1993) show that the Black–Scholes formula is misspecified, which would exacerbate the errors-in-variable problem if Black–Scholes implied volatility is used as a volatility forecast. Even so, a study of the relation between Black–Scholes implied and realized volatility is meaningful, but must be interpreted differently. Following the discussion in Section 1, our study can be regarded simply as a test of whether option prices are informative about future return volatility, without necessarily drawing any inferences about the efficiency of option markets.

3. The relation between implied and realized volatility

In this section, we analyze the information content of implied volatility via several different specifications. We show that implied volatility contains incremental information beyond that in past volatility. Similar conclusions are reached by all approaches, underscoring the robustness of our findings.

[3] The difference between true implied volatility and that estimated by ignoring dividends is roughly constant, provided the true volatility is not too small. For instance, for at-the-money options with 25-day maturities, the difference lies between 1.94% and 2% if true implied volatility lies between 10% and 50% and annualized dividend yields are about 4% per annum. We also replicate our study for a period beginning in 1983 and ending in 1992 using dividend-adjusted implied volatilities instead of unadjusted volatilities, with no material change in the results. Campbell Harvey kindly provided us the dividend data.

3.1. Conventional analysis

The information content of implied volatility is typically assessed in the literature by estimating a regression of the form

$$h_t = \alpha_0 + \alpha_i i_t + e_t, \tag{4}$$

where h_t denotes the realized volatility for period t and i_t denotes the implied volatility at the beginning of period t, as defined in Section 2.2.

At least three hypotheses can be tested using Eq. (4). First, if implied volatility contains *some* information about future volatility, α_i should be nonzero. Second, if implied volatility is an *unbiased* forecast of realized volatility, we should find that $\alpha_0 = 0$ and $\alpha_i = 1$. Finally, if implied volatility is *efficient*, the residuals e_t should be white noise and uncorrelated with any variable in the market's information set.

Ordinary least-squares estimates of Eq. (4) are reported in the first row of Table 3. The estimate of α_i is 0.76 and is statistically significant against a null of $\alpha_i = 0$. Hence, implied volatility contains some information about future volatility. However, it appears to be a biased forecast of future volatility since α_i is reliably different from unity and the intercept α_0 is different from zero. An F-test rejects the joint hypothesis $\alpha_0 = 0$ and $\alpha_i = 1$; the F (2, 136) statistic of 10.58 is significant at 1%.

That α_i is less than unity could be a consequence of an errors-in-variable (EIV) problem associated with implied volatility, an issue to which we will return soon. The fact that the intercept is negative is a consequence of the EIV problem and our use of the log-volatility series rather than the volatility level series to estimate the regression. If $\bar{h}_t$ and $\bar{i}_t$ denote the means of the two log-volatility series, then following Section 2.3, $\bar{h}_t < \bar{i}_t < 0$; additionally, the regression estimate of α_i is less than one. These inequalities lead to a negative intercept α_0, since this must satisfy the relation $\alpha_0 = \bar{h}_t - \alpha_i \bar{i}_t$. Consistent with this proposition, the intercept is smaller (0.03) and not significantly different from zero when Eq. (4) is estimated using the volatility level series. The Durbin–Watson statistic is not significantly different from two, indicating that the residuals from Eq. (4) are not autocorrelated.

How does the information content of implied volatility compare to that of past realized volatility? To address this question, we estimate the following multiple regression (the 'encompassing' specification in the nomenclature of the forecasting literature):

$$h_t = \alpha_0 + \alpha_i i_t + \alpha_h h_{t-1} + e_t. \tag{5}$$

OLS estimates of Eq. (5) are reported in Table 3. Past realized volatility, in isolation, explains future volatility (see the second row of Table 3). However, once implied volatility is added as an explanatory variable (third row of

Table 3
Information content of implied volatility: OLS estimates

Dependent variable: Log realized volatility h_t

Independent variables			Adj. R^2	DW
Intercept	i_t	h_{t-1}		
− 0.56[a]	0.76[a]		39%	1.89
(− 3.47)	(9.48)			
− 0.89[a]		0.57[a]	32%	2.32
(− 6.02)		(8.04)		
− 0.49[a]	0.56[a]	0.23[b]	41%	2.23
(− 3.01)	(4.78)	(2.38)		

[a] p-value < 0.01
[b] p-value $= 0.05$
OLS estimates of specification (5) in the paper,

$$h_t = \alpha_0 + \alpha_i i_t + \alpha_h h_{t-1} + e_t.$$

Here, i_t denotes the natural logarithm of the Black–Scholes implied volatility for at-the-money call options on the S&P 100 index, measured at the beginning of month t, and h_t denotes the natural logarithm of the ex-post daily return volatility of the index, over the life of the option whose log implied volatility is i_t. The data consist of 139 nonoverlapping monthly observations of each volatility series covering the period November 1983 to May 1995. Numbers in parentheses denote asymptotic t-statistics.

Table 3), the regression coefficient for past volatility (α_h) drops from 0.57 to 0.23, but it remains significant, suggesting that implied volatility is inefficient; the F (2, 135) statistic of 9.18 rejects the joint hypothesis $\alpha_0 = 0$, $\alpha_i = 1$, and $\alpha_h = 0$ at 1%. Nevertheless, the slope coefficient for implied volatility itself remains significant in the multiple regression and, in particular, is more than twice the coefficient for past volatility. Adding more lags and correcting standard errors along the lines of Newey and West (1987) does not alter these results.

Even these basic results differ from those reported in Canina and Figlewski. For Eq. (4), the CF estimates of α_i range from 0.14 to 0.22, whereas ours (0.76) is more than three times larger. In the multiple regression (5), the CF estimate of α_h is about 0.50; our largest estimate (0.23) is much smaller. For the same specification, the CF estimates of α_i range from 0.04 to 0.08 and are never significant. In contrast, our estimate (0.56) is an order of magnitude larger and is statistically significant.

To this point, our results indicate that OEX implied volatility is neither unbiased nor efficient. Nevertheless, implied volatility has more predictive power than past volatility, whether judged by the magnitude of the regression slope coefficients or by the R^2 of each of the regressions discussed above. At the very least, implied volatility contains information about future volatility beyond that contained in past volatility. In this respect, OEX index options do not appear to be structurally different from currency options (Jorion, 1995) or options on individual stocks (Lamoureux and Lastrapes, 1993).

3.2. Alternative specifications

In this section, we examine an alternative specification of the form

$$i_t = \beta_0 + \beta_i i_{t-1} + \beta_h h_{t-1} + e_{it}, \tag{6}$$

for at least two reasons. First, we use specification (6) in an instrumental variables framework to correct for EIV problems in OEX implied volatility. Second, we use it to test whether implied volatility is predicted by past volatility. If option prices reflect volatility information, implied volatility should not only predict future volatility but should also itself endogenously depend on past volatility, since past and future volatility are positively related. We test this implication using Eq. (6). Specification (6) is similar to a GARCH (1,1) specification (Bollerslev, 1986), with i_t playing the role of the underlying volatility series and h_{t-1} that of the lagged squared return in the standard GARCH model. However, the existence of the error term e_{it} moves Eq. (6) closer to a stochastic volatility model rather than a GARCH model (Nelson, 1990).

We characterize the effect of the EIV problem in estimated implied volatility on the conventional OLS estimators used in Section 3.1. We then reexamine the information content of implied volatility using alternative instrumental variable estimators of Eqs. (4) and (5). We start by establishing that the EIV problem causes implied volatility to appear both biased and inefficient. Proposition 1 describes the relevant results, based on standard least squares asymptotics (see, e.g., Greene, 1993, Chapter 9).

Proposition 1. *Suppose that observed implied volatility $i_t = i_t^* + \eta_t$, where i_t^* denotes true implied volatility and η_t denotes measurement error (uncorrelated with i_t^*). Then the probability limit of the OLS estimate of α_i in Eq.* (4) *is given by*

$$\alpha_{i,u}^* = \alpha_i \frac{1}{1 + (\sigma_\eta^2/\sigma_i^2)} \tag{7}$$

and the limits of the OLS estimates of slope coefficients α_h and α_i in Eq. (5) *are given by*

$$\alpha_{i,m}^* = \alpha_i \frac{1-\rho^2}{1-\rho^2 + (\sigma_\eta^2/\sigma_i^2)}, \tag{8a}$$

$$\alpha_h^* = \alpha_h + \alpha_{i,m}^* \frac{\sigma_\eta^2}{\sigma_i \sigma_h} \frac{\rho}{1-\rho^2}, \tag{8b}$$

where σ_i^2, σ_h^2 *and* σ_η^2 *denote the variances of true implied volatility, past realized volatility, and the measurement error, respectively, and* ρ *denotes the correlation between true implied volatility and past realized volatility.*

The EIV problem has two effects. First is the familiar attenuation result, whereby the slope coefficient associated with implied volatility is itself biased downward towards zero, and more so in multiple regression (5) than in univariate regression (4) since $\alpha^*_{i,m} \leqslant \alpha^*_{i,u}$. Second, the slope coefficient for past volatility (α_h) is biased upward, provided α_i and ρ (the correlation between implied volatility and past realized volatility) are positive (as, empirically, they are). In particular, even when past volatility has *no* explanatory power, i.e., $\alpha_h = 0$, the limiting OLS estimates α^*_h will be positive. Further, Eq. (8b) suggests that the upward bias will be greater when α_i and ρ are large. In other words, the *smaller* the true bias in implied volatility and the stronger its relation to past volatility, the *stronger* will be the apparent evidence that it is inefficient.

The preceding discussion suggests that OLS estimates of both Eqs. (4) and (5) will be inconsistent and will typically lead to the incorrect conclusion that OEX implied volatility is biased and inefficient. Consistent estimation is possible, however, through an instrumental variables (IV) procedure. Under this procedure, implied volatility i_t is first regressed on an instrument (and any other exogenous variables). Fitted values from this regression replace implied volatility i_t in Eqs. (4) and (5), and the specifications are then estimated by OLS.

The IV procedure requires us to specify an instrument for implied volatility. Past implied volatility i_{t-1} is a natural candidate for an instrument, since it is correlated with true time t implied volatility i^*_t but is quite plausibly unrelated to η_t, the measurement error associated with implied volatility sampled one month later. Following Section 2.4, the error η_t can be attributed to (i) nonsynchronous measurement of option prices and index levels, (ii) early exercise and dividends, which are ignored in the Black–Scholes formula, (iii) bid–ask spreads, (iv) the 'wild-card' option, or (v) misspecification of the stochastic process governing index returns. Regarding (i), the closing cash index level precedes the level implied in option prices by up to 15 min, due to different closing times on the two exchanges. However, the *difference* between the cash market closing and the level implied in closing option prices is unlikely to be correlated at a one-month lag. Furthermore, Day and Lewis (1992) find no material differences in their results when actual index levels are used instead of the levels implied by closing option prices. With regard to (ii)–(iv), the discussion in Section 2.4 suggests that these are not likely to be major sources of error, because OEX dividends are smooth and our sample excludes deep out-of-the-money options and options with very short maturities. Finally, regarding (v), following Cox and Rubinstein (1985) and Lamoureux and Lastrapes (1993), the Black–Scholes implied volatility is approximately equal to expected future return volatility for at-the-money options (used in our study) even when returns follow the (non-Black–Scholes) stochastic volatility model of Hull and White (1987). Thus, misspecification error should be small and there is little reason for it to be correlated across time.

With i_{t-1} as the instrument, IV estimation of Eq. (4) simply consists of replacing implied volatility i_t by fitted values from the regression of i_t on i_{t-1} and then estimating Eq. (4) by OLS. Specification (5) is handled in a similar manner except that in the first step, i_t must be regressed on both i_{t-1} and h_{t-1} rather than on i_{t-1} alone. The first-step procedure of regressing i_t on i_{t-1} and h_{t-1} can also be used in estimating specification (4), in which case the procedure is equivalent to estimating (4) with two instruments (i_{t-1} and h_{t-1}) for implied volatility rather than just one (i.e., i_{t-1}). Empirically, both approaches yield nearly identical results. Since implied volatility is a generated regressor in the IV procedure, OLS standard errors are inappropriate; see, e.g., Pagan (1984), Murphy and Topel (1985), or Schwert and Seguin (1990) for a discussion in the the context of volatility series. Consistent standard errors for the IV estimates can be obtained following, e.g., Greene (p. 602).

Table 4 reports the IV estimates. Panel A reports estimates of the first-step regression (6), while Panel B reports the IV estimates of Eqs. (4) and (5). The IV estimates in Panel B provide evidence that implied volatility is both unbiased and efficient. The point estimates of α_i in both specifications – 0.97 and 1.04 for (4) and (5), respectively – are not significantly different from unity, which suggests that implied volatility is unbiased. Additionally, the IV estimate of α_h (-0.06) is not significantly different from zero, indicating that implied volatility is efficient. The Hausman (1978) χ^2 statistic (one degree of freedom) for detecting the presence of measurement error is 7.11 for specification (4) and 7.13 for specification (5). With both statistics significant at 5%, we reject the null hypothesis that the EIV problem in observed implied volatility does not matter. Collectively, the results indicate that the EIV problem offers a plausible explanation for both the bias and inefficiency associated with implied volatility in the OLS results. There is little evidence of either bias or inefficiency of implied volatility, once we account for the EIV problem.

Additional evidence on the unbiasedness and efficiency of OEX implied volatility is provided by reduced form regressions implied by Eqs. (5) and (6). Specifically, substitute the right-hand side of Eq. (6) for i_t into Eq. (5) to get

$$h_t = a_0 + a_h h_{t-1} + a_i i_{t-1} + e_{ht}, \tag{9}$$

where $a_h = \alpha_h + \alpha_i \beta_h$ and $a_i = \alpha_i \beta_i$. Solving these equations for α_h and α_i, we get

$$\alpha_h = a_h - \frac{a_i \beta_h}{\beta_i} \tag{10a}$$

$$\alpha_i = \frac{a_i}{\beta_i}. \tag{10b}$$

Table 4
Information content of implied volatility: Instrumental variables estimates

Panel A: First stage regression estimates

Dependent Variable: i_t

Independent variables			Adj. R^2	DW
Intercept	i_{t-1}	h_{t-1}		
-0.37[a]	0.39[a]	0.40[a]	62%	2.19
(-3.41)	(7.25)	(5.77)		

Panel B: Second stage IV estimates

Dependent Variable: h_t

Independent variables			Adj. R^2	DW
Intercept	i_t	h_{t-1}		
-0.15	0.97[a]		36%	2.04
(-0.63)	(8.23)			
-0.14	1.04[a]	-0.06	34%	1.98
(-0.57)	(3.75)	(-0.33)		

[a] p-value < 0.01.
[b] p-value $= 0.05$.
Panel A reports OLS estimates of specification (6) in the paper,

$$i_t = \beta_0 + \beta_i i_{t-1} + \beta_h h_{t-1} + e_{it},$$

where i_t denotes the natural logarithm of the Black–Scholes implied volatility for at-the-money call options on the S&P 100 index, measured at the beginning of month t, and h_t denotes the natural logarithm of the ex-post volatility of the option whose log implied volatility is i_t. The data consist of 139 nonoverlapping monthly observations on each volatility series, covering the period from November 1983 to May 1995. Numbers in parentheses denote asymptotic t-statistics.
Panel B reports instrumental variables (IV) estimates of specifications (4) and (5) in the paper,

$$h_t = \alpha_0 + \alpha_i i_t + \alpha_h h_{t-1} + e_t,$$

using lagged implied volatility i_{t-1} as an instrument for implied volatility, and the same data as in Panel A. Numbers in parentheses denote asymptotic t-statistics.

OLS estimates of Eq. (9), which are not reported here, yield $a_h = 0.36$ and $a_i = 0.40$, while the results in Panel A of Table 4 indicate $\beta_h = 0.40$ and $\beta_i = 0.39$. Substituting these values into Eq. (10a), we get $\alpha_h = -0.05$, which is close to the corresponding IV estimate (-0.06) and smaller than the OLS estimate (0.23). Likewise, the α_i implied by Eq. (10b) is $a_i/\beta_i = 1.03$, which is also closer to the IV estimate (1.04) than the OLS estimate (0.56).

Finally, while we have used estimates of Eq. (6) primarily in an instrumental variables setting, the specification is also of independent economic interest for two reasons. First, estimates of the specification (Panel A, Table 4) reveal that implied volatility is positively related to past realized volatility. This result suggests that option prices reflect future volatility information, since past and future volatility are positively correlated. Second, specification (6) models time t implied volatility as a function of past implied and realized volatility, variables known to the market in advance of time t. Thus, the specification provides a constructive (though nontheoretic) means of forecasting future implied volatility, and hence future option prices, using variables in the market's information set. Our estimates (Panel A, Table 4) indicate that both past realized volatility and past implied volatility are significant determinants of one-period-ahead implied option prices, and both variables have roughly the same regression coefficients. Though parsimonious, the model explains an economically significant portion (62%) of the variation in implied volatility.

3.3. Further tests for robustness

We conduct additional tests to verify the robustness of the results reported above, with little change in the basic character of our results. To begin, we replicate all tests using alternative estimators of volatility, namely, the square root of the mean squared return (Jorion, 1995) and the standard deviation of returns corrected for the sample autocorrelation in daily returns (French et al., 1987). We also recompute realized volatility using two other estimators of the mean return, namely, the average return for the previous month and the risk-free rate, $r_{f,t}$. Our results are not sensitive to the manner in which realized volatility is computed.

A second question relates to how the October 1987 stock market crash should be treated. Should this observation be included as an extreme, though unlikely, draw from the same distribution as the other observations or should it be eliminated as an unusual outlier? In any event, it is useful to know whether any of our conclusions hinge on the inclusion of crash-related observations. To examine this issue, we reestimate all our specifications after eliminating observations in which realized volatility is estimated using returns on October 19, 1987, with no substantive change in our results.

A third issue is that all tests reported above are within-sample. However, we also examine the predictive power of implied volatility in an out-of-sample context. Specifically, we use 115 observations for estimating specifications (4) and (5) and then use the estimates to forecast the one-period-ahead volatility. Repeating this procedure for the last 24 months of our sample period, we obtain three sets of out-of-sample forecast errors for the two-year period ending May 1995, pertaining to volatility forecasts based on implied volatility alone, past realized volatility alone, and both implied and past realized volatility. The mean squared forecast error when using implied volatility alone is roughly equal to that obtained using both implied and past volatility, and is smaller in both cases than that obtained using past realized volatility alone. While the difference is not statistically significant (the *t*-statistic is 0.88) since the sample size (24 forecast errors) is small, we certainly find no evidence that the within-sample results reported before are reversed in an out-of-sample context.

4. Comparison with previous literature

Our results are most directly related to those of Canina and Figlewski, who find that, for the OEX options market, implied volatility has little correlation with future volatility and past volatility is significantly related to future volatility and subsumes whatever information implied volatility has. These results are clearly at odds with ours, and this section examines why the two studies arrive at different results.

Our research design differs from that of CF mainly along two dimensions. First, the CF data are drawn from a 48-month period from March 1983 to March 1987, whereas our data span a longer period of time, namely, the 139 months from November 1983 to May 1995. A second difference is the sampling procedure. Both aspects help explain the differences in results across the two studies, as we discuss below.

To control for the time period effect, we reestimate specifications (4) and (5) separately for two separate subperiods: a pre-crash subperiod (November 1983 to March 1987) similar to that used in CF, and a post-crash subperiod (December 1987 to May 1995) that begins after the crash. Our pre-crash subperiod begins only in November 1983 rather than March 1983 (as in CF) since options with monthly expiries became available only in November 1983. Our pre-crash subperiod ends in March 1987, as in CF. Our post-crash period begins in December 1987 in order to eliminate observations in which realized volatility for October 1987 appears as a dependent or independent variable.

Panel A of Table 5 reports estimates of Eqs. (4) and (5) for the pre-crash subperiod, and Panel B reports the corresponding post-crash results. We report the results using both the log-volatility and the volatility level series. As both sets of results are qualitatively similar, we focus the discussion on results obtained with the log-volatility series. We compare the pre-crash estimates to the post-crash and full period estimates as well as to the estimates reported by CF.

4.1. The slope coefficient for implied volatility

The pre-crash estimates of α_i, the slope coefficient for implied volatility, are 40–50% smaller than the corresponding post-crash (and full–period) estimates,

Table 5
Information content of implied volatility: Subperiod analysis

Dependent variable: log realized volatility h_t					Dependent variable: realized volatility σ_{ht}				
Independent variables			Adj. R^2	DW	Independent variables			Adj. R^2	DW
Intercept	i_t	h_{t-1}			Intercept	σ_{it}	$\sigma_{h,t-1}$		
Panel A: Pre-crash subperiod									
− 1.36[a]	0.33[a]		18%	1.78	0.09[a]	0.31[a]		16%	1.91
(− 6.44)	(3.09)				(5.55)	(2.90)			
− 1.39[a]		0.30	5.8%	2.05	0.11[a]		0.20	0.8%	1.96
(− 4.19)		(1.85)			(4.66)		(1.15)		
− 1.25[a]	0.29[a]	0.09	16%	1.95	0.09[a]	0.31[a]	0.00	14%	1.92
(− 3.90)	(2.40)	(0.48)			(3.88)	(2.58)	(0.03)		
Panel B: Post-crash subperiod									
− 0.27	0.92[a]		50%	2.03	0.00	0.92[a]		46%	2.00
(− 1.37)	(9.45)				(0.02)	(8.82)			
− 0.87[a]		0.60[a]	37%	2.42	0.05[a]		0.58[a]	37%	2.38
(− 4.99)		(7.38)			(4.57)		(7.36)		
− 0.26	0.74[a]	0.18	50%	2.27	0.00	0.70[a]	0.20	47%	2.26
(− 1.33)	(4.96)	(1.60)			(0.73)	(4.26)	(1.76)		

[a] p-value < 0.01.
[b] p-value = 0.05.
OLS estimates of specifications (4) and (5) in the paper. The regressions reported in the left side of each panel are of the form

$$h_t = \alpha_0 + \alpha_i i_t + \alpha_h h_{t-1} + e_t.$$

Here, i_t denotes the natural logarithm of the Black–Scholes implied volatility for at-the-money call options on the S&P 100 index, measured at the beginning of month t, and h_t denotes the natural logarithm of the ex-post daily return volatility of the index, over the life of the option whose log implied volatility is i_t. The regressions reported in the right side of each panel are of the form

$$\sigma_{ht} = \delta_0 + \delta_i \sigma_{it} + \delta_h \sigma_{h,t-1} + e_{1t},$$

where σ_{ht} and σ_{it} satisfy $h_t = \log \sigma_{ht}$ and $i_t = \log \sigma_{it}$. The results reported in Panel A are based on nonoverlapping monthly volatility observations for the 41 months from November 1983 to March 1987, while those reported in Panel B are based on nonoverlapping monthly volatility observations for the 90 months from December 1987 to May 1995. Numbers in parentheses denote asymptotic t-statistics.

and the difference between the two is significant. The Chow (1960) test statistic for a structural change around the crash takes the value 8.12 for specification (4) and 3.30 for specification (5) and both test statistics are significant at 5%. However, the Chow test is not robust to heteroskedasticity, and Section 2.3

demonstrates that the volatility series are heteroskedastic because the variance of volatility is greater after the crash than before. Hence, we also compute the likelihood ratio test statistic for a structural change around the crash. The likelihood ratio statistics for Eqs. (4) and (5) are 20.52 and 14.46, respectively, and both are significant at 5%. These results indicate that there was a regime shift following the crash, with implied volatility significantly more biased before the crash.

Our pre-crash subperiod estimates of α_i, which are 0.33 for Eq. (4) and 0.29 for Eq. (5), are larger than the largest relevant estimates obtained by CF, or 0.22 for Eq. (4) and 0.08 for Eq. (5), but the two are now closer than in Section 3.1. Thus, part but not all of the difference between our estimates of α_i and those reported in CF could be attributed to differences in time periods considered by the two papers. Qualitatively similar results obtain when the pre-crash subperiod is defined to end in September 1987 rather than in March 1987.

Why is the pre-crash estimate of α_i smaller than its post-crash and full period counterparts? One possibility, of course, is that OEX options were priced differently prior to October 1987, perhaps because the market was still in its infancy in the pre-crash period (index options began trading only in 1983) or because the stochastic process followed by index returns was different in the pre-crash subperiod. We cannot reject the possibility of such nonstationarities in the return-generating process or in the model for pricing index options. It has also been suggested to us that the pre-crash and post-crash results differ because post-crash implied volatilities price stock market crashes similar to that experienced in October 1987. This argument, however, suggests that the Black–Scholes prices and implied volatilities are more biased in the post-crash period; it does not explain our finding that implied volatility is more biased in the *pre-crash* period.

A second possibility suggested by Proposition 1 is that the *signal-to-noise* ratio is lower before the crash than after. This could happen either because there is less variation in true implied volatility (σ_i^2 from Proposition 1 is lower) during the pre-crash subperiod, or because implied volatility is noisier (σ_η^2 is higher) in this subperiod. The data in Table 1 are consistent with both possibilities, though we cannot empirically estimate the individual contribution of each factor to the lower signal-to-noise ratio.

Consider the variation in realized volatility in the 1983–1987 subperiod relative to that in the post-crash subperiod. Both numbers are reported in Table 1. The pre-crash subperiod variance of realized volatility is only about 29% of its post-crash subperiod counterpart, and the pre-crash subperiod variance of log realized volatility is only about 34% of the corresponding post-crash subperiod variance. Volatility is less variable in the pre-crash subperiod. While the pre-crash data in Table 1 pertain to the subperiod from November 1983 to September 1987, similar numbers obtain when the pre-crash subperiod is defined to end in March 1987, as in CF.

The data in Table 1 also suggest that implied volatility is noisier before the crash than after. Consider the following two facts. First, implied volatility is, somewhat counterintuitively, nearly twice as variable as realized volatility in the pre–crash subperiod, while it is smoother than realized volatility in the post–crash period as well as in the full period. Second, the ratio of the 1983–1987 subperiod variance of implied volatility to the post-crash subperiod variance is about 94% (107% for log implied volatility), whereas the same variance ratio for realized volatility is much smaller at 29% (34% for log realized volatility). Both facts are consistent with implied volatility being noisier in the pre-crash period.

Summing up, the difference between our study and that of CF on the bias of implied volatility could be attributed to a regime shift following the October 1987 stock market crash, which in turn could be attributed to nonstationarities in either the index return process or the pricing model or to a poor signal-to-noise ratio in the pre-crash subperiod.

4.2. *The slope coefficient for past volatility*

Canina and Figlewski find that past volatility subsumes the information content of implied volatility when the two series are sampled at daily frequency. They report that the slope coefficient for past volatility in Eq. (5), the associated t-statistic, and the regression R^2 all are virtually invariant to whether implied volatility is included as an explanatory variable or not. In contrast, we find that even in the pre-crash subperiod, *implied* volatility dominates past volatility (see Panel A of Table 5) when the two series are sampled in a nonoverlapping manner.

Why does past volatility appear to have limited explanatory power in our study but not in CF, even for data from a similar time period? We argue that the difference is because of differences in our sampling methods. Specifically, the CF procedure of sampling overlapping data tends to overstate the explanatory power of past volatility, as discussed next.

Accordingly, consider any one option expiration date, say τ, and the time series of past and future volatility built around this date. If t is any date prior to τ, CF define the time t future volatility ($\sigma_{f,\,\tau,\,t}$) and 60-day past volatility ($\sigma_{p,\,\tau,\,t}$) as follows:

$$\sigma^2_{f,\tau,t} = \frac{1}{\tau - t} \sum_{k=t}^{\tau} r_k^2 \tag{11}$$

and

$$\sigma^2_{p,\tau,t} = \frac{1}{60} \sum_{k=t-60}^{t} r_k^2. \tag{12}$$

For expositional ease, we ignore the correction for the squared mean return in defining volatility in both Eqs. (11) and (12). Consider now the next term in the daily time series of past volatility $\sigma_{p,\tau,t+1}$:

$$\sigma^2_{p,\tau,t+1} = \frac{1}{60} \sum_{k=t-59}^{t+1} r_k^2. \tag{13}$$

The extreme overlap between successive terms in the past volatility series is obvious from Eqs. (12) and (13). Given an expiry date, each element in the time series of realized volatility is a moving average of 60 squared returns, and $\sigma_{p,\tau,t}$ and $\sigma_{p,\tau,t+1}$ share 59 of these terms. Thus, for a given option expiration τ, successive elements in the past volatility series, $\sigma_{p,\tau,t}$ and $\sigma_{p,\tau,t+1}$, are virtually identical and only differ by a small amount of noise of the order of $\frac{1}{30}$ of $\sigma_{p,\tau,t}$. Thus, even if one-day volatility for day t and day $t+1$ had zero correlation, the procedure of sampling overlapping data would introduce dependence in the time series $\{\sigma_{p,\tau,t}\}_{\tau=1,2,\ldots}$, and, likewise, in the series $\{\sigma_{f,\tau,t}\}_{\tau=1,2,\ldots}$.

It is well known (e.g., Richardson and Smith, 1991) that moving average processes induced by such extreme overlap distort inferences in a regression context. Slope coefficients have poor finite sample properties, and the associated t-statistics and R^2 will typically be overstated. The usual Richardson and Smith analysis for overlapping data cannot, however, be applied directly in our context, since our volatility series $\{\sigma_{pt}\}$ and $\{\sigma_{ft}\}$ are heterogeneous composites of several more homogeneous moving average processes, namely, $\{\sigma_{p,\tau,t}\}$ and $\{\sigma_{f,\tau,t}\}$, one for each expiration date τ. Nevertheless, given the manner in which the daily series are constructed, biases due to overlap are likely to arise, and it is useful to analyze their effect.

Accordingly, we construct a sample of overlapping volatility series sampled at daily frequency along the lines of CF. Given an option expiration date τ (there is one per month), we first pick a day t such that $7 < \tau - t < 35$, i.e., such that there are at least seven and not more than 35 calendar days from date t to the next option's expiration. We record the implied volatility of an at-the-money call option on date t and compute 60-day past volatility and future realized volatility using Eqs. (11) and (12), respectively. The two steps are then repeated for every option expiration date τ preceding the October 1987 stock market crash. Daily option price data were hand-collected from The Wall Street Journal, and to the extent there are measurement errors in such data, our test procedures will understate the true explanatory power of implied volatility. Note that $\mathrm{cov}(\sigma_{p,\tau,t}, \sigma_{p,\tau,t+i})$ is never zero in the daily data series, simply by construction. To see why, note that $\sigma_{p,\tau,t}$ is the volatility for the last 60 days prior to t, whereas $\tau - t$ never exceeds 35. Thus, given τ, $\sigma_{p,\tau,t}$ and $\sigma_{p,\tau,t+i}$ share at least 25 squared returns for all i and t.

We regress future volatility on past realized volatility and implied volatility for the sample of daily volatilities, i.e., we estimate

$$\log \sigma_{f,\tau,t} = \gamma_0 + \gamma_p \log \sigma_{p,\tau,t} + \gamma_i \log \sigma_{i,\tau,t} + v_{\tau t} \tag{14}$$

and the corresponding univariate regressions. Table 6 reports the results.

Table 6
Information content of implied volatility: Daily estimates for pre-crash subperiod dependent variable: Realized volatility log $\sigma_{f\tau t}$

Independent variables			Adj. R^2	DW
Intercept	log $\sigma_{p\tau t}$	log $\sigma_{i\tau t}$		
Panel A: OLS estimates				
– 0.63	0.70		17%	0.26
– 5.04)	(11.43)			
– 1.63		0.20	7.5%	0.28
(– 29.77)		(7.39)		
– 0.66	0.63	0.06	17%	0.27
(– 5.33)	(8.64)	(2.03)		
Panel B: FGLS estimates				
– 3.16	– 0.57		11%	2.00
(– 9.03)	(– 3.31)			
– 2.00		0.01	0%	1.80
(– 36.82)		(0.80)		
– 3.15	– 0.57	0.01	11%	2.00
(– 8.99)	(– 2.24)	(0.54)		

Panel A reports OLS estimates of specification (14) in the paper,

$$\log \sigma_{f\tau t} = \gamma_0 + \gamma_p \log \sigma_{p\tau t} + \gamma_i \log \sigma_{i\tau t} + v_{\tau t}.$$

Here, log $\sigma_{f\tau t}$ denotes the natural logarithm of the realized volatility over the period from date t to the nearest date τ on which an OEX option expires, log $\sigma_{p\tau t}$ denotes the log realized volatility over a 60-day period prior to date t, annd log $\sigma_{i\tau t}$ denotes the natural logarithm of the date t Black–Scholes implied volatility of an at-the-money call option expiring at τ. The data consist of 653 observations on each volatility series, covering the time period from November 1983 to March 1987. Numbers in parentheses denote asymptotic t-statistics.
Panel B reports Cochrane–Orcutt feasible generalized least-squares (FGLS) estimates of the specification estimated in Panel A. Numbers in parentheses denote asymptotic t-statistics.

The univariate regression coefficient for past volatility is 0.70 while the multiple regression yields $\gamma_p = 0.63$. The corresponding numbers in CF are 0.57 and 0.49, respectively. The univariate and multivariate estimates of γ_i, the slope for implied volatility, are 0.20 and 0.06, respectively, while CF report estimates of 0.22 and 0.08. Our pre-crash daily results are thus qualitatively similar, though not identical, to those obtained by CF. Asymptotic t-statistics based on Newey and West standard errors are also similar to those obtained in CF when the number of lags used in standard error computations exceeds ten (the actual number of lags should be about 25, based on the extent of overlap within the past and future volatility series).

The daily results present a sharp contrast to the monthly results for the same period (Panel A, Table 5). In the monthly regressions, past volatility has little incremental explanatory power, whereas past volatility appears to dominate implied volatility in the daily data. However, consider the Durbin–Watson statistic for the daily data, which ranges from 0.26 to 0.28 (Panel A, Table 6). The low Durbin–Watson statistic indicates that the daily residuals are highly autocorrelated and raises the possibility of a spurious regression phenomenon (see, e.g., Granger and Newbold, 1974; Phillips, 1986). If so, the regression estimates are inconsistent and converge to a nondegenerate limiting distribution involving functionals of Brownian motion. Feasible generalized least-squares (FGLS) provides one means of handling the spurious regression phenomenon, and also provides alternative consistent estimates of Eq. (14) in the presence of autocorrelated errors (Hamilton, 1994; pp. 561–562). We report Cochrane–Orcutt FGLS estimates of specification (14) in Panel B of Table 6.

The Durbin–Watson statistic for the transformed regressions used in FGLS estimation is close to two, indicating that the FGLS procedure has virtually eliminated the autocorrelation in the daily residuals. Somewhat surprisingly, the FGLS estimates of γ_p are now negative (-0.57). This estimate, viewed in conjunction with the low Durbin–Watson statistic for the daily regression, calls for considerable caution in interpreting the daily OLS results.[4] On the other hand, the monthly regression (Table 6) has a Durbin–Watson statistic close to two, which indicates that monthly residuals exhibit little autocorrelation. The monthly results in Table 6 suggest that while implied volatility is certainly biased in the pre-crash period, it dominates past volatility in explaining future volatility. In fact, implied volatility appears to subsume the information contained in past volatility for the very subperiod in which the daily results seem to lead to exactly the opposite conclusion.

[4] That the daily regression estimates are unreliable is consistent with Richardson and Stock (1989) (RS). RS show that when the degree of overlap (J) is large relative to the time series length (T), OLS estimates are inconsistent and, furthermore, converge in the limit to a nondegenerate random variable. The RS distribution cannot, however, be used to analyze our Eq. (14), which differs in three significant ways from the RS model. Specifically, (i) regression errors are serially uncorrelated in RS under the null, whereas they are correlated in Eq. (14), from the time-varying volatility literature; (ii) only one independent variable (the lagged left-hand-side variable) appears in RS, while we have one extra variable (implied volatility); and (iii) past volatility is not just the lagged future volatility in Eq. (14). Rather, future volatility is computed using between seven and 35 squared returns, while past volatility uses 61 prior squared returns, which makes the overlap structure more heterogeneous and complicated than that in RS. Thus, derivation of the asymptotic distribution is significantly more complex than in RS and is not attempted here.

5. Conclusions

The fundamental question addressed in this study is whether the volatility implied in S&P 100 index option prices predicts ex-post realized volatility. Previous work indicates that implied volatility is both biased and inefficient when the series are sampled on a day-to-day or weekly basis. Such results have been construed as a vote against the joint hypothesis that option markets aggregate volatility information efficiently and the Black–Scholes option pricing formula is valid for OEX options.

This study introduces a different research design to examine the relation between implied volatility and realized volatility. Our analysis employs a lower (monthly) sampling frequency and nonoverlapping data that span a longer period of time, such that exactly one implied and realized volatility estimate pertain to each time period under consideration. Our conclusions are significantly different from those of previous literature, and the difference is robust to variations in econometric approach. We find that implied volatility does predict future realized volatility in isolation as well as in conjunction with the history of past realized volatility. In fact, OEX implied volatility subsumes the information content of past volatility in some of our specifications. We attribute the differences to our use of a longer time period relative to previous work and our use of nonoverlapping data.

Our analysis sheds light on an interesting aspect of the October 1987 stock market crash. While the crash itself was an extreme event accompanied by a period of high volatility (Schwert, 1990), we find that it is also associated with a structural change in the pricing of index options. Implied volatility is a significantly better predictor of future volatility following the crash. Our results also complement Jorion, who reports that implied volatility is an efficient volatility forecast in the foreign exchange futures market but is unable to definitively explain why the results differ from those reported for the OEX market. Jorion suggests two explanations for the difference, namely that index options are priced differently from foreign exchange futures options due to higher trading costs in the cash index markets or that OEX implied volatility has a greater errors-in-variable problem. Our results lend some support to the latter explanation. Our results also complement the findings of Harvey and Whaley (1992), who suggest that the errors-in-variable issue is important in the context of OEX options. Finally, the results provide an empirical justification for the common practice of interpreting the Black–Scholes implied volatility of OEX options as a volatility forecast rather than as just a convenient means of quoting option prices.

Virtually all studies in the literature have focused on predicting future *realized* volatility using options data, with less emphasis on the issue of predicting *implied* volatility. Our results suggest that the implied volatility of at-the-money call options is predictable using a parsimonious set of variables in the market information set. Interpreting implied volatility as an option's price, we have a constructive empirical approach for forecasting future prices of a subset of OEX options. An obvious extension of our paper would be to test whether prices of other options, such as out-of-the money options, puts, and longer maturity options, can be similarly forecast.

References

Bergman, Y., Grundy, B., Wiener, Z., 1996. General properties of option prices. Journal of Finance 51, 1573–1610.

Black, F., Scholes, M., 1973. The valuation of options and corporate liabilities. Journal of Political Economy 81, 637–654.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

Box, G.E.P., Pierce, D.A., 1970. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. Journal of the American Statistical Association 65, 1509–1526.

Brennan, M., 1979. The pricing of contingent claims in discrete time models. Journal of Finance 34, 53–68.

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

Chow, G., 1960. Tests of equality between sets of coefficients in two linear regressions. Econometrica 28, 591–605.

Constantinides, G., 1994. Option prices with transaction costs. Unpublished working paper. University of Chicago.

Cox, J., Ross, S., 1976. The valuation of options for alternative stochastic processes. Journal of Financial Economics 3, 145–166.

Cox, J., Rubinstein, M., 1985. Options Markets. Prentice-Hall, Englewood Cliffs, N.J.

Day, T., Lewis, C., 1988. The behavior of the volatility implicit in option prices. Journal of Financial Economics 22, 103–122.

Day, T., Lewis, C., 1992. Stock market volatility and the information content of stock index options. Journal of Econometrics 52, 267–287.

French, K., Schwert, G.W., Stambaugh, R., 1987. Expected stock returns and volatility. Journal of Financial Economics 19, 3–30.

Granger, C.W.J., Newbold, P., 1974. Spurious regressions in econometrics. Journal of Econometrics 2, 111–120.

Greene, W., 1993. Econometric Analysis. Macmillan, New York.

Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton, NJ.

Harvey, C.R., Whaley, R., 1992. Market volatility prediction and the efficiency of the S&P 100 index option market. Journal of Financial Economics 31, 43–73.

Hausman, J., 1978. Specification tests in econometrics. Econometrica 46, 1251–1271.

Heston, S., 1993. A closed form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies 6, 327–343.

Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatilities. Journal of Finance 42, 281–300.

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

Lamoureux, C.G., Lastrapes, W., 1993. Forecasting stock return variance: towards understanding stochastic implied volatility. Review of Financial Studies 6, 293–326.

Latane, H., Rendleman, R., 1976. Standard deviation of stock price ratios implied in option prices. Journal of Finance 31, 369–381.
Murphy, K.M., Topel, R., 1985. Estimation and inference in two-step econometric models. Journal of Business and Economic Statistics 3, 370–379.
Nelson, D.B., 1990. ARCH models as diffusion approximations. Journal of Econometrics 45, 7–39.
Newey, W., West, K., 1987. A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Pagan, A., 1984. Econometric issues in the analysis of regressions with generated regressors. International Economic Review 25, 221–247.
Phillips, P.C.B., 1986. Understanding spurious regressions in econometrics. Journal of Econometrics 33, 311–340.
Poterba, J., Summers, L., 1986. The persistence of volatility and stock market fluctuations. American Economic Review 76, 1142–1151.
Richardson, M., Smith, T., 1991. Tests of financial models in the presence of overlapping observations. Review of Financial Studies 4, 227–254.
Richardson, M., Stock, J., 1989. Drawing inferences from statistics based on multiyear asset returns. Journal of Financial Economics 25, 323–348.
Rubinstein, M., 1976. The valuation of uncertain income streams and the pricing of options. Bell Journal of Economics 7, 407–425.
Schwert, G.W., 1987. Effects of model specification on tests for unit roots in macroeconomic data. Journal of Monetary Economics 20, 73–103.
Schwert, C.W., 1989. Why does stock market volatility change over time? Journal of Finance 44, 1115-1153.
Schwert, G.W., 1990. Stock volatility and the crash of ’87. Review of Financial Studies 3, 77–102.
Schwert, G.W., Seguin, P., 1990. Heteroskedasticity in stock returns. Journal of Finance 45, 1129–1155.
Sheikh, A., 1989. Stock splits, volatility increases and implied volatility. Journal of Finance 44, 1361–1372.