International Review of Financial Analysis 35 (2014) 32–45

Contents lists available at ScienceDirect

International Review of Financial Analysis

Forecasting option smile dynamics☆

Van Le [a,*], Ralf Zurbruegg [b,1]

[a] Newcastle Business School, University of Newcastle, University Drive, Callaghan, NSW 2308, Australia
[b] Adelaide Business School, University of Adelaide, 10 Pulteney St, Adelaide, SA 5000, Australia

ARTICLE INFO

Article history:
Received 20 January 2014
Received in revised form 6 June 2014
Accepted 19 July 2014
Available online 1 August 2014

JEL classification:
G12
G19

Keywords:
Volume
Implied volatility
Option smile
Forecasting
VAR models

ABSTRACT

Practitioners have long tried to exploit the predictability of the option implied volatility smile. Motivated by the recent developments in the literature focusing on market-based option pricing arguments, this paper proposes the introduction of trading volume into a vector autoregressive (VAR) model to improve forecasts of the smile dynamics. The augmented VAR-volume model produces quality forecasts of the smile surface and explains its dynamic changes over time relatively well. Our results suggest that the incorporation of trading volume leads to it outperforming other alternative forecast approaches, as well as being robust to a variety of perturbations of the data and offers scope for investors to more accurately predict option implied volatility in the future.

1. Introduction

In the empirical option pricing literature, option price departures from the Black and Scholes (1973) model are often characterised by the implied volatility smile (IVS), whereby the implied volatility tends to vary with the option strike price. Despite the extensive literature examining alternative explanations of the smile's existence at any given time, it is only recently that its time varying dynamics have captured the attention of researchers in financial modelling. In fact, very few investigations exist on the topic of forecasting the dynamics of the smile, despite its importance to derivatives pricing, as well as portfolio and risk management.[2] This paper addresses the issue by proposing a model that accounts for the link that exists between trading volume and observed option prices when forecasting the IVS dynamics.

Previous studies report that the smile dynamics are time-varying (Bates, 1991; Bollen and Whaley, 2004; Gemmill, 1996; Gonçalves and Guidolin, 2006; Mixon, 2002). Gemmill (1996) shows that the smile behaviour in time, being characterised by a single measure of skewness, is dependent on the price level of the underlying index. Using over-the-counter options trading data, Mixon (2002) finds that variations over time in the IVS of the S&P 500 index are driven by multiple economic state factors. In particular, structural models which have been developed to provide explanations for the existence of the smile suggest that it is forecastable on the basis of information related to latent factors driving its changing shape. Recent evidence from Bollen and Whaley (2004) and others suggests that option pricing has partly been driven by the market trading process to incorporate new information reflected in the net market order. The form of the relationship between trading volume and the changing shape of the smile is, however, ambiguous. This gives rise to a line of inquiry which examines empirically the information content of trading volume in modelling the time series dynamics of the smile.

This paper examines the incorporation of trading volume into the time series model of the IVS to determine if it contains information relevant to the latent factors driving the time variations of option prices. This is important for the validation of the economic explanations underlying the smile evolution, one of which is directly related to the arrival of new information to the market. In fact, economic theory has long established that a slow diffusion of news into the market may cause a lead–lag (causal) relationship between volume and asset prices. While this proposition has clearly been shown to exist in the equity market

☆ We wish to thank an anonymous referee for a number of constructive comments that have significantly improved this paper.
* Corresponding author. Tel.: +61 2 4921 8798.
E-mail addresses: van.le@newcastle.edu.au (V. Le), ralf.zurbrugg@adelaide.edu.au (R. Zurbruegg).
[1] Tel.: +61 8 8313 5535.
[2] Option implied volatility is often utilised to extract information on the expected volatility (Day and Lewis, 1992; Harvey and Whaley, 1992) or the probability distribution (Jackwerth and Rubinstein, 1996; Melick and Thomas, 1997) of the underlying asset. In addition, implied volatility is commonly used by practitioners for option pricing and trading purposes.

http://dx.doi.org/10.1016/j.irfa.2014.07.006
1057-5219/

through the development of many influential, theoretical models, such as the sequential information hypothesis (Copeland, 1976) and the noise trading hypothesis (Brock and Lebaron, 1996; Harris and Raviv, 1993; Iori, 2002), empirical work on this line of inquiry has received little attention in the field of options research. This study therefore attempts to fill the gap in the literature by specifically seeking to examine whether past volume contains additional, useful information about the future dynamics of the smile. This is in contrast to the existing studies which have mainly focused on either a contemporaneous relationship between volume and option prices (Bollen and Whaley, 2004; Ni et al., 2008) or a lead–lag relationship between stock and option markets (Anthony, 1988; Chan et al., 2002).

Further, this paper's special interest in the role of trading volume in forecasting the smile dynamics rests on the scarcity of studies which link the volume-implied volatility relation with forecasting applications. Previous studies which examine the effect of trading on option pricing do so primarily to investigate the efficiency of the option markets, not to improve the forecasts of the smile dynamics per se. Some exceptions that do focus on using trading volume to forecast volatility include Brooks (1998), Donaldson and Kamstra (2005), and Le and Zurbruegg (2010). The latter two papers both find evidence that the usage of trading volume information can improve volatility forecasts. Donaldson and Kamstra (2005) use trading volume as a switching mechanism between using ARCH and option implied volatilty estimates, while Le and Zurbruegg (2010) incorporate trading volume directly into the variance equations of EGARCH models. They show that by doing this, volatility forecasts for highly liquid S&P 500 stocks improve. However, they caution that this effect may diminish for less traded stocks.

The above papers provide impetus for us to consider whether this lead–lag relationship between volume and option price may also be useful for IVS forecasts, as well as option implied volatility forecasts in general. To examine this, we incorporate trading volume into a vector-autoregressive (VAR) structure as part of a two-stage process to model the smile dynamics proposed by Gonçalves and Guidolin (2006) since this particular technique appears to fare relatively well compared to other competing forecasting models existing in the literature. This approach is intuitively appealing, not only to determine the importance of the link that exists between volume and option pricing but also to highlight the potential information captured by trading activity. It delivers a simple yet effective mechanism which may prove to yield a beneficial improvement in forecasting option implied volatility (price). To assess the performance of the proposed model, both statistical and economic criteria are used in the assessment. From the economic perspective, different simulated trading strategies are performed, both with and without transaction costs, to evaluate whether the inclusion of trading volume does lead to tangible benefits to option traders and practitioners.

Options traded on the S&P 500 index in the period from 2nd January 1996 to 8th September 2009 are examined in this study. Results produced by the investigation suggest that the proposed augmented VAR-volume model leads to quality forecasts of the smile surface and explains relatively well its time series properties. Further, out-of-sample evidence indicates that the inclusion of trading volume leads to improved forecasts, outperforming other alternate forecasting techniques. This is particularly true for options that are more liquid. This would provide support to the extant research aiming to examine the information reflected by market trading activities. The findings offer an important insight into the empirical relation between volume and asset prices in the context of option pricing, confirming that changes in the level of market trading activities are important for explaining the empirical time variations of the IVS.[3]

[3] The approach of this paper is closely related to studies which either feature option pricing models based on a general equilibrium framework or employ market-based arguments to explain the time varying properties and other stylised facts of the IVS (see Bollen and Whaley, 2004; Chan et al., 2004; Garcia et al., 2003; Gârleanu et al., 2009; Guidolin and Timmermann, 2003; Kang and Park, 2008; Ni et al., 2008).

The remainder of the paper is structured as follows. Section 2 describes the nature of the data set and the different assessment criteria employed to evaluate the forecast performance of the proposed VAR-volume model in comparison with other prominent forecasting techniques used in the literature. Section 3 is devoted to the discussion of some key findings which emerge from the empirical results, while Section 4 provides a conclusion and outlines some directions for future research.

2. Data and methodology

2.1. Data

The primary data set employed in this study is from the Securities Industry Research Centre of Asia-Pacific (SIRCA). The data consist of intra-day records of all trades of the S&P 500 index options transacted on the Chicago Board Options Exchange (CBOE) from 2nd January 1996 to 8th September 2009. Of this sample, the first period from 2nd January 1996 to 26th January 2004 is used for the in-sample hypothesis testing and model construction,[4] while the remaining data are utilised for the out-of-sample evaluation. It is important to note that S&P 500 index options are of European style and expire on the third Friday of each calendar month. Each option is categorised as a put or a call and characterised by its strike price and time to expiration, as distinctively identified by its Reuters Instrument Code (RIC). Minute-to-minute observations of the index level are also collected from the same SIRCA database. All the remaining data, including the dividend yield on the S&P 500 index and the 30-day bank accepted bill rate, which is utilised as the risk free rate, are collected from DataStream International.

In contrast to the OptionMetrics Ivy database, which only has records of the last trade of the day for each option series employed in similar studies, the data used in this paper report all transactions executed throughout the day, with the exact time stamp of trading to the nearest millisecond. This applies a tight time sampling window on the daily observation of option trades available. In fact, only those options which are actively traded during the next-to-closing trading session from 3:00 pm to 4:00 pm every trading day are used.[5] This choice is optimal for the minimisation of the effect of non-simultaneity of trades and other potential intra-day trading effects,[6] while maintaining a sufficient number of traded options to construct the implied volatility surface each day.

In addition, several other exclusionary filters, analogous to those in Bakshi et al. (1997), Bollen and Whaley (2004) and Gonçalves and Guidolin (2006), have been applied to ensure the integrity of the data when running our in-sample regressions. First, only options with less than 180 days to maturity are selected to avoid problems associated with no trades. Second, options with less than 5 days to maturity are excluded to eliminate any effect of option expiration.[7] Third, following Bollen and Whaley (2004), options with absolute delta below 0.02 and above 0.98 are also excluded due to distortions caused by price discreteness. Fourth, trading dates which have less than 5 option series traded within the trading window previously specified are eliminated to avoid the adverse effect of illiquid trades and price discreteness in fitting the daily smile surface. Finally, the data are filtered for any

[4] The in-sample period represents approximately 60% of the total number of observations.

[5] Because the stock market closes at 4:00 pm, option trades with time stamps later than 4:00 pm are eliminated to avoid the effect of non-synchronous trades when matching the index price level to each option trade.

[6] Please refer to Battalio and Schultz (2011) for a discussion of the problem of non-synchronous prices and microstructure issues responsible for most of the apparent arbitrage opportunities identified using the OptionMetrics IVY database. In particular, there may be a significant time lag between the closing trades of different option series transacted within the same reported trading day.

[7] Gonçalves and Guidolin (2006) further argue that their prices contain liquidity-related bias and contain little information on the time dimension of the IVS.