



Why do we smile? On the determinants of the implied volatility function

Ignacio Peña ^{a,1}, Gonzalo Rubio ^{b,*}, Gregorio Serna ^a

^a *Departamento de Economía de la Empresa, Facultad de Ciencias Económicas y Empresariales, Universidad Carlos III, Avda. de Madrid 126, 28903 Getafe, Madrid, Spain*

^b *Departamento de Fundamentos del Análisis Económico, Facultad de Ciencias Económicas y Empresariales, Universidad del País Vasco, Avda. L. Aguirre 83, 48015 Bilbao, Spain*

Received 10 November 1997; accepted 23 September 1998

Abstract

We report simple regressions and Granger causality tests in order to understand the pattern of implied volatilities across exercise prices. We employ all calls and puts transacted between 16:00 and 16:45 on the Spanish IBEX-35 index from January 1994 to April 1996. Transaction costs, proxied by the bid–ask spread, seem to be a key determinant of the curvature of the volatility smile. Moreover, time to expiration, the uncertainty associated with the market and the relative market momentum are also important variables in explaining the smile.

JEL classification: G10; G12; G13

Keywords: Smiles; Bid–ask spread; Volatility; Causality

* Corresponding author. Tel.: 34 4 479 7470; fax: 34 4 479 7474; e-mail: jepuirg@bs.ehu.es

¹ Tel.: 34 (9)1 6249625; fax: 34 (9)1 6249875; e-mail: ypenya@eco.uc3m.es

1. Introduction

There has recently been an increasing number of papers dealing with one prominent feature of option pricing data. It is well known that after the October 1987 crash the implied volatility computed from options on stock indexes in the US market inferred from the Black and Scholes (1973) formula (BS henceforth) appears to be different across exercise prices. This is the so-called “volatility smile”.² In fact, as pointed out by Dumas et al. (1996) (DFW henceforth), implied volatilities of the S&P 500 options decrease monotonically as the exercise price becomes higher relative to the current level of the underlying asset.

Of course, given the BS assumptions, all option prices on the same underlying security with the same expiration date but with different exercise prices should have the same implied volatility. However, the volatility smile pattern suggests that the BS formula tends to misprice deep in-the-money and deep out-of-the-money options. There have been various attempts to deal with this apparent failure of the BS valuation model. The stochastic volatility framework of Hull and White (1987) was the first systematic approach in option pricing literature to recognize nonconstant volatility.³ When volatility is stochastic but uncorrelated with the underlying asset price, they show that the price of a European option is the BS price integrated over the probability distribution of the average variance during the life of the option.⁴ Unfortunately, however, this framework generally requires a market price of volatility risk. In other words, with stochastic volatility, a second factor is introduced, requiring the option to satisfy a bivariate stochastic differential equation. Since the volatility – the second factor – is not spanned by existing securities, arbitrage pricing techniques are no longer valid. We must therefore introduce the explicit exogenous market price of volatility risk. We face similar problems if we introduce any other non-traded source of risks such as systematic jumps or, even more generally, random jumps.⁵

² See Rubinstein (1994), Jackwerth and Rubinstein (1996), and Das and Sundaram (1997) for a detailed discussion of this empirical regularity.

³ Ball and Roma (1994) show that the implied variance of the BS price when the true price process is subject to stochastic volatility is quadratic in the outness-of-the-money, and that the greatest downward bias occurs for at-the-money options. This suggests that the stochastic volatility option pricing model is consistent with the smile.

⁴ Another related (non-stochastic) approach allows the volatility to depend functionally on the underlying security price. Various alternative proposals for the functional volatility process have been suggested. The well-known constant elasticity of variance model due to Cox and Ross (1976) is the most prominent one.

⁵ See Merton (1976), Bates (1996), and Bakshi et al. (1997).

Recent advances in the literature related to implementable option pricing models that admit stochastic volatility include Stein and Stein (1991), Heston (1993), Bates (1996), Bakshi et al. (1997), and Das and Sundaram (1997). In particular, in his seminal article, Heston (1993) shows that a closed-form solution for a European call can be derived as an integral of the future security price density, which itself may be calculated by an inverse Fourier transform. This method may also be applied when correlation between the increments of the driving Brownian motions of the underlying asset and the volatility is non-zero. Thus, while Hull and White (1987) is an approximation, Fourier inversion methods are potentially more precise. Of course, estimation methods remain quite challenging.

Recently, the papers by Bakshi et al. (1997), and Das and Sundaram (1997) analyze the extent to which models under either stochastic volatility or random jumps are capable of resolving the well-known anomalies related to the BS model. They show that neither class of models is an adequate explanation of the biases found in the empirical literature associated with the BS model. However, in both papers, stochastic volatility models seem to behave slightly better than jumps.

An alternative approach for dealing with nonconstant volatility was suggested by Rubinstein (1994), Jackwerth and Rubinstein (1996) and Jackwerth (1996), and a related series of papers by Derman and Kani (1994), Dupire (1994), Chriss (1995), Derman et al. (1996). Instead of imposing a parametric functional form for volatility, they construct a binomial or trinomial numerical procedure so that a perfect fit with observed option prices is achieved. This procedure captures (by construction) the most salient characteristics of the data. In particular, the implied tree employed in the numerical estimation must correctly reproduce the volatility smile. The most popular models within this family use recombining binomial trees implied by the smile from given prices of European options. Once the appropriate prices and transition probabilities corresponding to the nodes and links of the tree are calculated, any American or path-dependent option may be priced consistently with the market. Also, to eliminate arbitrage opportunities, negative node transition probabilities are not allowed and the branching process must be risk-neutral at each step.

Empirical tests of implied binomial trees have been proposed by Dumas et al. (1996) and Jackwerth (1996). DFW point out that none of the previous studies analyzes the out-of-sample behaviour of the time-varying volatility function obtained by the in-sample implied binomial trees. The key empirical issue becomes the stability of the volatility function. Surprisingly enough, DFW find that the pricing (and hedging) out-of-sample performance of the implied binomial trees is worse than that of an ad hoc BS model with variable implied volatilities. They suggest that the BS model may be perfectly correct, but trading costs combined with option series clientele may produce systematic

patterns in implied volatilities.⁶ The point is that these patterns may have no relation with the distributional characteristics of the underlying asset.

On the other hand, Jackwerth (1996) tests the pricing performance of implied binomial trees, the BS model, and the constant elasticity of variance model. He chooses the parameters of these models to fit the observed prices of longer term options best and then price shorter options with those parameters. In the post-crash period, Jackwerth favours the pricing behaviour of implied binomial trees.

Finally, the papers by Corrado and Su (1996, 1997) contain a related way to cope with the smile effect featured by the option data. It is well understood that volatility smiles are a consequence of empirical violations of the normality assumption in the BS model. In other words, skewness and kurtosis in the option-implied distributions of stock returns are the source of volatility smiles. This is, of course, closely related to stochastic volatility models which can nicely explain the behaviour of option prices in terms of the underlying distribution of returns. In particular, the correlation between the Brownian motions associated with the underlying asset and the volatility affects the skewness of returns, while the volatility of volatility is directly related with kurtosis.⁷ Following this reasoning, Corrado and Su suggest an extended version of the BS model to account for biases induced by nonnormal skewness and kurtosis in stock return distributions. Their valuation formula is given by the sum of the BS option price plus adjustment terms for nonnormal skewness and kurtosis. They find that their adjusted formula yields significantly improved pricing performance for deep in-the-money or deep out-of-the-money options.

Despite the fact that we have theoretical models consistent with the smile pattern across exercise prices, it is also true that the empirical smiles are about twice as large as predicted by theory.⁸ It seems quite clear that something else is going on. In this regard, given the evidence provided by Longstaff (1995) and Dumas et al. (1996), a serious candidate to explain the pronounced pattern of volatility estimates across exercise prices might be related to liquidity and trading costs.

These remarks are at the origin of our research project. What is clearly missing in the extant literature is an analysis of the determinants of the implied volatility function. Surprisingly, none of the papers above has tried to explain directly the determinants of the smile, although this is a relevant issue. Otherwise, we may be missing an important point here; i.e., the reasons be-

⁶ For an alternative discussion of trading costs, see Longstaff (1995).

⁷ See the excellent and intuitive discussion provided by Heston (1993), and the formal analysis of Das and Sundaram (1997).

⁸ See Ghysels et al. (1996).

hind the “apparent” failure of the BS pricing model. Hence, the main objective of this paper is to study directly the determinants of the volatility smile. We employ an extensive database of intraday transaction prices for options on the Spanish IBEX-35 stock exchange index. This is one of the most popular option contracts traded in Europe. Given that we are particularly concerned with trading costs and liquidity effects, it may be relevant to explore alternative option markets which are probably narrower than the fully investigated S&P 100 index options traded at the Chicago Board Options Exchange (CBOE).

Our empirical results show that transaction costs, proxied by bid-ask spreads, and variables related to uncertainty about the return of the underlying asset and to relative market momentum seem to be key aspects regarding the shape of the implied volatility function. Moreover, complex and nonlinear causality effects on the dynamic interrelations between these variables and the volatility smile are also found.

This paper is organized as follows: the next section contains a brief summary of the Spanish options market. The data are described in Section 3. Some previous general results are reported in Section 4. In particular, smile seasonality is briefly discussed in this section. Section 5 presents the empirical results regarding the determinants of the smile volatility. We then conclude with a summary and discussion.

2. The Spanish IBEX-35 index options

The Spanish IBEX-35 index is a value-weighted index comprising the 35 most liquid Spanish stocks traded in the continuous auction market system. The official derivative market for risky assets, which is known as MEFF, trades a futures contract on the IBEX-35, the equivalent option contract for calls and puts, and individual option contracts for blue-chip stocks. Trading in the derivative market started in 1992. The market has experienced tremendous growth from the very beginning. Relative to the volume traded in the Spanish continuous market, trading in MEFF represented 40% of the regular continuous market in 1992, 156% in 1994, and 170% in 1995. The number of all traded contracts in MEFF relative to the contracts traded in the CBOE reached 20% in 1995.

The IBEX-35 option contract is a cash settled European option with trading during the three nearest consecutive months and the other three months of the March–June–September–December cycle. The expiration day is the third Friday of the contract month. Trading occurs from 10:30 to 17:15. During the sample period covered by this research, the contract size is 100 Spanish pesetas times the IBEX-35 index, and prices are quoted in full points, with a minimum

price change of one index point or 100 pesetas.⁹ The exercise prices are given by 50 index point intervals.

It is important to point out that liquidity is concentrated in the nearest expiration contract. Thus, during 1995 almost 90% of crossing transactions occurred in this type of contracts. Finally, it should be noticed that option and futures contracts are clearly associated. The futures contract has exactly the same contract specifications as the IBEX-35 options. This will allow us to employ the futures price rather than the spot price in our empirical exercise by using the Black (1976) model for pricing options on futures. In fact, this is what is usually done by practitioners.

3. The data

Our database is comprised of all call and put options on the IBEX-35 index traded daily on MEFF during the period January 1994–April 1996. Given the concentration in liquidity, our daily set of observations includes only calls and puts with the nearest expiration day. Moreover, we eliminate all transactions taking place during the last week before expiration. In other words, for each monthly expiration date cycle, we only take into account prices for the first three weeks of the cycle.

As usual in this type of research, our primary concern is the use of simultaneous prices for the options and the underlying security. The data, which are based on all reported transactions during each day throughout the sample period, do not allow us to observe simultaneously enough options with the same time-to-expiration on exactly the same underlying security price but with different exercise prices. In order to avoid large variations in the underlying security price, we restrict our attention to the 45-min window from 16:00 to 16:45. It turns out that, on average and during our sample period, almost 25% of crossing transactions occur during this interval. Moreover, care was also taken to eliminate the potential problems with artificial trading that are most likely to occur at the end of the day. Thus, all trades after 16:45 were eliminated so that we avoid data which may reflect trades to influence market maker margin requirements. At the same time, using data from the same period each day avoids the possibility of intraday effects in the IBEX-35 index options market. Finally, we eliminate from the sample all call and put prices that violate the well known arbitrage bounds.

These exclusionary criteria yield a final daily sample of 7947 observations. The implied volatility for each of our 7947 options is estimated next. Note that we take as the underlying asset the average of the bid and ask price quotation

⁹ This has recently been changed to 1000 pesetas.

given for each futures contract associated with each option during the 45-min interval.¹⁰ Recall that we are allowed us to use futures prices given that the expiration day of the futures and option contracts systematically coincides during the expiration date cycle. Moreover, note that dividends are already taken into account by the futures price. To proxy for riskless interest rates, we use the daily series of annualized repo T-bill rates with either one week, two weeks or three weeks to maturity. One of these three interest rates will be employed depending upon how close the option is to the expiration day.

As discussed by French (1984), volatility appears to be a phenomenon that is basically related to trading days. However, interest rates are paid by the calendar day. We therefore employ Black (1976) option pricing formula adjusted by two time measures to reflect both trading days and calendar days until expiration.

We next observe all calls and puts with the same exercise price for each day in the sample and for our 45-min interval. We average all implied volatilities previously estimated for each level of the exercise price available during each daily window. All underlying futures prices associated with each exercise price level are averaged to obtain the corresponding level of the underlying asset associated with each average implied volatility. We define moneyness as the ratio between the exercise price and the average of the futures price relative to each average implied volatility as previously obtained. We can now estimate our daily volatility smile. It should be pointed out that the number of observations within a day may vary according to the number of crossing transactions associated with different exercise prices available for each day. In any case, this procedure reduces our sample to 3016 observations from January 1994 to April 1996. This implies that, on average, we have between 5 and 6 options available for alternative exercise prices during each day.

Fig. 1 presents the representative smiles for the whole period and two consecutive subperiods. We employ five fixed intervals for the degree of moneyness, and compute the median over the alternative subperiods of the implied volatility within that fixed interval. These intervals are given by the following degrees of moneyness: 0.8598–0.9682; 0.9682–0.9913; 0.9913–1.0101; 1.0101–1.0321; 1.0321–1.1875. It is interesting to note that the Spanish market seems to be “smiling” independently of the subperiod employed in the estimation. Moreover, during the second subperiod, the implied volatility is lower than in the first subperiod independently of the degree of moneyness.

¹⁰ It might be that lack of liquidity in the futures market is responsible for the lack of variation in the price of the underlying asset during the 45-min window. However, this is not the case. In fact, the futures market is, at least, as liquid as the spot market in terms of comparable measures of trading volume.

VOLATILITY

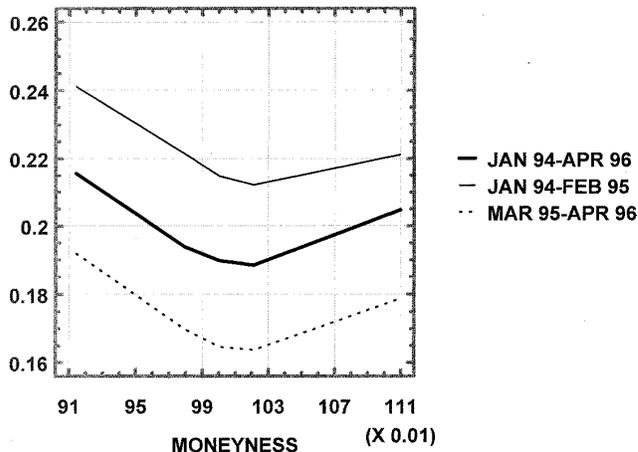


Fig. 1. Average smiles over the sample period.

Finally, Fig. 2 reports similar evidence when the smile is obtained for calls and puts independently. As before, a rather well defined smile seems to be a typical phenomenon in the Spanish options market. However, it should be recognized that a somewhat clearer picture emerges for puts than for calls.

VOLATILITY

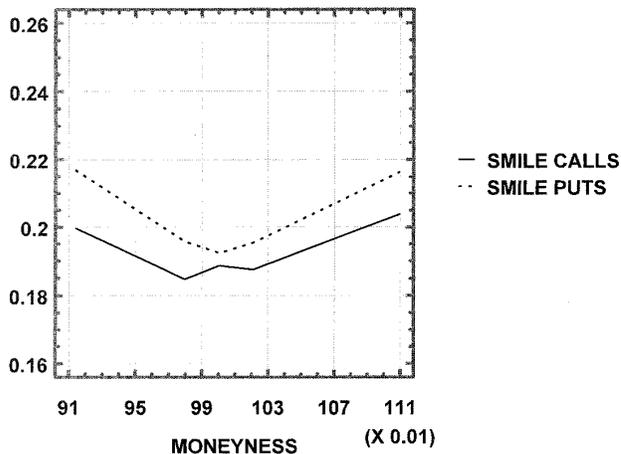


Fig. 2. Average smiles over the sample period for calls and puts.

This generally well behaved smile contrasts with the evidence found in the US market where the typical shape of the volatility function after the 1987 crash is closer to a “sneer”. Formal tests among the two alternatives are performed in the following section.

4. The implied volatility function and smile seasonality

4.1. The implied volatility function: Competing specifications

We next investigate the determinants of the smile. The idea is to estimate the volatility function by fitting the implied volatility through six alternative structural forms:

$$\begin{aligned}
 \text{Model 1 : } \sigma &= b_0 + \varepsilon, \\
 \text{Model 2 : } \sigma &= b_0 + b_1X + \varepsilon, \\
 \text{Model 3 : } \sigma &= b_0 + b_1X + b_2X^2 + \varepsilon, \\
 \text{Model 4 : } \sigma &= b_0 + b_1U + b_2D^2 + \varepsilon, \\
 \text{Model 5 : } \sigma &= b_0 + b_1U + b_2X^2 + \varepsilon, \\
 \text{Model 6 : } \sigma &= b_0 + b_1U + b_2X^2 + b_3D + \varepsilon,
 \end{aligned} \tag{1}$$

where X is the degree of moneyness; this is to say, the exercise price divided by the futures price. Let K be the exercise price and F the average futures price of all options with the same level of the exercise price, then X is equal to K/F . Thus, Model 1 is the volatility function of the BS constant volatility model. Model 2 posits a linear relation between volatility and the degree of moneyness. Model 3 incorporates a quadratic term to capture the typical smile shape. Finally, Models 4–6 employ three different ways of recognizing potential asymmetries in the shape of the volatility function. In particular, Model 4 assumes that the left side of the volatility function is linear on the degree of moneyness, but a quadratic term is necessary to capture some degree of curvature in the right side of the function. Thus, $U = (U_1, \dots, U_n)$ and $D = (D_1, \dots, D_n)$ where,

$$\begin{aligned}
 U_i = X_i & \text{ if } X_i < 1, & D_i = 0 & \text{ if } X_i < 1, \\
 U_i = 0 & \text{ if } X_i \geq 1, & D_i = X_i & \text{ if } X_i \geq 1,
 \end{aligned}$$

where n is the total number of exercise levels for a given day within our 45-min window.

For each day in the sample, we run the regressions given by Eq. (1). Given the number of observations available during each day, not all models can be run for every day. The average adjusted R^2 weighted by the number of observations available for each day within each model is estimated. The results

suggest that Model 3 is the best model at capturing variation in implied volatility attributable to moneyness. This quadratic model explains almost 63% of the variability of implied volatility. It should also be noted that Model 6 explains, in general, as well as the quadratic model, but Model 2, however, explains on average just 48% of the variation of the dependent variable.

The degree of moneyness, X , as specified by Eq. (1) might be more usefully defined as the logarithm of the ratio of the exercise price divided by the futures price.¹¹ Of course, in this way, zero implies being exactly at-the-money, and the regression intercept coefficient corresponds to an at-the-money implied volatility.¹² It is also the case that, under this specification, the regression coefficient for squared moneyness might now be a better descriptor of curvature of the smile. Given this reasoning, we also test the five alternative models of Eq. (1) written as:

$$\begin{aligned}
 \text{Model 1 : } \sigma &= b_0 + \varepsilon, \\
 \text{Model 2 : } \sigma &= b_{01} + b_{11} \ln X + \varepsilon, \\
 \text{Model 3 : } \sigma &= b_{01} + b_{11} \ln X + b_{21}(\ln X)^2 + \varepsilon, \\
 \text{Model 4 : } \sigma &= b_{01} + b_{11}U + b_{21}D^2 + \varepsilon, \\
 \text{Model 5 : } \sigma &= b_{01} + b_{11}U + b_{21}(\ln X)^2 + \varepsilon, \\
 \text{Model 6 : } \sigma &= b_{01} + b_{11}U + b_{21}(\ln X)^2 + b_{31}D + \varepsilon,
 \end{aligned} \tag{1'}$$

where as before, $U = (U_1, \dots, U_n)$ and $D = (D_1, \dots, D_n)$ and

$$\begin{aligned}
 U_i &= \ln X_i & \text{if } \ln X_i < 0, & & D_i &= 0 & \text{if } \ln X_i < 0, \\
 U_i &= 0 & \text{if } \ln X_i \geq 0, & & D_i &= \ln X_i & \text{if } \ln X_i \geq 0.
 \end{aligned}$$

The weighted average adjusted R^2 's obtained under Eq.(1') are basically the same as before. Once again Models 3 and 6 seem to be the best representation of the implied volatility function.

Given that the behaviour of the implied volatility pattern seems to be different in Spain than in the US market, it was decided to run a formal test to compare statistically the different performance of the ‘‘sneer’’ (Model 2) and the smile (Model 3).¹³ In order to investigate this issue, the regression below is estimated by stacking all of the observations and using OLS in the following pooled time-series and cross-sectional procedure:

¹¹ We thank one of the referees for pointing out this alternative specification.

¹² Note that, under the original specification, the at-the-money implied volatility is given by the sum of the three coefficients, b_0 , b_1 , and b_2 . Therefore, the specification given by Eq. (1) also allows for a quick recuperation of the at-the-money implied volatility.

¹³ It should be pointed out that the evidence found in the US market is best described by a ‘‘straight sneer’’. Rubinstein (1994) findings are a good example. In this sense, Model 2 becomes the relevant benchmark.

$$\sigma_{jt} = a_0 + a_1 X_{jt} + a_2 X_{jt}^2 + e_{jt}; \quad j = 1, \dots, J; \quad t = 1, \dots, T, \quad (2)$$

where σ_{jt} is the implied volatility of each option j available during the 45-min window and for each day t in our sample, and X is the degree of moneyness.

The idea is to test whether the coefficient associated with the quadratic term, a_2 , is statistically different from zero. Since our two models are nested, we are able to use a Lagrange Multiplier (LM) statistic that is asymptotically distributed as a chi-squared with one degree of freedom. It turns out that this statistic is 353.29 (p -value = 0.0000). This implies that the estimate of a_2 is statistically different from zero, so we favour Model 3 relative to Model 2.¹⁴

Finally, we also look at call and put options separately. Given that the put-call parity relationship implies that European call and put options of identical moneyness and maturity should have identical implied volatilities, there is no theoretical reason to expect significantly different behaviour between calls and puts. In fact, this turns out to be the case.¹⁵ Model 3 presents the highest average adjusted R^2 weighted by the number of observations available for each day within each model for both call and put options. Moreover, on running regression (2) stacking all available observations, we find that the LM statistic equals to 87.46 (p -value = 0.0000) and 148.16 (p -value = 0.0000) for calls and puts, respectively.

We also run the model under the logarithm specification

$$\sigma_{jt} = a_0 + a_1 \ln X_{jt} + a_2 (\ln X_{jt})^2 + e_{jt}; \quad j = 1, \dots, J; \quad t = 1, \dots, T. \quad (2')$$

The LM statistic is now 346.67 (p -value = 0.0000) when using all calls and puts available in the sample, and 94.04 (p -value = 0.0000) and 137.95 (p -value = 0.0000) for calls and puts, respectively.

It is therefore the case that, independently of the specification employed in the pooled regressions, the estimate of a_2 is statistically different from zero. We definitely favour Model 3 relative to Model 2. It seems certainly the case that the Spanish smile is characterized by a large degree of curvature.

Given these results, we will focus on the coefficients estimated with Model 3. It is important to emphasize that Model 3 is estimated every day in the sample period with enough observations. In other words, to run the corresponding regression for Model 3 every day we need to have enough levels of exercise prices throughout the 45-min interval. In particular for Model 3, using all call and put options at the same time, we have 446 days with enough observations.

¹⁴ It should be noted that Lagrange Multiplier Test \leq Likelihood Ratio Test \leq Wald Test. Thus, we would also reject the null with either of the alternative statistics.

¹⁵ This evidence is consistent with the empirical findings reported by Bates (1991).

Therefore in this case, on a daily basis, we can cross-sectionally estimate 446 coefficients for b_0 , b_1 and b_2 (and b_{01} , b_{11} , b_{21}).

Table 1 reports the average coefficients for both alternative specifications. All average coefficient estimates are statistically different from zero. This confirms the characteristics of the volatility smile intuitively suggested in Fig. 1. The key coefficient is, of course, the one related to the curvature of the smile. Our estimations suggest that the descriptor of the average curvature of the volatility smile is quite similar regardless of the specification used in the regressions. On the other hand, given the strongly U-shaped volatility smile, the analysis in terms of the slope seems meaningless under both specifications. It is true, however, that under the specification of Model 3 given by the equation $\sigma = b_0 + b_1X + b_2X^2$, there is a direct relationship between coefficients b_1 and b_2 . In particular, there is an almost perfectly negative correlation coefficient between b_1 and b_2 . If we take the derivative of Model 3 relative to the degree of moneyness, X , and equate to zero to find the minimum level of X , we note that $X_{\min} = -b_1/2b_2$. It turns out that X_{\min} is very close to one most of the time. That is to say, the minimum implied volatility on a daily basis is generally very close to at-the-money implied volatility. Hence, the estimate of b_1 should be approximately equal to minus two times the estimate of b_2 . This is certainly the case as reflected in the results reported in Table 1. It should be noted that this direct relationship does not hold under the specification of Model 3 given by $\sigma = b_{01} + b_{11}\ln X + b_{21}(\ln X)^2$. The behavior of the coefficient b_{11} is very similar to the slope of the sneer provided by b_1 in Model 2.

Table 1 also confirms the evidence suggested in Fig. 1 regarding the level of the implied at-the-money volatility across subperiods. It is remarkably lower in the second subperiod. Once again, the interpretation of the intercept under the specification with logarithms facilitates the interpretation of the results.

The discussion above suggests that it seems more appropriate to center the analysis around the specification of Model 3 given by $\sigma = b_{01} + b_{11}\ln X +$

Table 1
Daily averages of smile coefficients

Sample	$\sigma = b_{01} + b_{11}\ln X + b_{21}(\ln X)^2$			$\sigma = b_0 + b_1X + b_2X^2$		
	\hat{b}_{01}	\hat{b}_{11}	\hat{b}_{21}	\hat{b}_0	\hat{b}_1	\hat{b}_2
January 94–April 96	0.192	-0.245	4.762	5.384	-10.14	4.943
January 94–February 95	0.217	-0.286	3.403	5.993	-7.412	3.561
March 95–April 96	0.166	-0.202	6.158	6.736	-12.93	6.363

Two alternative specifications of a quadratic model relating implied volatility and moneyness are fitted from January 1994 to April 1996: $\sigma = b_0 + b_1X + b_2X^2$, and $\sigma = b_{01} + b_{11}\ln X + b_{21}(\ln X)^2$, where $X = K/F$, the exercise price divided by the underlying asset value. All calls and puts over the 45-min interval from 16:00 to 16:45 are employed in the estimation. The reported coefficients are the mean of the daily estimates of b_i for both specifications; $i = 0, 1$ and 2 .

$b_{2l}(\ln X)^2$. In fact, this will be our strategy although we will also report some results in terms of the curvature of the smile characterized by b_2 .

To end this preliminary discussion, it should be pointed out that we also synthesized the coefficients that characterized the volatility smile computing the first principal component of the 446×3 matrix of our daily estimates under Model 3 without logarithms. Given that, as discussed above, the correlation of the estimated parameters is very high, the first principal component explains almost 100% of the variability of these series. In particular, the principal component is positively correlated with b_0 and b_2 , and negatively correlated with b_1 . Any of the conclusions we obtain concerning the behavior of b_{2l} (and b_2) are also valid for the principal component. Hence, we will not report any result in terms of the principal component.

4.2. *The implied volatility function: Smile seasonality*

Daily return seasonalities have been a very popular research topic in recent years. Moreover, daily microstructure seasonalities have also been investigated by Foster and Viswanathan (1993) for the US market, and Lehmann and Modest (1994) for the Tokyo Stock Exchange market among others. They conclude that volume is lowest on Mondays, reflecting both reduced demand for liquidity traders who may fear increased adverse selection, and the higher trading costs on Monday since this is the day when bid–ask spreads are clearly largest. In the Spanish continuous auction market, Rubio and Tapia (1996) find both higher bid–ask spreads and less depth on Monday. They conclude that liquidity is unambiguously lower on Monday.

These findings may imply that the smile volatility function does not remain stable throughout all week days. There may be seasonalities in the shape of the volatility smile which may reflect different degrees of liquidity, institutional arrangements or a continuous learning process of market makers throughout the week which may suggest a different implied volatility function at the beginning of the week.

These issues are investigated by running the following regressions:

$$b_{it} = \beta_{\text{MO}}\text{MON}_t + \beta_{\text{TU}}\text{TUE}_t + \beta_{\text{WE}}\text{WED}_t + \beta_{\text{TH}}\text{THU}_t + \beta_{\text{FR}}\text{FRI}_t + \varepsilon_t, \quad (3)$$

where b_{it} is either b_{0it} , b_{1it} or b_{2it} (and b_2) and MON, TUE, WED, THU and FRI are dummy variables for Monday through Friday. The estimates of β_{MO} , β_{TU} , \dots , β_{FR} are the sample means corresponding to each day of the week for the three coefficients of Model 3. Newey–West consistent standard errors with five lags are employed in all estimations. Moreover, in this case the statistic to jointly test seasonalities across week days follows a χ^2 distribution asymptotically, under the null hypothesis.

The results are reported in Table 2. It is clear that, independently of the day of week, all three coefficients b_{0i} , b_{1i} and b_{2i} (and b_2) are, on average, significantly different from zero. Moreover, the magnitude of the degree of curvature of the smile seems quite different from one day to another. In particular, the results suggest that Monday presents a lesser degree of curvature than other

Table 2
Smile daily seasonality

DAYS	$\sigma = b_{0i} + b_{1i} \ln X + b_{2i} (\ln X)^2$			$\sigma = b_0 + b_1 X + b_2 X^2$	No. obs
	\hat{b}_{0i}	\hat{b}_{1i}	\hat{b}_{2i}	\hat{b}_2	
MONDAY	0.190 (47.2)	-0.252 (-9.43)	2.792 (4.55)	2.982 (4.84)	81
TUESDAY	0.191 (44.6)	-0.248 (-7.47)	4.554 (2.82)	4.722 (2.90)	94
WEDNESDAY	0.191 (41.0)	-0.218 (-6.90)	4.347 (5.04)	4.537 (5.23)	87
THURSDAY	0.195 (43.1)	-0.248 (-7.51)	5.461 (5.35)	5.613 (5.56)	92
FRIDAY	0.190 (41.5)	-0.258 (-6.29)	6.404 (5.46)	6.608 (5.57)	92

Two alternative specifications of a quadratic model relating implied volatility and moneyness are fitted from January 1994 to April 1996: $\sigma = b_0 + b_1 X + b_2 X^2$, and $\sigma = b_{0i} + b_{1i} \ln X + b_{2i} (\ln X)^2$, where $X = K/F$, the exercise price divided by the underlying asset value. All calls and puts over the 45-min interval from 16:00 to 16:45 are employed in the estimation. The reported coefficients are the mean of the daily estimates of b_2 for both specifications, and the mean of the daily estimates for both b_{1i} and the implied volatility of the at-the-money options as represented by the intercept of the specification with logarithms. The following regression is run:

$$b_{it} = \beta_{MO} \text{MON}_t + \beta_{TU} \text{TUE}_t + \beta_{WE} \text{WED}_t + \beta_{TH} \text{THU}_t + \beta_{FR} \text{FRI}_t + \varepsilon_t; \quad i = b_{0i}, b_{1i}, b_{2i}, b_2,$$

where MON, TUE, WED, THU, and FRI are dummy variables for Monday through Friday. Newey–West robust standard errors with five lags are employed (t -statistics in parentheses). The use of a consistent covariance matrix implies that the test statistic under the null follows asymptotically a χ^2 distribution.

(1) $\sigma = b_{0i} + b_{1i} \ln X + b_{2i} (\ln X)^2$:

(i) Test statistics for at-the-money implied volatility given by b_{0i} :

$$\chi^2(1) \{ \beta_{MO} = (\beta_{TU} + \beta_{WE} + \beta_{TH} + \beta_{FR})/4 \} = 0.323; \quad p\text{-value} = 0.56981,$$

$$\chi^2(4) \{ \beta_{MO} = \beta_{TU} = \beta_{WE} = \beta_{TH} = \beta_{FR} \} = 2.971; \quad p\text{-value} = 0.56259.$$

(ii) Test statistics for slope as approximated by b_{1i} :

$$\chi^2(1) \{ \beta_{MO} = (\beta_{TU} + \beta_{WE} + \beta_{TH} + \beta_{FR})/4 \} = 0.119; \quad p\text{-value} = 0.72993,$$

$$\chi^2(4) \{ \beta_{MO} = \beta_{TU} = \beta_{WE} = \beta_{TH} = \beta_{FR} \} = 1.305; \quad p\text{-value} = 0.86053.$$

(iii) Test statistics for curvature of the smile represented by b_{2i} :

$$\chi^2(1) \{ \beta_{MO} = (\beta_{TU} + \beta_{WE} + \beta_{TH} + \beta_{FR})/4 \} = 15.288; \quad p\text{-value} = 0.00009,$$

$$\chi^2(4) \{ \beta_{MO} = \beta_{TU} = \beta_{WE} = \beta_{TH} = \beta_{FR} \} = 16.397; \quad p\text{-value} = 0.00253.$$

(2) $\sigma = b_0 + b_1 X + b_2 X^2$:

(i) Test statistics for curvature of the smile represented by b_2 :

$$\chi^2(1) \{ \beta_{MO} = (\beta_{TU} + \beta_{WE} + \beta_{TH} + \beta_{FR})/4 \} = 15.016; \quad p\text{-value} = 0.00011,$$

$$\chi^2(4) \{ \beta_{MO} = \beta_{TU} = \beta_{WE} = \beta_{TH} = \beta_{FR} \} = 16.093; \quad p\text{-value} = 0.00290.$$

week days. In fact, our χ^2 statistic significantly rejects the equality of coefficients associated with the curvature of the smile across all days.¹⁶

The χ^2 statistic, however, does not reject the equality of the coefficients b_{01} and b_{11} . It is interesting to find out that the implied at-the-money volatility does not seem to present significant daily seasonality.

In summary, we may conclude that the Spanish options market smiles differently on Monday than on other week days. More specifically, the curvature of the Spanish smile is statistically different at the beginning of the week relative to the end of the week. This issue is further investigated in the next section of the paper.

5. On the determinants of the implied volatility function

As argued in the introduction, the key issue of this paper is the direct analysis of the reasons explaining the volatility smile. It is important to emphasize that, given that the IBEX-35 option contract is a European option, the pattern of implied volatilities across different exercise prices provides direct evidence of the shape of the risk-neutral density, relative to the lognormal benchmark. Of course, this is because the second derivative of the European call (put) option price with respect to the exercise price is proportional to the appropriate risk-neutral probability density. This argument implies that, in fact, our objective is to explain the true implicit distribution in actual option prices.

Under this line of reasoning, the results from our previous sections suggest that the implicit distribution in the Spanish market is leptokurtic in (both) the right and the left tail of the distribution. This means that out-of-the-money calls (in-the-money puts) and puts (in-the-money calls) which pay off under realizations in the tails are more valuable than predicted by the BS model with its lognormal distribution assumption. An important point of our research is therefore to investigate the characteristics of the (deep) out-of-the-money calls (in-the-money puts) and puts (in-the-money calls).

5.1. Data and preliminary findings

Analysis of the determinants of the volatility smile is based on three categories of economic variables. The economic determinants should include

¹⁶ Alternative formal tests of the difference between the coefficient on Monday and the rest of the week confirm that the curvature on Monday is statistically lower than in the rest of the week. Moreover, it turns out that the difference becomes more and more significant towards the end of the week. Thus, on average, Friday has the highest degree of curvature.

relevant characteristics of the underlying asset, economic variables that help to predict the future stock market, and some characteristics of the options market itself. In particular, violations of a constant implied volatility function may be due to the effects of trading costs or to the degree of options market liquidity. Proxies for these characteristics should be included in the list of relevant variables.

To capture the possibility of market-related effects on option pricing, we include the annualised standard deviation of the IBEX for each day in the sample estimated with minute by minute observations, and the natural log of the number of shares traded (volume) by the components of the IBEX during the 45-min interval for which we have option pricing data. The idea is to incorporate both a measure of uncertainty and a measure of the level of activity in the underlying asset.

Two variables are employed in order to incorporate variables that may help in predicting the market. Both measures reflect the relative market momentum of the underlying economic situation of the Spanish economy. The idea is to construct variables that reflect levels of asset prices. This is obviously somewhat arbitrary. However, there is well-known evidence that suggests some useful instruments in predicting general market conditions and expected returns of risky assets.¹⁷

Our first variable of this type is the log relative Treasury bill rate (RTB) given by the following expression:

$$\text{RTB}_t = \log \frac{r_t}{1/60 \sum_{\tau=t-1}^{t-60} r_\tau}, \quad (4)$$

where r_t is the one week Treasury bill repo rate available at day t . It provides a relative measure of the interest rate levels with respect to its three-month (60 trading days) moving average.

The second variable (MKT) is the log of the ratio of the previous short-run level of the IBEX, given by its three-month moving average, to its current level:

$$\text{MKT} = \log \frac{1/60 \sum_{\tau=t-1}^{t-60} \text{IBEX}_\tau}{\text{IBEX}_t}, \quad (5)$$

where IBEX_t is the level of the value-weighted Spanish stock exchange index at the end of day t .

The underlying justification for including both types of determinant in our analysis (two relevant characteristics of the underlying asset, and two economic variables that help predicting the future stock market,) lies in the possibility of their having path-dependent effects on option pricing. If such effects exist, they

¹⁷ See for example, Keim and Stambaugh (1986) and Campbell (1996).

may impact the market valuation of out-of-the-money calls (in-the-money puts) and puts (in-the-money calls).

The last group of variables which may be relevant in explaining implied volatility patterns across exercise prices is associated with the characteristics of the option market itself. As a measure of transaction costs, we employ the daily average relative bid–ask spread for the options transacted during the 45-min interval. They reflect the market-making costs and adverse selection risks faced by agents participating in the option market. Finally, as a measure of the level of activity in the option market, we include the natural log of the number of option contracts negotiated during the 45-min interval employed in this paper. This provides a reasonable estimate of the general liquidity of the option market.

Before presenting a time-series regression analysis relating the main characteristics of the volatility smile to the variables described above, we must analyze the potential non-stationarities in our chosen variables. Augmented Dickey–Fuller (DF) tests for unit roots are performed for all variables described above and the three time-series of coefficients characterizing the smile over time. Independently of the specification employed in the analysis, the results imply that the log relative Treasury bill repo rate (RTB) is nonstationary, while for the rest of our chosen variables, we are able to reject the existence of a unit-root. In the tests below, we therefore use the first daily differences of the log relative repo rate.

Our first test consists of simple regressions, with Newey–West robust standard errors, relating the variables described previously to either the principal component of the smile or the coefficients themselves. This section of the paper analyzes several factors (potentially) related to the volatility smile, but it does not test for *causes* of the smile. The hypothesis merely involves correlation between the volatility smile and some other variables.

In the regressions below, we also include a dummy variable for Monday, and two other control variables for moneyness and time to expiration. In particular, the average degree of moneyness of all options used in the analysis, and the time-to-expiration of the options employed in our database are taken into account. Note that, for a given day, all options available throughout the 45-min interval have the same time-to-expiration. However, the volatility smile may be changing throughout the life of the options.¹⁸

In order to explain the variability of the principal component and the coefficients which characterize the smile, the following time-series regressions are run:

¹⁸ Given that the degree of moneyness does not have any significant influence on the results, it is not included in the regressions shown in the paper.

$$b_{it} = \beta_0 + \beta_1 \text{MON}_t + \beta_2 \text{MKT}_t + \beta_3 \text{SIGMA}_{t-1} + \beta_4 \text{VMKT}_{t-1} + \beta_5 \text{DRTB}_t + \beta_6 \text{BA}_t + \beta_7 \text{VOPT}_{t-1} + \beta_8 \text{TIME}_t + \varepsilon_t, \quad (6)$$

where:

- b_{it} is any of coefficients (b_{0i} , b_{1i} , b_{2i} and b_{2i}) characterizing the volatility smile throughout the time period employed in the analysis;
- MON is the dummy variable for Mondays;
- MKT is the log of the relative market momentum given by expression (5);
- SIGMA is the annualized standard deviation of the IBEX for each day in the sample estimated by minute by minute observations;
- VMKT is the log of the number of shares traded by the individual stocks conforming the IBEX calculated during the 45-min interval;
- DRTB is the first daily difference of the log relative Treasury bill repo rate given by expression (4);
- BA is the daily average relative bid–ask spread for the options transacted during the 45-min interval considered in the analysis;
- VOPT is the log of the number of options contracts negotiated during the 45-min interval; and
- TIME is the annualized number of days to expiration of the options transacted during the 45-min interval.

The regressions are run with a one-period lag for the volatility variable, and for the volume related variables in the option and stock markets. In principle, it would be desirable to use available information at each day t if we want to make stronger statements about these regressions. However, the results reported are based on the contemporaneous bid–ask spread and the contemporaneous market momentum. It should be pointed out that very similar results are found when we also use a one-period lag for the market momentum variable and the bid–ask spread. Somewhat better statistical fit values are obtained, however, when we run the regression model given by Eq. (6).¹⁹

The results are shown in Table 3. They suggest that the degree of curvature of the volatility smile is positively and significantly related to transaction costs represented by the bid–ask spread. On average, whenever the bid–ask spread tends to increase, the degree of curvature of the volatility smile increases. Alternatively, when market makers tend to face higher adverse selection risks, out-of-the-money calls (in-the-money puts) and out-of-the-money puts (in-the-money calls) are more highly valued by the market relative to the BS model. This is a key result. It suggests that option pricing models with stochastic volatility and stochastic jumps will not be correctly specified as long as we do not allow for transaction costs. The compensation for market maker risks

¹⁹ Stepwise regressions are also employed in deciding the variables and the number of lags to be included in the regressions reported in the paper.

Table 3
The determinants of the implied volatility function

Coefficients associated with:	Smile: intercept (b_{01})	Smile: slope (b_{11})	Smile: curvature (b_{21})	Smile: curvature (b_2)
Intercept	0.139 (6.07)	0.236 (1.01)	2.353 (0.32)	2.617 (0.35)
MONDAY(t)	-0.005 (-1.94)	-0.002 (-0.05)	-1.113 (-1.37)	-1.083 (-1.34)
MKT(t)	0.497 (15.3)	0.510 (1.26)	-16.90 (-1.58)	-18.84 (-1.75)
SIGMA($t-1$)	0.057 (2.19)	-0.009 (-0.03)	-16.44 (-2.18)	-16.56 (-2.17)
VMKT($t-1$)	0.004 (1.48)	-0.047 (-1.67)	1.047 (1.08)	1.067 (1.09)
DRTB(t)	0.173 (1.43)	-0.134 (-0.10)	-58.86 (-1.37)	-55.81 (-1.29)
BA(t)	-0.006 (-0.46)	0.096 (0.63)	12.04 (2.38)	12.35 (2.38)
VOPT($t-1$)	0.002 (1.89)	-0.029 (-2.53)	0.127 (0.44)	0.110 (0.38)
TIME(t)	0.001 (1.04)	0.015 (1.91)	-1.310 (-5.97)	-1.342 (-6.09)

Two alternative specifications of a quadratic model relating implied volatility and moneyness are fitted from January 1994 to April 1996: $\sigma = b_0 + b_1X + b_2X^2$, and $\sigma = b_{01} + b_{11} \ln X + b_{21}(\ln X)^2$, where $X = K/F$, the exercise price divided by the underlying asset value. All calls and puts over the 45-min interval from 16:00 to 16:45 are employed in the estimation. We have 446 daily observations available. Time-series regressions are run to explain the variability of the coefficients which characterized the volatility smile. The following regressions are run:

$$b_{it} = \beta_0 + \beta_1 \text{MON}_t + \beta_2 \text{MKT}_t + \beta_3 \text{SIGMA}_{t-1} + \beta_4 \text{VMKT}_{t-1} + \beta_5 \text{DRTB}_t + \beta_6 \text{BA}_t + \beta_7 \text{VOPT}_{t-1} + \beta_8 \text{TIME}_t + \varepsilon_t; \quad i = b_{01}, b_{11}, b_{21}, b_2.$$

where MON is a dummy variable for Monday, MKT is the logarithm of the ratio of the previous short-run level of the IBEX (three-month moving average) to its current level, SIGMA is the annualized standard deviation of the IBEX for each day in the sample estimated by minute by minute observations, VMKT is the logarithm of the number of shares traded by the components of the IBEX calculated during the 45-min interval, DRTB is the first daily difference of the log relative (with respect to its three-month moving average) treasury bill rate, BA is the daily average relative bid-ask for the options available during the 45-min interval, VOPT is the logarithm of the number of option contracts negotiated during the 45-min interval, and TIME is the annualized number of days to expiration of the options available in the sample. Newey–West robust standard errors are employed (t -statistics in parentheses). The (average) adjusted R^2 is 0.23.

seems to be playing a crucial role in the behavior of the Spanish options market.

It is also the case that the degree of curvature is negatively and significantly related to the historical volatility of the underlying asset, and to time to expiration. Options with short times to expiration tend to have a higher degree of

curvature in the implied volatility pattern across exercise prices. Moreover, high volatility periods tend to be associated with lower curvature of the smile.

Finally, the relative momentum of the market seems to be weakly related to the degree of curvature of the smile. Whenever the current level of the stock market is better than the past, we find that, on average, the degree of curvature of the smile increases.

The results regarding the correlations between market momentum, historical volatility and the shape of the volatility smile suggest that periods which are relatively calm but at the same time have increasing current levels of the market stock exchange index tend to be associated with a higher degree of curvature of the volatility smile. Alternatively, the pattern across exercise prices becomes flatter whenever the volatility of the underlying asset goes up, and the relative market momentum gets worse. At these periods of time, out-of-the-money puts (in-the-money calls) and in-the-money puts (out-of-the-money calls) become more symmetrically valued by the market relative to at-the-money options.

To finish the discussion of this preliminary evidence, it should be pointed out that the Monday dummy variable relative to the descriptor of curvature does not seem to be significant once other variables are taken into account by the analysis.

In summary, we may conclude that transaction costs influence the valuation of out-of-the-money puts (in-the-money calls), and in-the-money puts (out-of-the-money calls) relative to at-the-money options. Higher transaction costs are associated with higher market values of extreme (in term of moneyness) options. These costs might be an explanation for the apparent failure of BS to explain the behavior of out-of-the-money puts (in-the-money calls) and in-the-money puts (out-of-the-money calls). At the same time, the degree of uncertainty and relative market momentum also seem to be relevant factors associated with the shape of the volatility smile. These factors, however, are surely associated with more traditional explanations along the line of stochastic volatility and stochastic jumps. Market conditions should be related to skewness and kurtosis effects on option prices. On the one hand, the correlation between volatility and the spot asset's price is a key issue for explaining skewness. In particular, negative skewness which is consistent with asymmetric GARCH effects found in the IBEX-35 index by León and Mora (1996) might be the explanation of our preliminary evidence regarding the correlation between the curvature of the smile, the historical volatility and the relative market momentum.²⁰ Note that our results are consistent with a negative relation

²⁰ Negative skewness or higher downside volatility might be explained either by the well known leverage effects or by wealth effects. The latter consists of economic agents becoming more risk averse as prices (wealth) go down. Hence, the arrival of new information causes a greater reaction among agents, so that volume of trading and volatility increase.

between volatility and the current level of the underlying asset. On the other hand it may be the case that, in terms of stochastic volatility models, the volatility of volatility increases whenever the current market conditions improve relative to the past. This would have the effect of increasing the tails of the underlying distribution and, therefore, it would be consistent with a higher and symmetric curvature in the volatility smile.

Table 3 also shows how the at-the-money implied volatility increases as long as historical volatility becomes higher, and the relative market momentum is stronger. This is again consistent with negative skewness effects in the Spanish market. Finally, it should be pointed out that the level of activity in the options market, as measured by the number of option contracts negotiated, is positively related to the at-the-money implied volatility, and negatively correlated with the coefficient b_{11} . If we interpret b_{11} as the (average) slope of the smile, the implication would be that the slope of the smile increases with volume. Given that volume is not significantly associated with the curvature of the smile, we may conclude that a higher volume in the option market gives more value to out-of-the-money puts (in-the-money calls) relative to the values of the in-the-money puts (out-of-the-money calls).

5.2. Linear Granger causality tests and the volatility smile

The general idea behind causality tests is that they can provide useful information on whether knowledge of past values of the variables employed in the previous section improves short-run forecasts of current and future variability on the shape of the volatility smile. In this section, we employ traditional Granger tests to investigate the presence of linear predicting power between the variables discussed previously and the shape of the volatility smile.

Let us assume that we observe two stationary and ergodic time series, X_t , and Y_t . Let $F(X_t | \mathbf{Z}_{t-1})$ be the conditional probability distribution of X_t given a bivariate information set \mathbf{Z}_{t-1} . This information set is formed with the Lx -length vector of past values of X_t , \mathbf{X}_{t-Lx} , and the Ly -length vector of past values of Y_t , \mathbf{Y}_{t-Ly} . Given these lags, the series, Y_t , does not strictly Granger cause X_t if

$$F(X_t | \mathbf{Z}_{t-1}) = F(X_t | (\mathbf{Z}_{t-1} - \mathbf{Y}_{t-Ly})); \quad t = 1, 2, \dots \quad (7)$$

Alternatively, of course, if the equality (7) does not hold, then knowledge of past values of Y_t is useful in predicting current and future values of X_t , and therefore Y is said to strictly Granger cause X .²¹

²¹ When the bivariate information set includes the current value of Y , we have the concept of instantaneous Granger causality.

In order to implement this test, we estimate the following bivariate vector autoregression (VAR) model:

$$\begin{aligned} b_{21t} &= \alpha + A_{11}(L)b_{21t} + A_{12}(L)DET_t + U_t, \\ DET_t &= \beta + A_{21}(L)b_{21t} + A_{22}(L)DET_t + W_t, \end{aligned} \tag{8}$$

where, α and β are two constant terms, b_{21t} is the descriptor of the curvature of the volatility smile,²² and DET represents the variables employed in the previous section of the paper. In particular, DET can be one of the following variables: the log relative market momentum (MKT), the annualized historical volatility of the underlying asset (SIGMA), the log of the number of shares traded in the underlying asset (VMKT), the first differences of the log relative Treasury bill repo rate (DRTB), the log of the number of contracts negotiated in the option market (VOPT), and the average relative bid–ask spread of the options transacted in our 45-min interval (BA). Moreover, $A_{11}(L)$, $A_{12}(L)$, $A_{21}(L)$, and $A_{22}(L)$ are lag polynomials of the same order in the lag operator L , and the residuals U_t and W_t are assumed to be mutually independent and individually i.i.d. variables with zero mean and constant variance.

To test for linear Granger causality from DET (MKT, SIGMA, . . . ,BA) to the curvature component or, alternatively, to the set of coefficients characterizing the smile under both specifications of Model 3, a standard joint F test of exclusion restrictions is carried out to determine whether lagged values of DET have significant linear predicting power for the smile coefficients. The appropriate number of lags is determined in each case on the basis of four alternative information criteria: the Akaike information criterion, the Schwarz specification test, the final prediction error criterion, and the Hannan–Quinn test. When conflicts are found, the Akaike criterion is employed.²³ The null hypothesis that DET does not Granger cause the smile coefficients is rejected if the coefficients on the elements in $A_{12}(L)$ are jointly significantly different from zero. When feedback causality exists, then the coefficients on the elements in both $A_{12}(L)$ and $A_{21}(L)$ are jointly different from zero.

Table 4 reports the results of testing linear Granger causality between the curvature of the smile described by b_{21t} , the average relative bid–ask spread (BA) and the rest of variables described above.²⁴ By analyzing rejections of the null hypothesis of Granger linear noncausality at the 5% level, our tests indicate that there is clear unidirectional causality from transaction costs, proxied

²² We also run the VAR model using the specification of Model 3 without logarithms (b_2). The same conclusions regarding causality are obtained in both approaches.

²³ Similar results across all criteria are generally obtained for all variables used in the analysis. In each case, the number of lags is always the same for the curvature component of the smile and DET (MKT, SIGMA, . . . ,BA).

²⁴ The same results are found when using White standard errors and a χ^2 test of exclusion.

Table 4

Linear Granger causality test results between the curvature of the volatility smile and the underlying market conditions, option activity, and the relative bid–ask spread: January 1994–April 1996

	Length	Stat- F	p -value
H_0 : BA does not cause the curvature of the volatility smile (b_{21})	3	3.486	0.017
H_0 : The curvature of the volatility smile (b_{21}) does not cause BA	3	9.246	0.000
H_0 : MKT does not cause the curvature of the volatility smile (b_{21})	2	0.321	0.726
H_0 : The curvature of the volatility smile (b_{21}) does not cause MKT	2	0.573	0.564
H_0 : SIGMA does not cause the curvature of the volatility smile (b_{21})	4	1.074	0.371
H_0 : The curvature of the volatility smile (b_{21}) does not cause SIGMA	4	0.602	0.662
H_0 : VMKT does not cause the curvature of the volatility smile (b_{21})	2	0.233	0.792
H_0 : The curvature of the volatility smile (b_{21}) does not cause VMKT	2	0.283	0.754
H_0 : DRTB does not cause the curvature of the volatility smile (b_{21})	3	1.556	0.201
H_0 : The curvature of the volatility smile (b_{21}) does not cause DRTB	3	2.159	0.093
H_0 : VOPT does not cause the curvature of the volatility smile (b_{21})	3	1.461	0.226
H_0 : The curvature of the volatility smile (b_{21}) does not cause VOPT	3	2.119	0.098

This table reports the results of the linear Granger causality test based on the following bivariate VAR model:

$$b_{21t} = \alpha + A_{11}(L)b_{21t} + A_{12}(L)DET_t + U_t,$$

$$DET_t = \beta + A_{21}(L)b_{21t} + A_{22}(L)DET_t + W_t,$$

where DET represents the following variables: BA is the daily average relative bid–ask for the options available during the 45-min interval, MKT is the logarithm of the ratio of the previous short-run level of the IBEX (three-month moving average) to its current level, SIGMA is the annualized standard deviation of the IBEX for each day in the sample estimated by minute by minute observations, VMKT is the logarithm of the number of shares traded by the components of the IBEX calculated during the 45-min interval, DRTB is the first daily difference of the log relative (with respect to its three-month moving average) treasury bill rate, VOPT is the logarithm of the number of option contracts negotiated during the 45-min interval, and b_{21} is the curvature of the volatility smile estimated from the following model: $\sigma = b_{01} + b_{11} \ln X + b_{21} (\ln X)^2$. The results are based on exclusion tests relative to an $F(q, T - k)$ where q is the number of excluded (lags) variables and $T - K$ is the number of observations minus the number of independent variables. Hence, the p -value denotes the marginal significance level of the computed F -statistic used to test the zero restrictions implied by the null hypothesis of Granger noncausality.

Table 5

Average transaction costs across five fixed intervals for the degree of moneyness: January 1994–April 1996^a

Classification	Moneyness	Relative bid–ask spread
Deep OMP (IMC)	0.8598–0.9682	0.3333
OMP (IMC)	0.9682–0.9913	0.1946
AMP (AMC)	0.9913–1.0101	0.1616
IMP (OMC)	1.0101–1.0321	0.2166
Deep IMP (OMC)	1.0321–1.1875	0.3668

This table reports the average transaction costs (relative bid–ask spreads) across five fixed intervals for the degree of moneyness. We define moneyness as the ratio between the exercise price level and the average of the future price relative to each average implied volatility.

^a (OMP) is out-of-the-money puts; (IMC) in-the-money calls; AMP (AMC) at-the-money puts (calls); (IMP) in-the-money puts; (OMC) out-the-money calls.

by the bid–ask spread, to the curvature of the smile. Hence, the bid–ask spread does Granger cause the curvature of the volatility smile.²⁵ At the same time it should be pointed out that the shape of the volatility smile also Granger causes transaction costs. Bi-directional causality is therefore found between transaction costs and the volatility smile.

What is really striking from the analysis of Table 4 is that Granger non-causality from the alternative variables to the curvature of the smile cannot be rejected at the 5% significance level. None of the other variables used previously in our simple regressions seems to be causing the curvature of the smile.

These results suggest that transaction costs, represented by the average bid–ask spread, are a key determinant of the shape of the implied volatility function. A quick way of checking the consistency of these results is reported in Table 5. This is not a formal test, but it provides an intuitive explanation of the results found in the paper. The table employs the five fixed intervals for the degree of moneyness used throughout the paper. The average of the relative bid–ask spread within each of the fixed intervals is calculated. As expected, given the empirical evidence reported in this paper, the extreme options (in terms of moneyness) have the highest bid–ask spreads. In other words, deep out-the-money (in-the-money) options have the highest transaction costs, while the at-the-money options present the lowest transaction costs. This pattern of transaction costs seems to be reflected in the pricing of options. They are precisely the options most highly valued on average relative to the BS model. It may easily be the case that these higher transaction costs reflect higher adverse selection costs faced by market-makers when negotiating these options.

As previously mentioned, smile patterns are consistent with leptokurtic distributions. However, as pointed out by Bates (1996), extremely high values

²⁵ This is also the case when we analyze instantaneous linear Granger causality.

of the volatility of volatility are necessary to obtain implicit leptokurtosis of a magnitude consistent with the empirically observed volatility smile. Our results suggest that the missing variable to explain the actual pattern of implied volatility across exercise prices is proxied by the bid–ask spread. Therefore, on the one hand, we have a theoretical justification for the smile – the volatility of volatility or leptokurtosis – and on the other hand, transaction costs seem to be the ultimate reason behind the actual magnitude of the smiles. In any case, it should be recognized that the linear causality found in the tests also runs in the opposite direction. In other words, the fact that these extreme options are more highly valued seems also to Granger cause higher transaction costs.

Finally, similar analyses are also performed for the rest of the coefficients characterizing the smile. It is interesting to mention that the relative market momentum unidirectionally does Granger cause the at-the-money implied volatility (b_{01}). Moreover, volume in the options market also causes the *slope* of the smile (b_{11}).²⁶ However, Granger noncausality from the alternative variables to b_{01} and b_{11} cannot be rejected at the 5% significance level.

There is increasing interest in the study of non-linearities in the dynamic interrelations between financial time series. The point is that on the removal of linear predictive power with a linear VAR model of Eq. (8), there may be a remaining incremental predictive power of one residual series to another. In this case, this incremental predictive power is considered to be nonlinear.

In order to check for the presence of nonlinear causality among our variables and the curvature of the smile (b_{21}), we follow the modified version of Baek and Brock (1992) nonlinear Granger causality tests as suggested by Hiemstra and Jones (1994). Let us consider two strictly stationary and weakly dependent time series U_t and W_t .²⁷ What we test under nonlinear causality is whether the conditional probability that two m -length lead vectors of U_t are within a distance $\delta > 0$ of each other, given that the Lu -length lag vectors of U_t and Lw -length lag vectors of W_t are within δ of each other is the same as the conditional probability that two m -length lead vectors of U_t are within a distance δ of each other, given that the Lu -length lag vectors of U_t are within δ of each other. In the implementation of this test²⁸, the scale parameter, δ , is chosen to be either $\delta = 1.5\sigma$ or $\delta = 0.5\sigma$, where $\sigma = 1$ is the standard deviation of the standardized residuals of Eq. (8).

Interestingly, the results are not robust to alternative scale parameters. For a δ of 1.5, there seem to be no clear signs of nonlinear causality. The null hypothesis of noncausality cannot be rejected when evaluated with right-tailed critical values of the asymptotic $N(0, 1)$ distribution. If anything, there is some

²⁶ Note that, in these two cases, bi-directional causality is not found.

²⁷ In the application, these series correspond to the residuals of the VAR model of Eq. (8).

²⁸ The details are available upon request.

slight evidence of unidirectional non-linear Granger causality between market momentum and the curvature of the smile.

On the other hand, for $\delta = 0.5$, independently of the number of lags assumed in the estimation for $Lu = Lw$, we find seemingly strong evidence of nonlinear Granger causality from the bid-ask spread, volume in both the underlying asset and the option market, market momentum, to the curvature of the smile. It is also the case that for shorter lags there seems to be evidence of nonlinear causality from the volatility of the underlying asset and even the relative level of the interest rates to the curvature of the smile.

Unfortunately, the interpretation of the results is rather complicated. It may be the case that, for this particular application, the nonlinear predictive power improves significantly when we use not only the variable itself, but also other variables such as the bid-ask spread or market momentum, as long as the required distance in the tests becomes smaller. In any event, our results suggest that the dynamic interrelations between the implied volatility function and economic variables such as transaction costs or market momentum and other relevant variables such as the volatility of the underlying asset are, to a certain extent, non-linear. Future research should probably be concentrated on using non-linear theoretical mechanisms when developing models of microstructure dynamic interrelations between information flow and the pricing of extreme (in terms of moneyness) options.

6. Conclusions

To the best of our knowledge, this work analyzes for the first time the underlying determinants of the well known pattern of implied volatilities across exercise prices for otherwise identical options, i.e., the so called volatility smile. We employ a database comprised of all call and put options on the IBEX-35 Spanish index traded daily during the 45-min interval from 16:00 to 16:45 from January 1994 to April 1996. Contrary to the US market, where the smile is rather a (straight) sneer, we find that the Spanish market tends to smile consistently throughout the sample period.

In order to understand the behavior of the implied volatility function, formal tests are performed under a simple regression framework together with more sophisticated techniques of both linear and nonlinear Granger causality tests. Our results suggest a strong seasonal behavior in the curvature of the volatility smile. However, this seasonality tends to disappear when we include several economic variables in the analysis. In particular, transaction costs proxied by the bid-ask spread of the negotiated options, the volatility of the underlying asset, time to expiration and relative market momentum seem to be key variables in explaining the variability of the implied volatility function over time. Both transaction costs and market conditions, play a simultaneous role in

explaining the shape of the implied volatility pattern across exercise prices. Taking into account the significant and inverse relationship between time to expiration and degree of curvature, we are tempted to conclude that market conditions and transaction costs are relatively more important whenever there is a short way to go in the life of the option.

Linear causality tests point to bidirectional linear Granger causality between the shape of the smile and transaction costs. No other economic variable seems to linearly cause the curvature of the smile. This is a key result in our paper. It suggests that illiquidity costs are a crucial determinant of the magnitude we observe in the volatility smile. There is a feedback between transaction costs and the curvature of the smile. Option pricing models with both stochastic volatility and stochastic jumps may not be sufficient to explain well known biases found under the BS framework. In fact, this is the summarizing conclusion reached by Bakshi et al. (1997). Of course, to explicitly model transaction costs on option pricing models with stochastic volatility and/or stochastic jumps is a challenging task that is left for future research.

Acknowledgements

We would like to thank José Luis Serna and Ainhoa Zárrega for capable computational assistance, and Rafael Salinas of BBV Interactivos for helpful conversations. We have received valuable comments from Lars Norden, Richard Stapleton, Stephen Taylor and two anonymous referees, as well as seminar participants at the sixth Annual Meeting of the European Financial Management Association, Istanbul, the 25th Annual Meeting of the European Finance Association, Fontainebleau, and I Foro de Segovia-Finanzas. Partial financial support was provided by Dirección General Interministerial Científica y Técnica (DGICYT) grant PB95-0298 (Ignacio Peña and Gregorio Serna) and Dirección General Interministerial Científica y Técnica (DGICYT) grant PB94-1373 (Gonzalo Rubio). Ignacio Peña and Gonzalo Rubio would also like to acknowledge the financial support provided by Fundación BBV. The contents of this paper are the sole responsibility of the authors.

References

- Baek, E., Brock, W., 1992. A general test for non-linear Granger causality, Working paper, Iowa State University and University of Wisconsin, Madison.
- Bakshi, G., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. *Journal of Finance* 52, 2003–2049.
- Ball, C., Roma, A., 1994. Stochastic volatility option pricing. *Journal of Financial and Quantitative Analysis* 29, 589–607.

- Bates, D., 1991. The crash of '87: Was it expected? The evidence from options markets. *Journal of Finance* 46, 1009–1044.
- Bates, D., 1996. Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options. *Review of Financial Studies* 9, 69–107.
- Black, F., 1976. The pricing of commodity contracts. *Journal of Financial Economics* 3, 167–179.
- Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–659.
- Campbell, J., 1996. Understanding risk and return. *Journal of Political Economy* 104, 298–345.
- Chriss, N., 1995. How to grow a smiling tree, Working paper. Harvard University, Department of Mathematics, Cambridge, MA.
- Corrado, C., Su, T., 1996. Skewness and kurtosis in S&P 500 index returns implied by option prices. *Journal of Financial Research* 19, 175–192.
- Corrado, C., Su, T., 1997. Implied volatility skews and stock return skewness and kurtosis implied by stock option prices. *European Journal of Finance* 3, 73–85.
- Cox, J., Ross, S., 1976. The valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3, 145–160.
- Das, S., Sundaram, R., 1997. Of smiles and smirks: A term-structure perspective, Working paper, Harvard University, Cambridge, MA. December.
- Derman, E., Kani, I., 1994. Riding on a smile. *Risk* 7, 32–39.
- Derman, E., Kani, I., Chriss, N., 1996. Implied trinomial trees of the volatility smile. *Journal of Derivatives* 3, 7–22.
- Dumas, B., Fleming, J., Whaley, R., 1996. Implied volatility functions: Empirical tests, Working paper 5500. National Bureau of Economic Research, Inc., Cambridge, MA.
- Dupire, B., 1994. Pricing with a smile. *Risk* 7, 18–20.
- Foster, D., Viswanathan, S., 1993. Variations in trading volume, return volatility, and trading costs: Evidence on recent price formation models. *Journal of Finance* 48, 187–211.
- French, D., 1984. The weekend effect on the distribution of stock prices: Implications for option pricing. *Journal of Financial Economics* 13, 547–559.
- Ghysels, E., Harvey, A., Renault, E., 1996. Stochastic volatility. In: Maddala, G.S., Rao, C.R. (Eds.), *Statistical Methods in Finance*, Handbook of Statistics 14, North Holland.
- Heston, S., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* 6, 327–344.
- Hiemstra, C., Jones, J., 1994. Testing for linear and nonlinear Granger causality in the stock price-volume relation. *Journal of Finance* 49, 1639–1664.
- Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatilities. *Journal of Finance* 42, 281–300.
- Jackwerth, J.C., 1996. Implied binomial trees: Generalizations and empirical tests. Working paper RPF-262. University of California at Berkeley, Berkeley, CA.
- Jackwerth, J.C., Rubinstein, M., 1996. Recovering probability distributions from option prices. *Journal of Finance* 51, 1611–1631.
- Keim, D., Stambaugh, R., 1986. Predicting returns in the stock and bond markets. *Journal of Financial Economics* 17, 357–390.
- Lehmann, B., Modest, D., 1994. Trading and liquidity on the Tokyo Stock Exchange: A bird's eye view. *Journal of Finance* 49, 951–984.
- León, A., Mora, J., 1996. Modelling conditional heteroskedasticity: Application to the IBEX-35 stock-return index, Working paper AD 96-11, Instituto Valenciano de Investigaciones Económicas, Universidad de Valencia, Valencia.
- Longstaff, F., 1995. Option pricing and the martingale restriction. *Review of Financial Studies* 8, 1091–1124.
- Merton, R., 1976. Option pricing when the underlying stock returns are discontinuous. *Journal of Financial Economics* 4, 125–144.

- Rubinstein, M., 1994. Implied binomial trees. *Journal of Finance* 49, 771–818.
- Rubio, G., Tapia, M., 1996. Adverse selection, volume and transactions around dividend announcements in a continuous auction system. *European Financial Management* 2, 39–67.
- Stein, E., Stein, J., 1991. Stock price distributions with stochastic volatility: An analytical approach. *Review of Financial Studies* 4, 727–752.