Panel Regression of Out-of-the-Money S&P 500 Index Put Options Prices

Prakher Bajpai*

(May 8, 2014)

1 Introduction

In 1973, two economists, Myron Scholes and Fischer Black, developed a mathematical model to correctly price European style options.[1] These options were priced using a simple probabilistic model that assumed that over a large period of time the asset returns were normally distributed with mean, μ, and standard deviation (or volatility), σ. The μ is the return obtained through the risk free rate, normally assumed to be the U.S. Treasury rate. And, once the σ is known one should be able to calculate the expected return of the asset. Then assuming that there are no opportunities to earn a riskless profit over the risk-free rate one could come up with the "fair" price for the option. Thus, one of the key assumptions of the Black-Scholes model is the constant volatility for options with different strike prices.[2]

However, in the post-1987 crash world, the constant volatility assumption for the options prices appeared to have been violated. The options prices for options with strike rates not equal to the present expected value of the underlying asset appeared to be much greater than the theoretically obtained options prices through the Black-Scholes model. One reasoning was that the investors now knew that the stock returns, which were assumed to be normally distributed were not actually normally distributed and instead had fat tails. Another explanation is that after the 1987 crash, investors became permanently wary of the stock market and the options underwriters charged a risk premium in addition to the "fair price." Regardless of why the Black-Scholes model failed, what we do know for certain is that a better model would be needed to explain the prices. While I hold no such lofty targets in this paper so as to come up with a whole new model to price the options, what I do attempt to do is explain empirically the factors that have affected the S&P 500 index European-style out-of-the-money put options prices, i.e. we will be attempting to explain the out-of-the-money put option prices with respect to the at-the-money put options [3][4]

2 Data

The primary source of data for this report is ivolatility.com, a third-party market data vendor for Chicago Board Options Exchange. The panel data consists of monthly Black-Scholes implied volatilities for $25\%, 20\%, 15\%, 10\%, 5\%,$ and 0% out-of-the-money S&P 500 put options from January 1990 to July 2013 with a time period of 3 months, i.e. options

[1] Options are derivatives that give the buyer the option or the right to either buy or sell the underlying asset at a predetermined price either before or at the option's expiration date. European style options are options that can only be exercised upon the option's expiration date and not before.

[2] Strike price is the predetermined exercise price for the underlying asset in the option.

[3] Put options refer to options which only allow the holder to sell the underlying asset.

[4] At-the-money refers to options that have the same strike price as the present value of the underlying asset, and out-of-the-money put options refer to options with strike price less than the present value of the underlying asset.

expiring in 3 months. [5] The 0% out-of-the-money, or otherwise known as at-the-money, options' implied volatilities were used as a substitute for "actual" or realised volatility. The reasoning for using the at-the-money volatility as the "actual" volatility is the assumption that since at-the-money options are the least risky options, the risk premium for such options will not be as big and the implied volatility will act as a fair approximation for the realised volatility. As such, in our analysis, we will not be including the at-the-money datapoints, and the out-of-the-money option prices will be represented as a fraction of the corresponding out-of-the-money put option price. Similarly, the at-the-money strike prices also represent the value of the underlying asset, S&P 500 index.

$$O_P(S) = N(-a_2)K\exp(-r(3)) - N(-a_1)S, \quad (1)$$

where O_P is the put options price, S the present price of the underlying asset, K the strike price, r the risk-free rate, σ the implied volatility, and

$$a_1 = \frac{1}{\sqrt{3}\sigma}\left(\log(S/K) + 3(r+\sigma^2/2)\right) \quad (2a)$$

$$a_2 = \frac{1}{\sqrt{3}\sigma}\left(\log(S/K) + 3(r-\sigma^2/2)\right). \quad (2b)$$

The put option prices for our regression themselves were calculated using equation (1), the Black-Scholes equation for put options with a time period of 3 months, in *Mathematica*. The dataset also includes 3-month London Interbank Offer Rate (LIBOR) to represent the risk-free interest rate.

3 Methodology and Results

We appear to have one outlier for the option price in the 95% out-of-the-money (OTM) put option (priceoveratm) at t=146 or February, 2002 in figure 1. However, since it is just a single outlier it is not of much concern.

Now, when regressing the priceoveratm data we either have the option to choose a fixed effect model or a random effect model. First let us choose the fixed effect model with and without an autoregressive term. The results are shown in figure 2. In fixed effect we assume that the time-invariant characteristics are unique to the strike price. The only variable in both the regressions that appears to be statistically significant at the 1% significance level is the at-the-money volatility. In addition to that the AR(1) term coefficient is not statistically significant and both the AIC and the BIC values are lower for the simple linear regression (please see Appendix A).

However, there is no reason for us to believe that our model presents time-invariant characteristics unique to the strike prices and that the variation across the strike prices is not random. However, the random effect regression also produces the same coefficients and the standard error as shown in 3.

	priceoveratm
atmstrike	0.000
	(1.55)
atmivol	1.162
	(15.21)**
libor	-0.242
	(1.07)
_cons	0.000
	(0.00)
N	1,415

* $p<0.05$; ** $p<0.01$

Figure 3: A table with the coefficients assumed to have a random effect.

Normally, the Hausman test with the null hypothesis that the model used should be the random effects model would tell us which model should be preferred. However, since both the standard errors and coefficients for the two regressions are the same the p-value for the Hausman test will be 1.0 indicating that the random effects model should be used. Thus, we move ahead with the random effects model.

[5] Black-Scholes Implied volatility refers to the "volatility" obtained by solving for volatility in the Black-Scholes model using the actual options price. Similarly, $25\%, 20\%, 15\%, 10\%, 5\%$, and 0% out-of-the-money put option refers to options with strike prices of $75\%, 80\%, 85\%, 90\%, 95\%$, and 100% of the present value of the underlying asset.

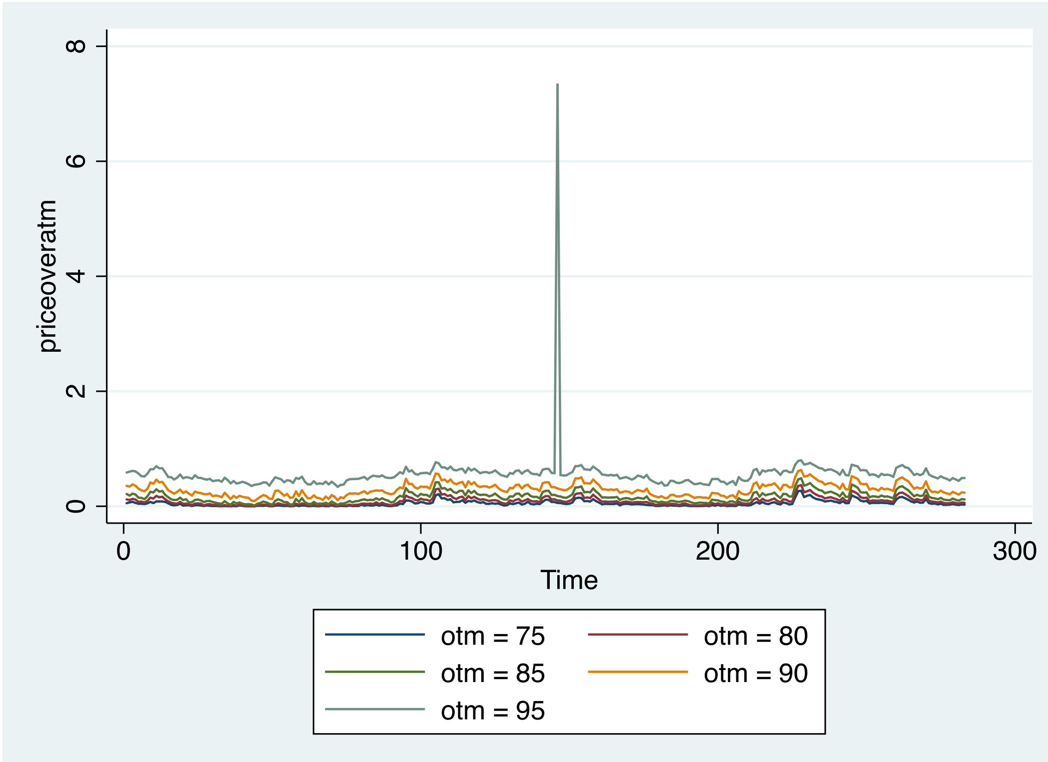

Figure 1: A timeplot of options price as a fraction of at-the-money option price for different strike prices as a % of the underlying asset's price.

We can also test to see whether a simple OLS regression can be used instead of a random effects regression using a Breusch-pagan Lagrange Multiplier test. The null hypothesis is that the variance across the different out-of-the-money strikes is zero, i.e. a simple OLS regression would be sufficient or that there are no panel effects, and quickly find that p value is almost 0 allowing us to reject the null hypothesis and choosing the random effects model as shown in fig. 4.

Also, in the fixed effects model we find evidence of groupwise heteroskedasticity when using a modified Wald test as shown in figure 5. Thus, our regressions also suffer from heteroskedasticity and random effects model with robust errors is preferrable to the fixed effects model. This proves to be a little problematic because in our model a fixed effect regression would make more sense. We would expect there to be some sort of time invariant characteristics specific to the strike prices which allows the options

```
priceoveratm[otm,t] = Xb + u[otm] + e[otm,t]

Estimated results:
                         |       Var     sd = sqrt(Var)
                ---------+-----------------------------
               priceov~m |   .0739596       .2719552
                       e |   .0338664       .1840282
                       u |    .042506       .2061699

Test:   Var(u) = 0
                             chibar2(01) = 50154.10
                          Prob > chibar2 =   0.0000
```

Figure 4: Results from the Breusch and Pagan Lagrange Multiplier test for random effects.

prices to differ so significantly even when everything else is the same. We, thus, move ahead with finding alternative regression models.

But before we do, we also perform a serial correlation test for the presence of autocorrelation in panel data. We utilise a Wooldridge test for autocorrelation, as shown in fig. 6. As expected, we are unable

	priceoveratm	priceoveratm
libor	-0.255	-0.243
	(1.12)	(1.06)
atmivol	1.161	1.130
	(15.16)**	(13.97)**
atmstrike	0.000	0.000
	(1.58)	(1.53)
L.priceoveratm		0.030
		(1.16)
_cons	0.000	-0.001
	(0.02)	(0.03)
R^2	0.15	0.16
N	1,410	1,410

* p<0.05; ** p<0.01

Figure 2: A table with and without an autoregressive term with the coefficients assumed to have a fixed effect.

```
. xttest3

Modified Wald test for groupwise heteroskedasticity
in fixed effect regression model

H0: sigma(i)^2 = sigma^2 for all i

chi2 (5)  =      5.1e+05
Prob>chi2 =       0.0000
```

Figure 5: Results from the modified Wald test with the null hypothesis of groupwise homoskedasticity.

```
. xtserial priceoveratm atmstrike atmivol libor

Wooldridge test for autocorrelation in panel data
H0: no first-order autocorrelation
    F(  1,        4) =      1.565
           Prob > F =      0.2791
```

Figure 6: Results from the Wooldridge test for serial correlation. The null hypothesis is that there is no first-order correlation. As expected we are unable to reject the null hypothesis in this case.

to reject the null hypothesis that there is no first oder autocorrelation and thus we find no significant evidence of serial correlation.

With evidence for heteroskedasticity and no evidence for first-order autocorrelation, we utilise and compare both a feasible generalised least squares (GLS) model and a panel-corrected standard error (PCSE) model. The PCSE model is an alternative to the feasible GLS model. We, thus, run both the regressions assuming that the errors are heteroskedastic in nature. The results are shown in fig. 7. We see that when using the feasible GLS model both the at-the-money strike price and the constant term in addition to the at-the-money implied volatility are statistically significant at the 1% level, whereas we find no such statisticaly significance for the two coefficients in the PCSE model.

Beck and Katz (1995) shows how PCSE models are more preferable to feasible GLS models. Beck and Katz (1995) show through a monte carlo simulation that the feasible GLS models often underestimate the standard errors in the regression. If that is the case then the newly found statistically significant variables' may have underestimated standard errors. So in order to have a conservative estimate we instead rely on the results of the PCSE model for our analysis and thus we are once again left with just one statistically significant variable. The PCSE also takes into account the heteroskedasticty across panels and possible cross-sectional correlation.

The coefficient of 1.162 to the *atmivol* indicates

	priceoveratm	priceoveratm
atmstrike	0.000	0.000
	(3.11)**	(1.52)
atmivol	1.315	1.162
	(29.72)**	(14.89)**
libor	-0.066	-0.242
	(0.51)	(1.04)
_cons	-0.066	0.000
	(4.80)**	(0.02)
N	1,415	1,415
R2		0.08

* p<0.05; ** p<0.01

Figure 7: Results from the feasible GLS and the PCSE models. The first column represent the feasible GLS coefficients while the second column represent PCSE coefficients.

that a tenth of a rise in the volatility of the at-the-money put options leads to a rise of 0.1162 in the price of the out-of-the-money put options with respect to the price of the at-the-money put option. In addition to that, the above tests revealed that in our dataset there was no reason to believe that there were time invariant characteristics unique to the strike rates. This would indicate that we are missing the factors that would provide the different strike rates with their unique characteristics, and our choice of variables is in no way exhaustive.

4 Conclusion

Thus, we began our empirical study of the out-of-the-money put options prices for S&P500 index with three dependent variables: at-the-money implied volatility, at-the-money strike price and LIBOR values. The put options prices were obtained by substituting the implied volatilities in the Black-Scholes equations, and at-the-money implied volatility was used as a substitute for the "actual" volatility.

We began by testing whether a fixed effect panel regression or a random effect panel regression should be used and quickly learn that even though a fixed effect regression makes more sense, but since both the regresions produce extremely similar coefficients and standard errors the Hausman test would reveal that the random effects model is preferrable over the fixed effects model. Similarly, the Breusch-Pagan Lagrange Multiplier test also revealed that the random effects modelwould be preferrable over a simple OLS model. We also found no significant evidence of serial correlation, while there was esignificant evidence of groupwise heteroskedasticity.

However, in order to account for cross-panel heteroskedasticty, we move towards a feasible Generalised Least Squares model and a Panel-Corrected Standard Error. The GLS model revealed two new statistically significant coefficient while the PCSE model did not, mostly due to smaller standard errors. However, the GLS models often underestimate the standard errors. Thus, the PCSE model was chosen for our analysis. We found that the only coefficient that was statistically significant was the coefficient for at-the-money implied volatility, which was found to be 1.162. Thus, our model indicates that a 0.1 rise in at-the-money implied volatility gives a rise 0.1162 in the out-of-the-money option price to at-the-money option price. However, the explanatory power of our model is extremely low, 0.08. Thus, our models are far from complete and none of the variables included in our dataset could account for the unique characteristics of different strike rates.

A AIC and BIC values for panel data fixed effect regression with and without an AR(1) term.

As can be seen from figures 9 and 8, both the AIC and the BIC values are lower when there is no lag.

References

[1] Beck, N, and J.N. Katz (1995). What to do (and not to do) with time-series cross section data, *American Political Science Review*, **89**, 634-47.

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	1415	272.9819	391.3362	4	-774.6724	-753.6529

Figure 8: AIC and BIC values for the panel regression with no lagged variables.

[2] Black, Fischer, and Myron Scholes (1973). The pricing of options and corporate liabilities, *Journal of Political Economy*, **81**, 637-654.

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	1410	269.6258	388.3639	5	-766.7277	-740.471

Figure 9: AIC and BIC values for the panel regression for an AR(1) model.