CALIBRATING AN OPTION PRICING MODEL WITH MARKET DATA

Jacques Burrus [jacques.burrus@bfi.lat]

This term project is an opportunity for students to apply a theoretical pricing model to real financial market data.

Specifically, the project consists in calibrating a given model onto a volatility smile. Students will have to assess the relevance of the model. Nuanced conclusions highlighting the pros and cons of the model are expected.

Students should underline the economic context of their analyses.
Models:

The students will be assigned one of the following models:

Local volatility:

1. Bachelier volatility
2. Constant Elasticity Volatility
3. Displacement diffusion volatility
4. Hyperbolic volatility
5. Ramp volatility

Stochastic volatility:

6. Heston model
7. Stein and Stein model (Stein and Stein 1991 - Stock price distributions with stochastic volatility: an analytical approach)
8. Hull and White stochastic volatility model (Hull White 1987 - The pricing of options on assets with stochastic volatilities)
9. Perello, Sircar, and Masoliver model (Perello Sircar Masoliver 2008 - Option pricing under stochastic volatility The exponential Ornstein-Uhlenbeck model)

Jump diffusion:

11. Constant jump size
12. Log-normal jump size

PART I: The data

Step 1:

Good quants usually get acquainted with the market data before plunging head first into complex mathematics.
The quant must be aware of the significant historical events that affected the market and the resulting regime changes. Indeed, a historical calibration is only meaningful within a given regime. Simple statistics such as the mean, standard deviation, maximum or minimum of times series may turn out useful. Comments and additional relevant graphs can provide extremely valuable insight.

Step 2:

Quants must absolutely verify that the data consistency. This step also allows to check that they fully understand the data.

In this project, the students will:

- Compute the forward price based on the spot price and relevant discount factors for all tenors,
- Compute the strike prices based on the delta, forward price, relevant discount factor, implied volatility, and term for all tenors and pillars (the delta convention may depend on the tenor and pillar),
- Compute the option value based on the strike, spot price, relevant discount factors, and implied volatility. This value may not necessarily be the value of a call. It could be a put, straddle, collar, or butterfly. The student must figure out the nature of the contract.

NB: Notice that these calculations do not require to transform the discount factors into interest rates.

PART II: The pricing engine

Before proceeding to the calibration, the students must demonstrate that they are able to properly apply the model to the traded European vanilla options. The grading committee will assume that the calculations are wrong. Students must henceforth rigorously justify the contrary. This is called benchmarking.

Step 3:

Benchmarking requires pricing the same vanilla contract with the same model with at least two independent numerical methods.

The students must choose two pricing methods depending on the model and justify their choice. For example, they can compare a Monte-Carlo simulation with finite differences or a theoretical formula.
The report must contain the precise description of the employed numeric methods. In particular, the students are strongly advised to use variance reduction techniques and/or customize the standard methods seen in class so as to differentiate themselves from their other classmates.

Step 4:

First, before attempting the pricing of vanilla options, each numerical method must be tested against:

- Term deposits for the different tenors whose value is thee domestic discount factor: $V = D^{DOM}(T)$
- Forward contracts for the different tenors and strikes (C10, C25, P10, P25, ATM): $V = D^{FOR}(T) \cdot S - D^{DOM}(T) \cdot K$

Specifically, the student must compare the numeric and theoretical values.
In the case of Monte-Carlo simulations the student must report the 99% confidence interval and check that the numeric values are 99% within the interval.
NB: Students might be surprised that even plain Monte-Carlo simulations hide tricky details.

Step 5:

Second, the quant will price European vanilla calls options in the special case where the model parameters coincide with a model with a known analytic formula such as the Black-Scholes model.
The student must explain how he sets the model parameters so as to enforce a constant volatility.
The European call options will have the same tenors and strikes as the forwards in the previous step.
The student will test constant volatilities of 5%, 10%, 20%, and 50%.
Sometimes, an academic article provides a price for a set of model and contract parameters. It is a good practice to reproduce the price.

Step 6:

Third, the quant will price the previous European vanilla calls options with arbitrary values for the model parameters.
The model parameters must now be chosen so as to significantly differ from the Black-Scholes benchmark. These deviations are to whole point of choosing more advanced models.
The student will modify each model parameter separately and verify that both numeric methods still yield similar values (that is within the confidence interval).
Naturally, the student will carefully justify the chosen range of parameter values based on the economic interpretation of said parameter or academic paper.
NB: Sometimes, an academic article provides a price for a set of model and contract parameters. It is a good practice to reproduce these prices.

Step 7:

The previous step has given a set of prices for European call options with various tenors and pillars under different model parameters.
In order to build intuition, the student will convert these prices into implied volatilities, plot the implied volatility smiles, and explain how each model parameter affects the smile.

Step 8:

At this stage, the student has verified that both numeric methods are trustworthy but must choose one of them for the calibration with real market data.
There is a trade-off between computational speed and pricing accuracy that must be quantified. For a given set of model parameters and the 3M ATM call option, the student will plot the implied volatility for both methods as a function of computation time. To do so, the student will have to fine-tune the numeric parameters such as e.g. the number of simulation paths, the number and range of nodes, and/or the size of the time intervals.
He should choose the method which yields an accuracy of 10bp on the implied volatility in the least computation time.

PART III: The model calibration

Last but not least, the students must provide reasonable model parameters and comment how suitable the model is for trading.

Step 9:

The calibration will be based on the minimization of some error metric wrt the model parameters and will therefore require a starting set of parameter values.
Unfortunately, the "optimal" parameters tend to depend strongly on the initial values and a poor starting point will almost inevitably lead to unreasonable results.
A good initial guess is in fact essential.
Hence, before engaging into a complex calibration, the financial engineer must base his initial guess on empirical data, which is the purpose of this question.
For each regime identified in the first part of the project, the student will perform simple empirical estimations based on the nature of his model as follows:

- Local volatility models: Plot the implied volatilities as a function of the spot price. The local volatility model specifies a parametric relationship between the instantaneous volatility -approximated here by the implied volatility as a quick-and-dirty proxy- and the spot price. A linear Ordinary Least-Squares (OLS) regression yields values for the parameters.
- Stochastic volatility models: Plot the implied volatility as a function of time. The model here specifies the stochastic process of the instantaneous variance approximated here as the squared implied volatility. An OLS regression can also provide reasonable values for the dynamics parameters. The variance of the residuals is generally related to the volatility of the instantaneous variance.
- Jump-diffusion models: The student must identify the jumps of the spot price. For instance, he can look at the returns greater the 3 standard deviations and then estimate the frequency and size of these jumps.

In all cases, the student must consider these values with a pinch of salt and decide whether they are economically reasonable. Shall the initial guess be unsatisfactory, he must not hesitate to choose other values based on his educated judgment or academic paper. For instance, starting the calibration of the Black-Scholes model with a volatility of 2000% is ridiculous. Anywhere between 5% and 50% is ok.

Step 10:

The calibration usually requires the minimization of some distance acting as an objective function. There exist many types of distances (for instance the grid, Euclidean, or maximum distances), many ways to express an error (for instance the difference, the linear ratio, or the log-ratio), and even various ways to express a price (the premium or implied volatility).
A good objective function must be sensitive wrt the model parameters.
The student should therefore work in terms of implied volatilities rather than raw option prices.
The professor suggests the following objective function ϵ for the smile at date t:

$$\epsilon = \frac{1}{N} \cdot \sum_{i,j} |\sigma_{ij}^{market} - \sigma_{ij}^{model}| \tag{0.1}$$

The student will here simply carefully explain how he will compute this metric ϵ.

Step 11:

Naively, a quant may think of minimizing the distance on the entire set of historical data. This approach is in practice insanely time-consuming and unrealistic.

Instead, it is advised to start with day by day calibrations. That is calibrating the model for a given smile at same specific date. This problem is much easier as a smile consists of few options (about 25 data points instead of thousands).
Of course, the first day is the most complicated because its parameters rely solely on the initial guess.

Here, the student will calibrate the model parameters for the first day.
In order to verify the reasonableness of these parameters, he will plot the market smile and compare it with the smile from his model with the obtained parameters.
The match may not be perfect but both smiles should be in broad agreement.

Step 12:

Next, the student must calibrate the model parameters for the subsequent days.
The initial guess for day $t+1$ will be the optimal parameters obtained from day t.
The student will observe that the calibration is must faster than for the first day because the initial values are already much closer to the optimal. Indeed, the smile does not fundamentally change between two consecutive days.
The student will plot the optimal parameters and the calibration error as a function of time and comment the evolution. Normally, the regimes identified in the first part of the project should appear.

Step 13:

Model parameters should -by definition- be constant in time at least within each regime. This step is maybe the most creative as students must imagine a calibration strategy on the historical data. On the one hand, it is desirable to use as much data as possible. But on the other hand, computational resources are finite.

The student will compute a set of model parameters for each historical regime. The initial guess is now straightforward given the results from the previous step.

Step 14:

It is understood that different calibration strategies will provide different model parameters. Henceforth, there is a model risk, i.e. an uncertainty on the model parameters.

To conclude, students can discuss how this uncertainty impacts the market prices of typical vanilla options.