MIXTURE OF MODELS: A SIMPLE RECIPE FOR A ... HANGOVER?

VLADIMIR V. PITERBARG

Abstract. The idea of using a weighted average of derivative security prices computed using different "simple" models (the so-called "mixture of models", or "ensemble of models", approach) has been put forth recently by a number of authors. Some view it as a simple way to add stochastic volatility to virtually any model, and others advocate it on the grounds that it provides a simple and tractable method for capturing certain market characteristics, most importantly volatility smile. Ease of calibration to market prices of vanilla and exotic instruments is also cited as the approach's redeeming quality. While not disputing the fact that such "models" are easy to calibrate, we explain that these models are under-specified (leading to multiple possible prices of derivatives). We also demonstrate that the "weighted average" valuation formula, the main selling point of the "mixture of models" approach, is self-inconsistent and cannot be used for valuation.

1. Mixtures of distributions

It has long been known that a mixture of normal distributions (i.e. a distribution with a density that is a weighted average of two or more Gaussian densities with different volatilities) has heavy tails. This observation has been explored by numerous authors aiming to describe and quantify volatility smile effects (see e.g. [2]). A European option price in a model with returns following such a distribution is easily computed as a weighted average of Black-Scholes prices calculated with different volatilities (see e.g. [5]). This approach is probably the easiest to incorporate volatility smile in a model, the fact that has undoubtedly lead to its popularity.

A model based on a mixture of Gaussian distributions for European options has another alluring interpretation. It is a simple case of a stochastic volatility model with the volatility randomly taking one of the pre-specified values with certain probabilities. This approach is trivial to extend to any other type of models, including interest rate models, commodity models, and so on. If a model has a concept of a volatility, one can formulate a number of volatility scenarios, value a derivative contract using each one of them, and average the values with certain weights. This more likely than not will introduce a volatility smile effect in valuation.

It what follows we will call such models "mixture models". They are different from what is known conventionally as "stochastic volatility models" as the latter

Date: July 18, 2003.

Key words and phrases. Mixtures of models, model ensembles, stochastic volatility, lognormal mixture.

I am grateful to Leif Andersen for his help and support. I also acknowledge useful discussions with Carol Alexander.

specify a *stochastic process* for volatility, whereas the former only specify a *discrete distribution* for it.

It has long been realized ([5]) that for European-style options, the actual volatility path has no bearing on the option value, as only the average volatility (what we call "term volatility" to distinguish it from instantaneous volatility) between now and the option's expiry matters. Because of that, mixture models can legitimately be applied to European options. Non-dynamic nature of stochastic volatility is of no importance as only the total effect of randomness on the *term* volatility is important. We can pretend that there is a stochastic volatility *process* that drives changes in volatility but we do not care to specify it fully as only its effects on the term volatility, embodied in the volatility scenarios chosen, is important.

Apparent success of this approach for European options has led some to propose a sweeping generalization of it to all other kinds of derivative contracts. Consider for example a problem of valuing a Bermuda swaption using a BGM model. If one has in his disposal a log-normal BGM model but wants to capture a volatility smile effect, what could be easier than valuing the Bermuda swaption in the log-normal BGM model a few times, each time scaling all BGM volatilities by some amount, and then averaging the results with some weights (weights and volatility scalings can be calibrated to match European swaption volatility smiles quite easily). Compared to the difficulty of adding "real", dynamic stochastic volatility to a BGM model (see [1]), this method suddenly seems very alluring. We argue however, that this extension takes the method simply too far – far enough to make this application simply wrong. (Using such non-dynamic stochastic volatility in the context of BGM models was recently proposed in [4]).

Some authors extend (stretch?) this method even further. They propose to take weighted averages (mixtures, ensembles) of models that do not even come from the same family – for example mixing local volatility models with stochastic volatility models and, presumably, jump models thrown in for a good measure (see [6], [7]). This supposedly allows the authors to calibrate the mixture model to all sorts of inputs including vanilla options, barrier options, and cliquets. In addition they claim to capture the dynamics of the volatility smile. Presumably the intention of the proponents of the approach is to price derivatives other than the ones to which their "model" was calibrated to.

The main attraction, to these authors, of mixture models is the simplicity of derivatives valuation. Let us write the formula down, so we can refer to it later. We denote by $\mathcal{S}_1, \ldots, \mathcal{S}_N$ the N "scenarios", and by $p_1, \ldots, p_N$ the weights associated with them. Scenarios can be different volatility scenarios, different alternative representations of reality, or different models. Let $V(\mathcal{S}_i)$ be the value of a derivative in the scenario $\mathcal{S}_i$. Then the value of the derivative in the "mixture model" V_{MM} is given by

$$V_{MM} = \sum_{i=1}^{N} p_i V(\mathcal{S}_i) \tag{1.1}$$

We will demonstrate that the formula (1.1) *cannot be used for valuation* of any but the simplest derivative contracts.

2. Mixture model dynamics

To use a model in practice, we typically need to specify what happens to the model variables as time evolves. The basic interpretation of a mixture model is that one flips a coin and, depending on the outcome, chooses the model A (with its dynamics of the underlying variables) or the model B (with its own dynamics). Such an interpretation looks quite shaky. At time 0 (today) there is uncertainty about which volatility scenario will be realized. Because of this uncertainty, prices are computed as weighted averages over different possible volatility scenarios. However, this uncertainty is resolved in the next millisecond. After that, volatility uncertainty does not exist at all. In math-speak, all volatility uncertainty is concentrated in the sigma-algebra (i.e. information flow) associated with time 0.

Relying on such unorthodox dynamics to generate volatility smiles leaves one with an unpleasant aftertaste and a sense of vague dissatisfaction. Putting all volatility uncertainty into a single instance of time makes the behavior extremely non-time-homogeneous. At the very least, such localized "information explosion" is certainly not consistent with the observed behavior of financial markets.

Another issue that comes to mind is how one can hedge consistently with the model's prediction of future dynamics. The answer is of course one really cannot. Hedges initiated at time 0 are "average" hedges, they offset an "average" move, either in volatility or the underlying. However, over an infinitesimal time period one of the volatility scenarios will be realized for sure, and the hedge will not compensate for that jump.

3. Are mixture models fully specified?

A proponent of mixture models will say that she does not necessarily believe in instantaneous resolution of uncertainty at time 0, but just does not specify how the uncertainty is resolved over time. It is enough to specify possible different states of nature (or uncertainty of our knowledge of the states of nature) at some future time, and not how the model "gets there". This brings up an important point. A mixture model only specifies what is going to happen at a fixed time in the future, and not what happens in between now and then. In such a formulation of a mixture model, there is no information on how the probabilities (or weights) of different scenarios are updated through time. It is akin to specifying only marginals of a multidimensional probability distribution, and not the dependence structure. Therefore, a mixture model is inherently under-specified. There are multiple dynamics that exist that lead to the same mixture at a given future time.

Is that necessarily bad? So what if there are many dynamics that are consistent with the "terminal" mixture? Well, for one, different dynamics can lead to different values of derivatives, depending on which dynamics are chosen. And that is, of course, less than ideal.

One can ask why do we even care about the dynamics of the model. We have the formula (1.1) that we can use to value any derivative, right? Wrong! The formula (1.1), the main selling point of the mixture model approach, *is inconsistent with any dynamics* that can be imposed on a mixture model. As such, it is not suitable for valuing all but the simplest (i.e. European-style options) derivatives. We present an example illustrating our point in the next section. The example is elementary and is presented for a compound put option.

For the FX markets, barrier options are probably more relevant. They are specifically mentioned in [6] and [7] as instruments for which mixture models are particularly useful. We discuss applications of mixture models to barrier options in Appendix.

4. Example

We consider the simplest possible case of a mixture model. We assume that a stock price S follows a lognormal distribution. The volatility is uncertain and can take two values, σ_1 with probability p_1 and σ_2 with probability $p_2 = 1 - p_1$. Then, according to the mixture model prescription, the price of any contract should be computed as p_1 times the Black-Scholes price of a contract with volatility σ_1 plus p_2 times the Black-Scholes price of a contract with volatility σ_2.

For simplicity we assume that the stock pays no dividends and the interest rate is zero.

Consider a simple compound put option. Today is time 0. There are two exercise dates T_1 and T_2, $0 < T_1 < T_2$. The holder has the right to exercise on either of these two dates (or not at all). If the option is exercised at time T_1, the holder receives $K_1 - S$ (and the option is gone). If the holder exercises at time T_2 he receives $K_2 - S$.

Let us denote the value of the compound option, at time t with stock at Sa by $V_{MM}(t, S)$. Denote the expectations under the two volatility scenarios by $\mathbf{E}^1$ and $\mathbf{E}^2$, respectively. Then, the mixture model prescription specifies that the value of the compound option at time $t = 0$ is equal to the weighted sum of the values of this option under the two volatility scenarios,

$$\begin{aligned} V_{MM}(0, S_0) &= p_1 \mathbf{E}_0^1 \left(\max \left(K_1 - S_{T_1}, \mathbf{E}_{T_1}^1 (K_2 - S_{T_2})^+ \right) \right) \\ &\quad + p_2 \mathbf{E}_0^2 \left(\max \left(K_1 - S_{T_1}, \mathbf{E}_{T_1}^2 (K_2 - S_{T_2})^+ \right) \right). \end{aligned}$$

So far so good. But let us look at the problem from a different prospective. Even though our model assumes that the volatility scenario will be revealed at the next instant of time, this of course will not be the case. In fact, come tomorrow, if nothing changes, we will be using the exact same mixture model, i.e. a weighted average of the two Black-Scholes models (same weights, same volatilities). In particular, on the date T_1, we will compute the "continuation value", the value of the remaining option should we decided not to exercise, using the same mixture model. In particular, the critical exercise price, S^*, will be determined at time T_1 using the equation

$$\begin{aligned} K_1 - S^* &= H(S^*), \\ H(S) &= p_1 \mathbf{E}_{T_1}^1 \left((K_2 - S_{T_2})^+ \middle| S_{T_1} = S \right) \\ &\quad + p_2 \mathbf{E}_{T_1}^2 \left((K_2 - S_{T_2})^+ \middle| S_{T_1} = S \right). \end{aligned}$$

Choosing the best option at time T_1 gives us

$$V_{MM}(T_1, S) = 1_{\{S > S^*\}} \times H(S) + 1_{\{S \le S^*\}} \times (K_1 - S).$$

Today's value of the compound option must be the expected value of the payoff at time T_1. The expected value of course should be done under our "mixture" measure. This gives us another way of computing the compound option value,

$$\tilde{V}_{MM}(0, S_0) = p_1 \mathbf{E}_0^1 V_{MM}(T_1, S_{T_1}) + p_2 \mathbf{E}_0^2 V_{MM}(T_1, S_{T_1}).$$

The two values must of course be equal for the model to be self-consistent. It turns out that in this mixture model, the two values $V_{MM}(0, S_0)$ and $\tilde{V}_{MM}(0, S_0)$ **are not equal**,

$$V_{MM}(0, S_0) \neq \tilde{V}_{MM}(0, S_0).$$

The table below summarizes numerical results for particular values of the parameters.

Quantity	Value
S_0	100
σ_1	10%
σ_2	50%
p_1	90%
p_2	10%
T_1	1
T_2	2
K_1	110
K_2	100
$V_{MM}(0, S_0)$	12.97
$\tilde{V}_{MM}(0, S_0)$	12.83

On a theoretical level, this result shows that our "non-dynamic" stochastic volatility specification cannot be used to price contracts that depend on the dynamics of the underlying market variables (unlike European options that depend only on their terminal distribution).

Practical implications are also pretty clear – we cannot use the mixture model approach to valuing any (but the simplest) contracts as weighted averages of their values over "simple" volatility scenarios. It is just plain wrong. A trader who, in the example above, buys the option for 12.97 and follows the model's recommendations about optimal exercise and hedging, is in for a nasty surprise – he will only manage to earn 12.83 over the life of the option, a loss of 14 cents. His risk manager (and, probably, a model validation Quant) will be in for a nasty surprise as well.

5. Conclusions

We have demonstrated that specifying a model that is a "mixture" of other models presents two problems. One is that the model is not really fully specified. There are multiple dynamics that are consistent with the "mixture", and different dynamics lead to different prices of derivatives. The second problem is that none of these possible dynamics are consistent with the valuation formula (1.1) for complex derivatives.

If one is only interested in valuing European options, using mixture models is appropriate, as the (multiple possible) dynamics have no bearing on their values, and the formula (1.1) can be used. For anything more complex, one has to do two things. First, fully specify the evolution of all the state variables in the model through time. Second – *do not use the formula* (1.1). This has been the approach of Brigo and Mercurio ([3]). Brigo and Mercurio start with a model for a stock price such that all one-dimensional distributions in the model are mixtures of lognormals. They, however, do not propose to use a weighted average of Black-Scholes models to value *all* derivatives. Instead, they derive a local volatility model for the stock price *consistent* with the assumption of lognormal distribution mixture. A local volatility model is of course fully self-consistent. It is the latter model that is then applied to price all derivatives of interest.

This approach however is not easy to extend to the situations discussed in [6] and [4]. How does one derive consistent dynamics for a model which is a mixture of local volatility, stochastic volatility and jump models? Or even a mixture of BGM models with different deterministic volatilities? This is a highly non-trivial step, a step that the proponents of mixture models seem to have overlooked (save for Mercurio and Brigo, of course).

References

[1] L.B.G.Andersen and R. Brotherton-Ratcliffe. Extended Libor market models with stochastic volatility. *Gen Re working paper*: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=294853

[2] B. Bhupinder. Implied risk-neutral probability density functions from option prices: A Central Bank prospective, In *Forecasting Volatility in the Financial Markets*, 137-167, Oxford.

[3] D. Brigo and F. Mercurio. A mixed-up smile, *Risk*, September 2000

[4] D. Gatarek. LIBOR market models with stochastic volatility, *SSRN working paper*: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=359001

[5] J.Hull and A.White. The pricing of options on assets with stochastic volatilities. *Journal of Financial and Quantitative Analysis* 3, 1987, 281-300

[6] S. Johnson and H. Lee. Capturing the smile, *Risk*, March 2003

[7] S. Johnson and H. Lee. Application of ensemble methods to interest rate options, Forthcoming

Appendix A. Can barrier options be valued with the "weighted average" formula?

Market wisdom states that market prices of FX barrier options lie somewhere in between two model prices, one obtained in a local volatility model and the other obtained in a stochastic volatility model (both calibrated to European options on the FX rate). It has been suggested to use a weighted average of the two model prices to match the market. We have shown that the formula (1.1) is inconsistent with any dynamics at least for some instruments (i.e. compound-type options). Could it be that we can still use the formula to value barrier options?

Let us state the problem formally. Again we consider the simplest possible mixture model, a weighted average of two Black-Scholes models with weights p_1 and p_2. The two Black-Scholes models have volatilities σ_1 and σ_2 (these are the two scenarios $\mathcal{S}_1$ and $\mathcal{S}_2$ from the formula (1.1)). We ask the following two questions. First, is it true that under *all* dynamic models consistent with the mixture model we defined, all barrier options can be valued by (1.1)? And second, is there *any* dynamic model consistent with the mixture such that all barrier options are valued by (1.1)?

For simplicity we assume zero interest rates and no dividends.

To answer the first question, it is enough to construct a model that is consistent with the mixture assumption in which barrier option prices do not satisfy (1.1).

Let $\pi_1(t,x)$, $\pi_2(t,x)$ be lognormal densities for time t, with volatilities σ_1.and σ_2 respectively. A local volatility model that is consistent with the mixture is easy to construct (see [3]):

$$\begin{aligned}\frac{dS(t)}{S(t)} &= \sigma(t,S(t))\,dW(t), \\ \sigma^2(t,x) &= \frac{p_1\sigma_1^2\pi_1(t,x)+p_2\sigma_2^2\pi_2(t,x)}{p_1\pi_1(t,x)+p_2\pi_2(t,x)}.\end{aligned} \tag{A.1}$$

Next, let us fix a barrier $L > 0$. Let $B(t,x)$ be the value of an up-and-out continuous barrier call option with strike x and maturity t. Its price is given by

$$B(t,x) = \mathbf{E}_0 \left[1_{\left\{\max_{0\le s\le t} S(s)<L\right\}} (S(t)-x)^+ \right].$$

We claim that **it is not true** that

$$B_{MM}(t,x) = p_1 B_1(t,x) + p_2 B_2(t,x) \text{ for all } t,x, \tag{A.2}$$

where B_{MM}, B_1, B_2 are the prices of the barrier option in the model (A.1) and the two Black-Scholes models, respectively.[1]

Assume (A.2) is correct. By differentiating the equality twice with respect to x we obtain

$$f_{MM}(t,x) = p_1 f_1(t,x) + p_2 f_2(t,x) \text{ for all } t,x, \tag{A.3}$$

where

$$f(t,x) = \mathbf{E}\left[1_{\left\{\max_{0\le s\le t} S(s)<L\right\}} \delta_x (S(t)) \right]$$

is the the so-called *survival density* (in all three models). The survival density in the model (A.1) satisfies the following PDE:

$$\begin{aligned} \frac{\partial}{\partial t} f_{MM}(t,x) &= \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2(t,x)\, x^2 f_{MM}(t,x)\right), \\ f_{MM}(t,x) &= 0 \text{ for } x > L. \end{aligned} \tag{A.4}$$

Plugging (A.3) into this equation, and using the fact that f_1 and f_2 satisfy similar equations with volatilities σ_1 and σ_2, we obtain

$$\begin{aligned} \frac{\partial}{\partial t}\left(p_1 f_1(t,x) + p_2 f_2(t,x)\right) &= \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2(t,x)\, x^2 \left(p_1 f_1(t,x) + p_2 f_2(t,x)\right)\right) \\ p_1\sigma_1^2 \frac{\partial^2}{\partial x^2} x^2 f_1(t,x) + p_2\sigma_2^2 \frac{\partial^2}{\partial x^2} x^2 f_2(t,x) &= p_1 \frac{\partial^2}{\partial x^2}\left(\sigma^2(t,x)\, x^2 f_1(t,x)\right) + p_2 \frac{\partial^2}{\partial x^2}\left(\sigma^2(t,x)\, x^2 f_2(t,x)\right). \end{aligned}$$

Integrating both sides with respect to x yields

$$p_1\sigma_1^2 f_1(t,x) + p_2\sigma_2^2 f_2(t,x) = \sigma^2(t,x)\left(p_1 f_1(t,x) + p_2 f_2(t,x)\right).$$

Expanding $\sigma^2(t,x)$ with the help of (A.1) results in

$$\frac{p_1\sigma_1^2 f_1(t,x) + p_2\sigma_2^2 f_2(t,x)}{p_1\sigma_1^2 \pi_1(t,x) + p_2\sigma_2^2 \pi_2(t,x)} = \frac{p_1 f_1(t,x) + p_2 f_2(t,x)}{p_1 \pi_1(t,x) + p_2 \pi_2(t,x)}.$$

It is obvious that this relation cannot hold for all t,x, unless there are very strong linear relationships between π_1, π_2, f_1 and f_2 (and we know there aren't any).

With the first question answered, let us move on to the second one. Can we come up with any dynamic model in which (A.2) is satisfied? It turns out that one does exist. It consists of a Bernoulli random variable ξ that has a probability p_1 of being 1 and probability p_2 of being 0. This random variable ξ lives in the time-0 sigma-algebra $\mathcal{F}_0$, and the dynamic equation governing S is given by

$$\frac{dS(t)}{S(t)} = \sigma_\xi\, dW(t). \tag{A.5}$$

This model is of course nothing other than the "coin-toss" model discussed in the beginning of Section 2. While this is a dynamic model, its implications, as discussed earlier, make the model completely unsuitable for application.

---

[1] The idea of the proof was suggested to us by Leif Andersen

Does there exist a "real" and "reasonable" dynamic model, in which uncertainty is revealed over time, and not in an instant explosion of information as in (A.5), such that all European options and all barriers are priced using (A.2)? The answer is most likely no, but we do not have a formal proof.

Bank of America, 233 South Wacker Drive, Suite 2800, Chicago IL 60606
E-mail address: vladimir.piterbarg@bankofamerica.com