



Comment

Innovations in testing the stability of risk measures over time and across models

Peter J.G. Vlaar

De Nederlandsche Bank, Westeinde 1, 1017 ZN Amsterdam, The Netherlands

The paper by Robert Bliss and Nicolaos Panigirtzoglou, deals with the stability of implied probability density functions (PDFs) retrieved from cross-sections of observed option prices. Implied PDFs are gaining increasing attention amongst both academics and practitioners. The main advantage of implied PDFs is that these provide, in principle, a complete picture of market expectations, including the probability attached to unlikely outcomes. Knowledge of these PDFs is necessary for pricing complicated derivatives. Moreover, the PDFs can serve as indicators of market sentiment, which is especially interesting around major economic events. Changes in PDFs can for instance be used to evaluate policy actions.

Notwithstanding their theoretical appeal, the practical usefulness of PDFs depends of course on the reliability of the results found. The analysis of Bliss and Panigirtzoglou is not very encouraging in this respect. They first give a review on the five different methods to derive PDFs. Then, they evaluate two of the most popular candidates: the double-lognormal approximating function, and the smoothed implied volatility smile. Whereas previous studies focussed on in-sample goodness of fit or price perturbations assuming correctly specified models, the current paper evaluates the stability of PDFs on the basis of actual option prices. The tests are conducted using Short Sterling futures options and FTSE 100 index options.

As a first test, the day-to-day changes in summary statistics of implied PDFs are calculated. The test reveals that the PDFs are extremely volatile, especially for the double-lognormal method. Moreover, the results found for the two

E-mail address: p.j.g.vlaar@dnb.nl (P.J.G. Vlaar).

different methods are hardly correlated. Consequently, conclusions regarding the effects on market expectations of, for instance, monetary policy actions based on changes of the implied PDFs can hardly be called robust.

Then the influence of perturbing the option prices by at most one half of the quotation tick size is given. These perturbations are only the minimal unavoidable pricing errors that are due to discrete price notations. Probably much more severe pricing errors due to non-synchronous trading of the options and the underlying are not even included. Nevertheless, the consequences are already severe. Even if the 10% largest outliers are disregarded, higher order statistics can hardly be estimated with precision. The confidence intervals can be so large as to make the estimation useless. This is especially true for the Short Sterling contracts, for which the number of strikes per cross-section is rather limited. For almost all calculations the results for the double-lognormal are much less stable than those for the smoothed implied volatility smile.

As for the reasons for this instability, several arguments can be mentioned. First, most option series are not very liquid. Indeed, most option series are not quoted or traded at all on most dates. Consequently, irregular quotations might lead to huge day-to-day changes. Second, the discrete tick size of option prices is relatively large compared to its time value. Only for at-the-money options the tick size is negligible, but the shape of the PDF is much more influenced by the out-of-the-money options. Third, the cross-sections on which the PDFs are based are quite small. For the Short Sterling contract the number of strikes can be as low as five. Consequently, these data are easily overfitted. Fourth, option prices do not contain information on the far tails of the distribution, beyond the highest and lowest strike prices. Interpolation is much easier than extrapolation. Therefore, especially the higher moments and the extreme percentiles of the distribution are hard to predict. Finally, the double-lognormal distribution is probably too flexible. This discrete mixture of two lognormal distributions contains five free parameters. This flexibility can lead to odd-looking bimodal PDFs, especially if the number of strikes in the cross-section is low. An outlier in the data is particularly problematic as one of the two lognormal densities might completely concentrate at this observation. If the variance of this lognormal would not be bounded below, it would even become zero, leading to an unlimited likelihood value. The economic interpretation of these spikes is of course absent, and indeed PDFs with spikes related to a single observation are discarded in the analysis.

So, what should we learn from the analysis. First of all, we should stop using the double-lognormal approximating function. The smoothed implied volatility smile method performs much better. Moreover, this method can probably be improved further by imposing less flexibility in the smoothing function. Imposing the same fit for both methods is useful for the comparison, but as the double-lognormal probably leads to overfitting, this is probably not the optimal trade off between fit and stability. Second, the problems related to mea-

surement errors – either related to the discrete tick size or to non-synchronous trading – are severe. These can only be reduced by increasing the sample size. This will in general not be possible in the cross-section dimension. Even if more option contracts would be introduced, it is not likely that they will be frequently traded. It might be possible however, to increase the number of observations in the time dimension. If transaction data would be available, these could be used to extract information at every option transaction. This would also reduce the non-synchronicity problem. Otherwise, distributions might be extracted from the panel of several days, instead of using just one day. The clear advantage of using the time dimension is that measurement errors average out. This comes at the price of having to assume homogeneity of the distribution of price changes over the time span. Given the apparent lack of stability of implied PDFs based on just one cross-section, this price is certainly worth considering.

I now turn to the two papers that are concerned with the stability of credit rating transition matrices. Rating migration is an important element in credit risk models. These models are becoming more widespread in their use among financial institutions as a way to attribute economic capital and to price credit derivatives. Also in the latest proposals for the new Basle capital accord, the potential use of credit risk models based on credit ratings is emphasised, although it is not yet clear whether banks will indeed be permitted to base their capital requirements for credit risk on these models in the future. At the moment, the most common assumption in these models is that the probability to migrate to another rating category, as well as the associated credit spreads and recovery rates in case of default are constant in expectation over time.

Both papers question the validity of this unconditional view of rating transitions. Anil Bangia, Francis Diebold and Til Schuermann first investigate the Markovian assumption for the credit migration process. Based on an analysis of eigenvalues and eigenvectors of rating transition matrices for different horizons, they conclude that the assumption of a first order Markov process is a reasonable approximation. This means that the expected transition matrix for a long horizon can indeed be computed as the matrixpower of a short horizon transition matrix. If the history of rating transitions is taken into account however, deviations from Markovian behaviour are detected. Based on a comparative static analysis, it is shown that the probability of a downgrade this year is higher if the same bond was downgraded last year. Especially bonds that were downgraded last year to the C to CCC class are more prone to default than those that were in this class already the year before.¹

¹ This result might be due to the fact that all bonds in the C, CC, and CCC categories are grouped in the same class. As the probability of default is higher for the C category than for the CCC bonds, the population of downgraded bonds will contain more C rated bonds.

They then investigate the influence of the business cycle on rating migration. Based on the monthly NBER business cycle indicator, they compare the average quarterly transition matrices in recessions and expansions for US-based firms over the 1981–1998 period. As to be expected, the probability of a downgrade, or a default, is higher in a recession than in an expansion. This information can be used to stress test credit portfolio models. The authors focus in this respect on the determination of the right probability to enter a recession or an expansion period. They compare the NBER business cycle data for the 1981–1998 data with the sample 1959–1998. The data reveal that in the longer period, recessions were slightly more frequent and lasted on average longer. If indeed this longer sample represents recession probabilities best, using the unconditional migration matrices would understate credit risk.

One may wonder whether the sixties and seventies are still containing relevant information for current business cycle dynamics. The original sample already contained eighteen years of data, so this can hardly be considered a small sample. As a method however, this technique might prove very useful as in many cases, the internal ratings gathered by banks cover much shorter periods. Of course, in order to be able to discriminate between recessions and expansions, the rating transition data should include at least one recession, otherwise the recession transition matrix can not be computed.

Irrespective of the usefulness of determining the proper probabilities to enter and remain in a recession, it can be questioned whether this is relevant for a stress test. A stress test is supposed to analyse the least favourable conditions, so one should always assume a recession. Whether the actual probability to enter a recession is 15% or say 25% is irrelevant for a stress test. It is very relevant however in calculating the unbiased value-at-risk.

David Lando and Torben Skødeberg emphasise the use of continuous-time rating transitions instead of the more common discrete time ‘cohort’ method. They state that rating transition is a continuous process, and as the dates on which the ratings of firms are adjusted are publicly available, these exact dates should be taken into account in order to efficiently calculate transition probabilities. If the transition process is time homogeneous, the resulting solution is of the matrix exponential form.² A major advantage of this approach is that the expected transition probabilities can be computed for arbitrary time periods, whereas these probabilities can only be computed for integer times if the cohort method is used. Moreover, even for arbitrary small time periods, the probability of a large rating change, for instance from AAA to default, will in general be non-zero, even if such a transition has never taken place over such a small time period in the past. Indeed, this probability should be non-zero as a

² This result is the matrix equivalent of the homogeneous solution of a first order differential equation.

AAA bond could in principle default via successive downgrades over the period considered. For longer time horizons, these advantages are less important as the non-zero default probability for AAA bonds also arises from the matrix power solution for the cohort method.

Unfortunately, the neat solution for the time-homogeneous case might not be applicable most of the time as indeed expected rating migrations are not constant over time. A continuous time non-homogeneous solution can still be computed. It is very similar to the discrete time one. This so-called product-limit estimator differs from the cohort method in that a transition matrix is calculated every time a migration occurs. The product-limit estimator for a certain period is subsequently calculated as the matrix product of all (mostly single migration) transition matrices over this period.

In the third part of the paper, it is investigated whether the rating transition process exhibits non-Markov behaviour. In particular, it is shown that the probability of an upgrade (downgrade) is significantly higher if this bond was upgraded (downgraded) before. Moreover, a longer history in a certain rating category significantly increases the probability of staying in that class. The method used in these exercises is the Cox proportional hazard model. In this model, the probability of transition for a particular bond at a particular time is proportional to the time-varying baseline hazard. The proportional part is a function of the covariates, whereas the baseline hazard is left unspecified. The influence of the covariates can be analysed without specifying the baseline hazard by considering only transition times. Consequently, the transition probabilities are calculated conditional on the knowledge that one bond is indeed migrating. Therefore, only the relative hazards are important. The baseline hazard absorbs all other time-varying elements, such as economic fluctuations.

An advantage of the Cox proportional hazard model is that it gives a consistent estimate of the influence of the covariates, even if other important sources for time variation are neglected. However, that also implies that knowledge on the influence of the covariates is not enough to predict future migration probabilities. The method gives a clear signal that something is wrong, but is not very illuminating on what we should do instead. Moreover, the proposed model does not take into account the fact that rating migrations are not independent of each other. Each element in the intensity matrix is calculated separately.

A more efficient way to calculate transition probabilities might be to use an ordered probit model, as in Nickell et al. (2000). This discrete time model does take into account the natural ordering of rating transition probabilities. Moreover, conditional on the knowledge of the covariates, the migration probabilities can be estimated. The resulting model can be considered 'conditional homogeneous'. The advantages of the continuous time models can easily be incorporated in this model if one is willing to assume that the covariates are

constant over the discrete time spans over which the migration matrices are calculated. Given constant covariates, the way to proceed is to first estimate the, say, quarterly continuous time generator matrices. Second, add the identity matrix to each of these matrices in order to get quarterly rating transition matrices.³ Third, use these matrices to estimate the ordered probit model. Conditional rating probability matrices can subsequently be computed using the exponential continuous-time solution, where the conditional generator matrix equals the conditional migration matrix minus the identity matrix. Of course, in order to calculate rating transition probabilities more than one period ahead, one needs a forecast of the covariates in the next period. If these are not the same as the current period the migration probabilities for the total time span can be computed as the product of the probability matrices of the sub-periods.

To conclude, both rating transition papers clearly demonstrate that using unconditional transition matrices might at times underestimate credit risk. In as far as the path dependence is concerned, the risk for a total portfolio will probably be only minor as this risk is likely to be diversified. The influence of the business cycle on the other hand is systematic, and therefore seems more dangerous. Moreover, this danger will be further enlarged by the fact that also the credit spreads, the loss given default and possibly the correlation between rating transitions is higher during recessions. These risk factors clearly deserve further research before credit risk models are granted a crucial role in banking supervision.

Reference

- Nickell, P., Perraudin, W., Varotto, S., 2000. Stability of rating transitions. *Journal of Banking and Finance* 24, 203–227.

³ The only difference between these transition matrices and the ordinary ‘cohort’ ones is the more efficient calculation of the denominator, that is the size of the group that might migrate.