Probability of default as a function of correlation: The problem of Non-uniqueness

Toby Daglish and Wei Li

September 26, 2005

Abstract

It is common practise in industry for traders to use copula models, combined with observed market prices, to calculate implied correlations for firm defaults. The actual feasibility of this calculation depends on the assumption that there is a one-to-one mapping between values of CDO tranches, and the correlation implicit in the copula. This paper presents several proofs which demonstrate that, for sufficiently large portfolios of underlying credits, the probability of certain number of default are hump shaped as a function of the correlation. We follow our analytical results with some numerical examples of pricing CDOs, demonstrating the non-uniqueness problem of implied correlations.

1 Introduction

In the fast growing credit derivatives market, products whose payoffs depend on multiple credits have gained in popularity. In this paper, we focus on one particular example of these products: collateralised debt obligations (CDOs). A CDO consists of a portfolio of corporate bonds, or a portfolio of credit default swaps. The portfolio manager then securitises the portfolio into multiple tranches of securities. These tranches are categorised as senior, mezzanine, and subordinated/equity. In cases when defaults occur, the more junior tranches take responsibility for the loss due to default. Their equity is reduced by the loss in portfolio capital or (in the case of a portfolio of credit default swaps) they are responsible for the payment of the swaps. As a result, junior tranches take on the majority of the credit risk, leaving senior tranches almost immune to credit risk.

The issue at the heart of pricing CDOs is the modelling of the correlations among defaults. Currently the most popular technique is the copula approach. This technique is used in the CreditMetrics package (see Gupton, Finger, and Bhatia (1997)) and summarised in the article by Li (2000). Copulas allow one to specify an arbitrary joint distribution for firm default times, and then (using this joint distribution) construct a new distribution which has appropriate marginal distributions. Depending on the choice of copula, different interdependencies can be achieved. The most common practise in industry is to use a normal distribution for the joint distribution and then force its marginal distributions to be consistent with risk neutral probabilities of default obtained from corporate bond and credit default swap prices. In a recent paper, Hull and White (2004) apply the copula approach to develop a fast procedure to evaluate CDOs and n^{th} to default swaps without recourse to Monte Carlo simulation.

In keeping with the calculation of implied volatility for regular derivatives, it is also popular for traders to use their models, combined with observed market prices, to calculate implied correlations for firm defaults. As is the case with implied volatility, these correlations vary over tranches written on the same basket of underlying securities.

Although this is standard practise in industry, the actual feasibility of this calculation is somewhat unclear. In order for this technique to work, there must be a one-to-one mapping between values of CDO tranches, and the correlation implicit in the copula. In many realistic situations, this condition is violated, leading to non-uniqueness of implied correlations: observed market prices may be explained by several different copulas.

This non-uniqueness poses a two-fold problem. On the one hand, the fact that one price can be explained by two different levels of correlation poses difficulties for inferring true correlations from CDO tranche prices, an obvious application of the technique. The second problem is related to the hedging of correlation risk. If CDO spreads are hump-shaped in correlation, then observing a given spread does not tell the user whether the spread would increase or decrease in response to an increase in correlation. In a market whose raison d'etre is the joint distribution of firm defaults, this shortcoming is very worrying.

This paper presents several results which demonstrate that this phenomenon will always occur for sufficiently large portfolios of underlying credits. Since CDOs traded in practise depend on 100 or more underlying credits, this represents a very real challenge to the implied correlation methodology. We follow our analytical results with some numerical examples of pricing CDOs.

Although the phenomenon of non-unique implied correlations has been frequently observed in practise, and is regarded as a "stylised fact" of the CDO market, this paper presents the first concrete analysis to show how and why this effect occurs. We note that in acknowledgement of this problem, market quotations changed during late 2004 to use the "base correlation" approach (see section 4).

Our hope is that this rigorous analysis of the problem will not only highlight some of the drawbacks of the implied correlations approach, but will also shed some light on the behaviour of the normal copula model. We hope that further study of the structure of the model will yield more intuition as to the risk structure implied by the model for individual tranches.

2 Analysis of a Simple Model

2.1 The Model

We use a simple model for illustrative purposes. This is a special case of the Hull and White (2004) one factor model in which there is only one period, default is symmetric across firms, and the normal copula determines defaults. Thus the probability of default for each bond is set to be constant p. Following Hull and White, we define:

$$x_i = aM + \sqrt{1-a^2} Z_i. \tag{1}$$

We also define x_0 such that $\Phi(x_0) = p$. When $x_i < x_0$, bond i defaults. Denote $d_i = (x_i < x_0)$.

2.2 Tails of the distribution

A slightly reduced (dimension of the fundamental probability space is lowered by one), but equivalent (equivalence should be seen in the correlation matrix), form of the probability model is that:

$$Z_0 = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} Z_i, \text{ and } x_i = \lambda Z_0 + \left(\sqrt{1 - \frac{N-1}{N}\lambda^2} - \frac{\lambda}{\sqrt{N}} \right) Z_i, \text{ with } \lambda \in [0, 1].$$

When $\lambda = 0$, the x_i's are independent. When $\lambda = 1$, they are identical. In this model, we have the following intuitive result.

Claim 1 *If $p \leq 0.5$ (i.e. $x_0 \leq 0$), we have that $f(\lambda) \equiv P[\text{All bonds default}]$ is a monotonically increasing function. The upper bound is $f(1) = p$ and the lower bound is $f(0) = p^N$.*

Proof: See Appendix A.

In Claim 1, we illustrate through an intuitive geometrical argument that it becomes more likely for all the firms to default at the same time as the correlation increases. However, we show it under the condition that each firm's default probability is below 0.5. Such a restriction is unnecessary, as we demonstrate in Claim 2.

Claim 2 *$f(\lambda) \equiv P\left[(x_j(\lambda) < x_0)\right]$ is a monotonically increasing function.*

Proof: See Appendix B.

Using the preceding results, we are able to state our first proposition regarding the distribution of defaults.

Proposition 1 *The probability of all bonds defaulting is an increasing function of λ. The probability of at least one default is a decreasing function of λ.*

Proof: The first part follows directly from claim 2. For the second part, notice that

$$P[\text{at least one default}] = 1 - P[\text{no defaults}] = 1 - P\left[\cap_{j=1,\ldots,N}(-x_j(\lambda) < (-x_0))\right],$$

which is a decreasing function of λ by claim 2 again.

□

The question still left untouched is what happens for the middle cases: how does the probability of at least m firms' default change with respect to correlation, with $1 < m < N$? We approach this question with some preliminary results regarding the mean and the dispersion of the number of defaults. These results provide some intuition as to why we might expect the non-uniqueness problem.

2.3 The Mean and the Variance

Denote $d_i = (x_i < x_0)$ where we introduce the indicator notation:

$$(y < x) = \begin{cases} 1 & \text{if } y < x, \\ 0 & \text{if } y \geq x, \end{cases} \tag{2}$$

with similar definitions for other relational operators. In this framework, firm i will default if and only if $d_i = 1$. The total number of defaults is thus $\sum_{i=1}^{N} d_i$, which we denote D.

Proposition 2 $E(D) = Np$. *The variance, var*$[D(\alpha)]$, *is an increasing function of* α, *with the minimum var*$[D(0)] = (p - p^2)N$ *and the maximum var*$[D(1)] = (p - p^2)N^2$.

Proof: It is clear that

$$E(D) = \sum_{i=1}^{N} E(d_i) = Np.$$

On the other hand,

$$\begin{aligned} \text{var}(D) &= E(D^2) - E^2(D) \\ &= \sum_{i,j=1}^{N} E(d_i d_j) - N^2 p^2 = N(p - p^2) + \sum_{i,j=1; i \neq j} \big(E(d_i d_j) - p^2\big) \\ &= N(p - p^2) + N(N-1)\big(E(d_1 d_2) - p^2\big) \end{aligned}$$

Note that $E(d_1 d_2)$ is equal to the probability that both firms default in the case where there are only two firms in total. Then by claim 1, $E(d_1 d_2)$, which is equal to the probability of both firms defaulting, is increasing in α, with the minimum value p^2 reached when $\alpha = 0$, the maximum value p reached when $\alpha = 1$.

□

Proposition 2 states that the total number of defaults, as a random variable, has mean $E(D) = Np$ regardless of the value of α, and has variance an increasing function of α. That is, as α increases, the distribution of the total number of defaults becomes more dispersed. One can imagine that the density mass is pushed away from the mean as α increases, with mass being pushed toward the two extreme cases where all firms default or all firms survive. This is consistent with the result of Proposition 1, i.e., the probabilities of the occurrence of the two extreme cases are increasing in α. Furthermore, in the process of the density mass being dispersed toward the two extreme cases, it is intuitive that the probabilities of a small number of defaults and that of a small number of survivals are increasing initially. This initial increase will create a hump shape in the probabilities of a small number of defaults and that of a small number of survivals, since as α approaches one, the probabilities will eventually approach zero. We formalise this intuition in the following subsection.

2.4 The Hump Shaped Probability of Exactly k Defaults

Thus far, we have proved that the size of the two tails of the distribution of the number of defaults are an increasing function of α. As the sum of the probabilities of all the possible numbers of defaults has to be one, the probability of the portion of defaults lying *between* zero and one has to go down. We show below that the probability of certain number of defaults will first increase and then decrease.

First, we make the assumption $x_0 = 0$. This is to say that an individual firm has a probability of 0.5 to default. Under this case, we are able to derive some general and precise results regarding the occurrence of the hump shape. We later relax this assumption, and derive some less precise results in the more general case. Our main goal is to demonstrate that the hump shape will occur under some loose conditions.

We denote the probability of exactly k defaults occurring as $P[k]$, and accordingly the probability, conditional on M, $P[k|M]$.

Proposition 3 *Under the condition that $x_0 = 0$, the probability of exactly k defaults occurring, $P[k]$, has a humped shape as a function of α, for all k such that $0 < k < \frac{N-\sqrt{N}}{2}$ and $\frac{N+\sqrt{N}}{2} < k < N$*

proof: See Appendix C

Proposition 3 shows that, under the assumption that $x_0 = 0$, the probability of exactly k defaults occurring will first increase and then decrease for certain k's as functions of the correlation. Certainly the assumption the assumption that $x_0 = 0$ is too restrictive. In the following proposition, we relax this assumption to derive a more general result regarding the shape of $P[k]$ as a function of correlation.

Proposition 4

1. *When $x_0 < 0$, for any k, there exists a positive integer K_1 such that whenever the total number of firms N is larger than K_1, the probability of exactly k defaults occurring, $P[k]$, has a humped shape as a function of α;*
2. *When $x_0 < 0$, for any k, there exists a positive integer K_2 such that whenever N is larger than K_2, the probability of exactly $N-k$ defaults occurring, $P[N-k]$, has a humped shape as a function of α;*
3. *When $x_0 > 0$,for any k, there exists a positive integer K_3 such that whenever N is larger than K_3, the probability of exactly $N-k$ defaults occurring, $P[N-k]$, has a humped shape as a function of α;*
4. *When $x_0 > 0$,for any k, there exists a positive integer K_4 such that whenever N is larger than K_4, the probability of exactly k defaults occurring, $P[k]$, has a humped shape as a function of α;*

Proof: See Appendix D

Proposition 4 claims, among other things, that when $x_0 < 0$ there will be a humped shape for the probability of exactly k default, when there are sufficiently large number of firms, i.e. large N.

2.5 The Hump Shape of Probability of at least k Defaults

We now turn our attention to the probability of at least k defaults; a probability which we will see is critical to the valuation of n^{th} to default CDSs and CDOs. As before, we use the change of variable $z = \frac{x_0-\alpha M}{\sqrt{1-\alpha^2}}$. The probability of at least k defaults, conditional on z, is:

$$P[D > k|z] = \sum_{n=k+1}^{N} \binom{N}{n} \Phi(z)^n \Phi(-z)^{N-n} = \frac{B(x; k+1, N-k)}{B(k+1, N-k)}$$

where $x = \Phi(z)$, and the beta function $B(k+1, N-k)$ and incomplete beta function $B(x; k+1, N-k)$ are given by:

$$B(k+1, N-k) = \frac{\Gamma(k+1)\Gamma(N-k)}{\Gamma(N+1)} = \frac{k!(N-k-1)!}{N!}$$

and

$$B(x; k+1, N-k) = \int_0^x u^k (1-u)^{N-k-1} du$$

Thus,

$$\begin{aligned} P[D > k] &= E\left[P[D > k|z]\right] \\ &= \int_{-\infty}^{\infty} P[D > k|z] \frac{\sqrt{1-\alpha^2}}{\sqrt{2\pi}\alpha} \exp\left[-\frac{\left(z - \frac{x_0}{\sqrt{1-\alpha^2}}\right)^2 (1-\alpha^2)}{2\alpha^2}\right] dz. \end{aligned}$$

and

$$\frac{\partial}{\partial \alpha} P[D > k] = \frac{1}{\alpha^3} E\left[P[D > k|z] \left(\left(\left(z - \frac{x_0}{\sqrt{1-\alpha^2}}\right)^2 - \frac{\alpha^2}{1-\alpha^2}\right) + \frac{\alpha^2 x_0}{\sqrt{1-\alpha^2}} \left(z - \frac{x_0}{\sqrt{1-\alpha^2}}\right)\right)\right]$$

In the above expression, as α approaches zero, the expectation is equivalent to an integration with the weight function being the delta function. Therefore, we can use a Taylor expansion to approximate the function in the expression. The following will be based on this intuition. The argument can be made precise in a similar fashion to the above proof of Proposition 4.

Let π be the proportion of defaults, that is, $\pi = k/N$. Denote

$$h(z) = P[D > k|z] = P[\frac{D}{N} > \pi|z].$$

If we let $\Theta(\alpha) = \frac{1}{\alpha}\frac{\partial}{\partial \alpha} P[D > k]$, it is straightforward to demonstrate that $g(\alpha)$ is continuous, and $\Theta(0)$ well defined as $\lim_{\alpha \to 0} \Theta(\alpha)$. We have

$$\Theta(0) = 2h''(x_0) + x_0 h'(x_0).$$

Substituting the function form of $h(z)$, and by taking derivatives of it, we arrive at

$$\Theta(0) = h'(x_0) \left(\frac{N\left(1 - \pi - \frac{1}{N}\right)}{\Phi(x_0)} \left(\frac{\pi}{1 - \pi - \frac{1}{N}} - \frac{\Phi(x_0)}{1 - \Phi(x_0)} \right) + \frac{\phi'(x_0)}{\phi(x_0)} \right)$$

When $\pi < \Phi(x_0) - \frac{1}{N}$, we have $\Theta(\alpha) < 0$ for sufficiently small α, which further implies that $\frac{\partial}{\partial\alpha} P[D > k] < 0$ for sufficiently small α. On the other hand, we can fix a value for π, with $\pi > \Phi(x_0)$. Then, by taking a sufficiently large value for N, we have that $\Theta(0) > 0$. Hence due to continuity of $\Theta(\alpha)$, we have $\Theta(\alpha) > 0$ for sufficiently small α, which further implies that $\frac{\partial}{\partial\alpha} P[D > k] > 0$ for sufficiently small α. Thus we have proved the following proposition:

Proposition 5 *For constant π, with $\pi < \Phi(x_0) - \frac{1}{N}$, the probability of the proportion of firms defaulting being at least π is initially decreasing in α. On the other hand, for constant π, with $\pi > \Phi(x_0)$, and for sufficiently large N, the probability of the proportion of firms defaulting being at least π is initially increasing in α.*

When $\alpha = 1$, the probability of the proportion of firms defaulting being at least π is equal to the unconditional probability that each individual firm defaults, i.e. $\Phi(x_0)$. When $\alpha = 0$, this probability is $\sum_{n>N\pi} \Phi(x_0)^n (1 - \Phi(x_0))^{N-n}$. We have the following corollary:

Corollary 1 *For constant π such that $\pi > \Phi(x_0)$, and $\sum_{n>N\pi} \Phi(x_0)^n (1 - \Phi(x_0))^{N-n} < \Phi(x_0)$, the probability that the proportion of firms defaulting is at least π is hump shaped as a function of α, for N is sufficiently large.*

Some appreciation of the conditional in the corollary can be achieved by the help of central limit theorem for the case that N is large. TO BE COMPLETED

2.6 The Hump Shape of n^{th} to default CDS and CDO rates

In this section, we demonstrate that the spread of n^{th} to default CDS for certain values of n, and that the spreads for certain tranches of a CDO will have a humped shape as functions of the correlation coefficient of the model, α.

2.6.1 n^{th} to default CDS

Let R denote the recovery rate. In the event of an n^{th} default occurring, the seller pays the notional principal times $1 - R$. This contract can be valued by calculating the expected present value of payments and the expected present value of payoffs in a risk-neutral world. The break even CDS spread is the spread for which the expected present value of the payments equals the expected present value of payoffs. Therefore, the payment rate of an n^{th} to default CDS is given by the following expression.

$$r = (1 - R)\frac{P[D > n - 1]}{1 - P[D > n - 1]}$$

Notice that the rate of an n^{th} to default CDS and $P[D > n - 1]$ are positively related. Therefore, the hump shape of $P[D > n-1]$ established in the corollary above for $\pi = \frac{n-1}{N}$ such that $\pi > \Phi(x_0)$, and $\sum_{i>N\pi} \Phi(x_0)^i (1-\Phi(x_0))^{N-i} < \Phi(x_0)$, carries over to the rate of the n^{th} to default CDS directly.

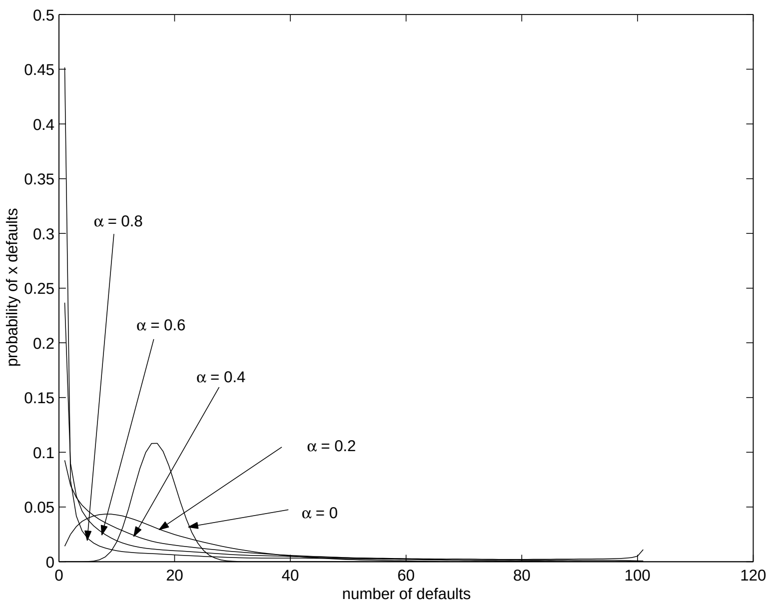

Figure 1: The probability of exactly k defaults occurring when defaults are generated by the model defined in equation (1), corresponding to $\alpha = 0, 0.2, 0.4, 0.6, \text{and } 0.8$

2.6.2 CDO

We consider a CDO with lower attachment point $\frac{m}{1-R}$ and upper attachment point $\frac{n}{1-R}$. Similar to the case of an n^{th} to default CDS, we have the following expression for the rate of this CDO:

$$r = (1 - R)\frac{\sum_{i=m-1}^{n} P[D > i]}{\sum_{i=m-1}^{n}(1 - P[D > i])}$$

For the case where both m and n satisfy the requirement of corollary following Proposition 5, we have the result that $P[D > i]$, for $i = m - 1, ..., n$, are all initially increasing, and are eventually decreasing in α. Therefore, the tranches of a CDO which satisfy this restriction all have hump shape as function of α.

3 Numerical Analysis of the Model

In this section, we further our analysis of the simple model presented in the previous section using numerical techniques. This analysis provides a more detailed picture of the behaviour of the model.

3.1 The Non-Monotonicity in Default Probability

In this subsection, we use numerical integration to locate the points (as function of α) where the default probability changes from increasing in α to decreasing in α. See figure 1.

In the figure, the x-axis is the number of defaults, and the y-axis is the probability of this number of defaults occurring. There are one hundred firms, we set parameter $x_0 = -1$, and the correlation coefficient α is set to be 0, 0.2, 0.4, 0.6, and 0.8 in this example. As one can see, when $\alpha = 0$ (i.e. there is no correlation across different firms' default) the picture resembles a normal distribution, which is assured by the central limit theorem. It is apparent from this figure, for intermediate values of number of defaults (for example, between 9 and 12, and between 25 and 28) non-monotonicity occurs. The probability of exactly k default, for $9 < k < 12$ and for $25 < k < 28$, initially increases for increasing α,, but then goes down.

We can also see that when α's value is large, i.e. close to one (and in the figure, for $\alpha = 0.8$), the distribution is pushed to the two end points, i.e., the point with no default and the point where all firms default. Eventually, when $\alpha = 1$, all the distribution will be concentrated in these two points only, with other numbers of defaults occurring with probability zero. This is so because when $\alpha = 1$, all firm default events are determined by the same indicator variable, and therefore, either all default or none default. In the picture, for k small enough (smaller than 8) or large enough (larger than 28), the probability of exactly k defaults conditional upon α has the lowest value when $\alpha = 0$. However, what is not immediately apparent from the picture is that for the case when alpha is very close to one, the distribution will be more and more concentrated on the two end points, and therefore will exhibit even lower probabilities of default for small or large values of k. Therefore, for all the cases with k either very small or very large, we will have that the probability of exactly k defaults is initially increasing in α, and eventually decreasing in α. This exhibits the non-monotonicity of the relation between the probability of exactly k defaults and α, as proved to be the case in Proposition 4.

3.2 The Hump Shape of CDO Rates

We have illustrated the hump shape, in a single period model, of the relation between the probability of exactly k defaults occurring and α. Such a non-uniqueness relation eventually manifests itself in the relation between the spread of a CDO tranche and the fundamental variable α. In this subsection, we illustrate that the non-uniqueness manifests itself in the more general framework used in Hull and White. More specifically, we allow for a multiple-period model of firm default, with quarterly defaults and spread payments. We consider four tranches in a CDO, which are responsible for between 3% and 6% of defaults, between 6% and 10% of defaults, between 7% and 10% of defaults, and between 10% and 15% of defaults respectively. The portfolio consists of 100 names. The total period is set to be 5 years. The expected recovery rate is 40%. The term structure of interest rates is assumed to be flat at 5%. We use a symmetric model, where all entities have a probability of about 1% of defaulting each year. See figure 2.

This figure shows the coupon earned by different CDO tranches as functions of the correlation coefficient alpha. It clearly demonstrates the hump shape in the function. Moreover, for more senior tranches, the maximal point of the hump shape function also moves to the right.

TO BE COMPLETED.

4 Base Correlations

The standard solution to the non-uniqueness problem in the CDO industry has been to consider "base" correlations. A base correlation is a correlation synthesised from regular CDO tranche prices,

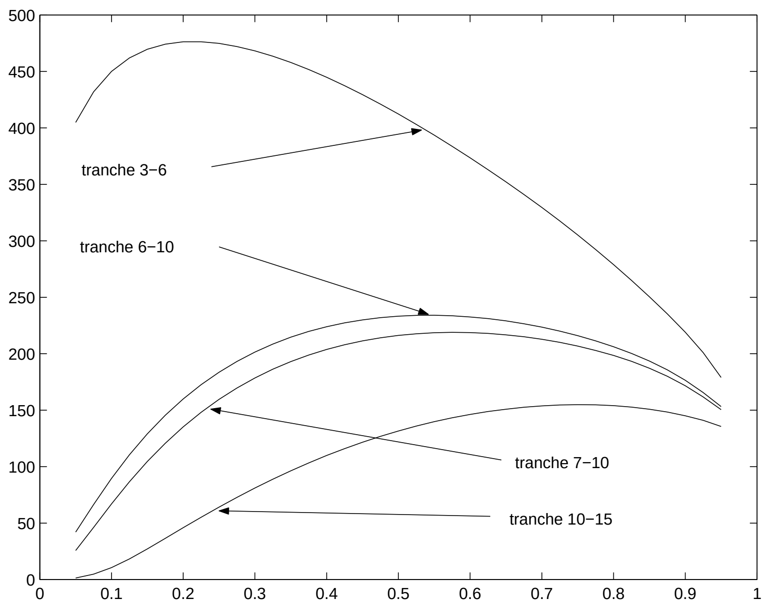

Figure 2: The CDO rates in basis points versus α.

so that its lower attachment point is zero, and its upper attachment point is the upper attachment point of a given tranche. The correlations are generated by a bootstrap approach.

We take Bear Stearns' approach (see Reyfman (2004)) as an example to illustrate the bootstrapping steps. Consider three tranches: $0-3\%$, $3-7\%$, and $7-10\%$. The starting point will be that the $0-3\%$ base correlation is equal to the compound correlation. We use the following notation: $i(0-a)$ stands for the implied correlation for $0-a\%$ tranche, $s(a,b)$ stands for the observed market spread for $a-b\%$ tranch, and $P(a,b,s;\rho)$ is the price for $a-b\%$ tranche with the spread s if the correlation is ρ. Then the bootstrapping method contains of the following steps:

1. Solve equation $P(0,3,s(0,3);\rho)=0$ for ρ. This gives $\rho(0,3)$.
2. Solve equation $P(0,7,s(3,7);\rho)=P(0,3,s(3,7);\rho(0,3))$ for ρ. This gives $\rho(0,7)$.
3. Solve equation $P(0,10,s(7,10);\rho)=P(0,7,s(7,10);\rho(0,7))$ for ρ. This gives $\rho(7,10)$.
4. Repeat the procedure till we have all the needed base correlation.

The second step above relies intuitively on: $P(0,7,s(3,7);.)=P(0,3,s(3,7);.)+P(3,7,s(3,7);.)=P(0,3,s(3,7);.)$ due to the fact that the market fair price $P(3,7,s(s,7);.)=0$. Notice that this equality is different from what has been used in the second step above to solve for $s(0,7)$. The difference is that in this equality, both sides are computed with the same correlation, but in the step two above, we use the correlation from the junior tranches on the right hand side to solve for the correlation from the more senior tranches on the left side. Also, the identity, $P(3,7,s(s,7);.)=0$, is used in the derivation. This identity is what is used in the traditional computation of implied correlation. The base correlation approach has great practical value, but with a cost of logical consistency. JP Morgan's approach (see McGinty (2004)) follows similar bootstrapping steps. The only difference being that JP Morgan base their computation on the expected loss for each tranche, while the Bear Sterns approach above is based on the value of each tranche.

The valuation of a new tranche requires interpolation between known base correlations to find the appropriate correlation for calculating expected loss and the risky duration of the portfolio.

5 Conclusions

The Normal Copula model has been widely used in practise for the purpose of valuing and hedging synthetic CDO tranches. One common application of this model in the past was to back out the implied correlation from the observed spread of a tranche. In this paper, we demonstrate through rigorous proofs, the non-uniqueness problem that results from this. From our analysis, we can see that a naive application of the model can give contradictory suggestions for appropriate hedge ratios. We note that practitioners are aware of this non-uniqueness problem, and the current standard approach to bypass this shortcoming of direct application of the Normal Copula model is the Base Correlation approach.

We note that the Base Correlation approach does not present a different model; it is still based on the Normal Copula model as in compound correlations, but applies the model in a different way. However, even the Base Correlation approach leaves a more fundamental conflict between the model and the practise unresolved. On one hand, the model assumes a single correlation for the entire portfolio. On the other hand, implying correlations leads to different estimated base correlations for different attachment points. To take up the challenge of resolving such a conflict is beyond the scope of this paper. Here we hope that a better understanding of the simple Normal Copula model can provide a more solid foundation for future development of a more flexible model. Therefore, as properties of the basic Normal Copula model, the non-uniqueness problems identified in this paper are worth studying. We hope our analytical approach provides some intuitive understanding of the Normal Copula model, its shortcomings and its insights.

References

Gupton, Greg M., Christopher C. Finger, and Mickey Bhatia, 1997, CreditMetrics – Technical Document, Morgan Guaranty Trust Co.

Hull, John, and Alan White, 2004, Valuation of a CDO and nth to Default CDS Without Monte Carlo Simulation, *Journal of Derivatives* 12, 8–23.

Li, David X., 2000, On Default Correlation: A Copula Function Approach, *Journal of Fixed Income* 9, 43–54.

McGinty, Lee, 2004, A Model for Base Correlation Calculation, J.P. Morgan Technical Paper.

Reyfman, Alex, 2004, Valuing and Hedging Synthetic CDO Tranches Using Base Correlations, Bear Stearns Technical Paper.

A Proof of claim 1

$f(\lambda) = P[\max(x_i(\lambda)) < x_0] = P[\max d_i(\lambda)]$ Assume $0 \leq \lambda_1 < \lambda_2 \leq 1$. Define $x_{i,j} = x_i(\lambda_j)$. When $x_{i,2} < x_0 < 0$ and $Z_i \geq 0$, we have $Z_0 < 0$. Therefore,

$$\begin{aligned}
& \left(X_{i,1} + (\lambda_2 - \lambda_1)Z_0 - \left(\sqrt{1 - \frac{N-1}{N}\lambda_1^2} - \sqrt{1 - \frac{N-1}{N}\lambda_2^2} + \frac{\lambda_2 - \lambda_1}{\sqrt{N}}\right) Z_i < x_0\right) \\
& .(Z_i \geq 0). \left(Z_i = \max_{j=1,\dots N} Z_j\right) \\
= & \left(X_{i,1} + (\lambda_2 - \lambda_1)Z_0 - \frac{\lambda_2 - \lambda_1}{\sqrt{N}} \left(\frac{(N-1)(\lambda_1 + \lambda_2)}{\sqrt{N - (N-1)\lambda_1^2} + \sqrt{N - (N-1)\lambda_2^2}} + 1\right) Z_i < x_0\right) \\
& .(Z_i \geq 0). \left(Z_i = \max_{j=1,\dots N} Z_j\right) \\
\leq & \ (X_{i,1} < x_0).(Z_i \geq 0). \left(Z_i = \max_{j=1,\dots,N} Z_j\right).
\end{aligned}$$

On the other hand, notice that

$$1 - \frac{\lambda_1 + \lambda_2}{\sqrt{N - (N-1)\lambda_1^2} + \sqrt{N - (N-1)\lambda_2^2}} > 0$$

for $0 \leq \lambda_1 < \lambda_2 \leq 1$. Therefore, we have

$$\begin{aligned}
& \left(X_{i,1} + (\lambda_2 - \lambda_1)Z_0 - \left(\sqrt{1 - \frac{N-1}{N}\lambda_1^2} - \sqrt{1 - \frac{N-1}{N}\lambda_2^2} + \frac{\lambda_2 - \lambda_1}{\sqrt{N}}\right) Z_i < x_0\right) \\
& .(Z_i < 0). \left(Z_i = \max_{j=1,\dots N} Z_j\right) \\
\geq & \left(X_{i,1} + \sqrt{N}(\lambda_2 - \lambda_1)Z_i - \left(\sqrt{1 - \frac{N-1}{N}\lambda_1^2} - \sqrt{1 - \frac{N-1}{N}\lambda_2^2} + \frac{\lambda_2 - \lambda_1}{\sqrt{N}}\right) Z_i < x_0\right) \\
& .(Z_i < 0). \left(Z_i = \max_{j=1,\dots N} Z_j\right) \\
& \left(X_{i,1} + \frac{(\lambda_2 - \lambda_1)(N-1)}{\sqrt{N}} \left[1 - \frac{\lambda_1 + \lambda_2}{\sqrt{N - (N-1)\lambda_1^2} + \sqrt{N - (N-1)\lambda_2^2}}\right] Z_i < x_0\right) \\
& .(Z_i < 0). \left(Z_i = \max_{j=1,\dots N} Z_j\right) \\
\geq & \ (x_{i,1} < x_0).(Z_i < 0).(Z_i = \max_{j=1,\dots,N} Z_j).
\end{aligned}$$

Combining these two inequalities, we have

$$
\begin{aligned}
& (X_{i,2} < x_0)\,.\left(Z_i = \max_{j=1,\ldots N}\right) \\
= \;& \left(X_{i,1} + (\lambda_2 - \lambda_1)Z_0 - \left(\sqrt{1-\frac{N-1}{N}\lambda_1^2} - \sqrt{1-\frac{N-1}{N}\lambda_2^2} + \frac{\lambda_2-\lambda_1}{\sqrt{N}}\right) Z_i < x_0\right) \\
& .\left(Z_i = \max_{j=1,\ldots N} Z_j\right) \\
= \;& \left(X_{i,1} + (\lambda_2 - \lambda_1)Z_0 - \left(\sqrt{1-\frac{N-1}{N}\lambda_1^2} - \sqrt{1-\frac{N-1}{N}\lambda_2^2} + \frac{\lambda_2-\lambda_1}{\sqrt{N}}\right) Z_i < x_0\right) \\
& .(Z_i < 0).\left(Z_i = \max_{j=1,\ldots N} Z_j\right) \\
& + \left(X_{i,1} + (\lambda_2 - \lambda_1)Z_0 - \left(\sqrt{1-\frac{N-1}{N}\lambda_1^2} - \sqrt{1-\frac{N-1}{N}\lambda_2^2} + \frac{\lambda_2-\lambda_1}{\sqrt{N}}\right) Z_i < x_0\right) \\
& .(Z_i \geq 0).\left(Z_i = \max_{j=1,\ldots N} Z_j\right) \\
\geq \;& (x_{i,1} < x_0)\,.(Z_i < 0).\left(Z_i = \max_{j=1,\ldots,N} Z_j\right) + (x_{i,1} < x_0)\,.(Z_i \geq 0).\left(Z_i = \max_{j=1,\ldots,N} Z_j\right) \\
= \;& (X_{i,1} < x_0)\,.\left(Z_i = \max_{j=1,\ldots N} Z_j\right)
\end{aligned}
$$

Therefore, $P\left[(X_{i,2} < x_0)\,.\,(Z_j = \max_{j=1,\ldots N})\right] \geq P\left[(X_{i,1} < x_0)\,.\,(Z_i = \max_{j=1,\ldots N})\right]$ Notice that the probability that at least two z_i's are the same is zero. Also, in the case $\lambda > 0$, $(x_i > x_j) = (z_i > z_j)$. Therefore,

$$
\begin{aligned}
f(\lambda_2) = \;& \sum_{i=1}^{N} P\left[(X_{i,2} < x_0).\left(X_{i,2} = \max_{j=1,\ldots,N} X_{j,2}\right)\right] \\
= \;& \sum_{i=1}^{N} P\left[(X_{i,2} < x_0).\left(Z_i = \max_{j=1,\ldots,N} Z_j\right)\right] \\
= \;& \sum_{i=1}^{N} P\left[\left(X_{i,1} + (\lambda_2 - \lambda_1)Z_0 - \left(\sqrt{1-\frac{N-1}{N}\lambda_1^2} - \sqrt{1-\frac{N-1}{N}\lambda_2^2} + \frac{\lambda_1-\lambda_2}{\sqrt{N}}\right) Z_i < x_0\right)\right. \\
& \left. .\left(Z_i = \max_{j=1,\ldots,N} Z_j\right)\right] \\
\geq \;& \sum_{i=1}^{N} P\left[(X_{i,1} < x_0).\left(Z_i = \max_{j=1,\ldots,N} Z_j\right)\right] \\
= \;& f(\lambda_1)
\end{aligned}
$$

Thus f is an increasing function in λ. Furthermore, $f(1) = p$ and $f(0) = p^N$. That is, as the correlation increases, the probability of all bonds defaulting increases, with the upper bound being $f(1) = p$, and the lower bound $f(0) = p^N$.

□

B Proof of Claim 2

Notice that by symmetry, we have $P[x_{i,1} < x_0] = P[x_{i,2} < x_0]$. Now notice that compared to $x_{i,1}$, $x_{i,2}$ puts more weights on z_i and less weight on z_0, which is the average across all z_j with $j = 1, \ldots, N$. Therefore, conditional upon $z_i = \max_{j=1,\ldots,N} z_j$, $x_{i,2}$ is likely to be larger than $x_{i,1}$. It is clear that

$$P\left[x_{i,1} < x_0 | z_i = \max_{j=1,\ldots,N} z_j\right] < P\left[x_{i,2} < x_0 | z_i = \max_{j=1,\ldots,N} z_j\right].$$

Therefore, $P[x_{i,1} < x_0; z_i = \max_{j=1,\ldots,N} z_j] < P[x_{i,2} < x_0; z_i = \max_{j=1,\ldots,N} z_j]$. The remainder of the proof follows as in the end of the proof of claim 1.

□

C Proof of Proposition 3

$$P[k|M] = \binom{N}{k} \Phi\left(\frac{x_0 - \alpha M}{\sqrt{1-\alpha^2}}\right)^k \Phi\left(-\frac{x_0 - \alpha M}{\sqrt{1-\alpha^2}}\right)^{N-k}$$

Notice that $P[k|M] = E[(D = k)|M]$. Thus,

$$P[k] = E[(D = k)] = E[P[k|M]]$$

Change variables such that $z = \frac{x_0 - \alpha M}{\sqrt{1-\alpha^2}}$. Notice that $z \sim N\left[\frac{x_0}{\sqrt{1-\alpha^2}}, \frac{\alpha^2}{1-\alpha^2}\right]$. And in the case that $x_0 = 0$, $z \sim N\left[0, \frac{\alpha^2}{1-\alpha^2}\right]$. We then have

$$P[k|z] = \binom{N}{k} \Phi(z)^k \Phi(-z)^{N-k}$$

$$\begin{aligned} P[k] &= E[P[k|z]] \\ &= \int_{-\infty}^{\infty} P[k|z] \frac{\sqrt{1-\alpha^2}}{\sqrt{2\pi}\alpha} \exp\left[-\frac{z^2(1-\alpha^2)}{2\alpha^2}\right] dz. \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial \alpha} P[k] &= \int_{-\infty}^{\infty} P[k|z] \frac{\sqrt{1-\alpha^2}}{\sqrt{2\pi}\alpha} \exp\left[-\frac{z^2(1-\alpha^2)}{2\alpha^2}\right] \frac{1}{\alpha^3}\left(z^2 - \frac{\alpha^2}{1-\alpha^2}\right) dz \\ &= E\left[P[k|z] \frac{1}{\alpha^3}\left(z^2 - \frac{\alpha^2}{1-\alpha^2}\right)\right] \\ &= \binom{N}{k} \frac{1}{\alpha^3} E\left[\left(\Phi(z)^k \Phi(z)^{N-k} + \Phi(-z)^k \Phi(z)^{N-k}\right) \cdot \left(z^2 - \frac{\alpha^2}{1-\alpha^2}\right)(z > 0)\right] \end{aligned}$$

Notice that $E\left[z^2 - \frac{\alpha^2}{1-\alpha^2}\right] = 0$.
Denote $h(z) = \Phi(z)^k\Phi(-z)^{N-k}$. Then

$$\frac{\partial}{\partial\alpha}P[k] = \binom{N}{k}\frac{1}{\alpha^3}E\left[(h(z)+h(-z))\cdot\left(z^2-\frac{\alpha^2}{1-\alpha^2}\right)(z>0)\right]$$

$$\begin{aligned}\frac{\partial}{\partial z}h(z) &= h(z)\frac{N\phi(z)}{\Phi(z)\Phi(-z)}\left(\frac{k}{N}-\Phi(z)\right)\\ &= N\phi(z)\Phi(z)^{k-1}\Phi(-z)^{N-k-1}\left(\frac{k}{N}-\Phi(z)\right)\end{aligned}$$

Several special cases warrant consideration:

1. For the case that $k = \frac{N}{2}$, we have $\frac{\partial}{\partial z}(h(z)+h(-z)) = 2N(\Phi(z)\Phi(-z))^{N/2-1}\left(\frac{1}{2}-\Phi(z)\right)\phi(z) < 0$, for all $z > 0$. Therefore

$$\begin{aligned}\frac{\partial}{\partial\alpha}P\left[\frac{N}{2}\right] &= \binom{N}{N/2}\frac{1}{\alpha^3}E\left[(h(z)+h(-z))\left(z^2-\frac{\alpha^2}{1-\alpha^2}\right)(z>0)\right]\\ &= \binom{N}{N/2}\frac{1}{\alpha^3}E\left[(h(z)+h(-z))\left(z^2-\frac{\alpha^2}{1-\alpha^2}\right)\right.\\ &\quad\left.\times\left((0<z<\frac{\alpha}{\sqrt{1-\alpha^2}})+(\frac{\alpha}{\sqrt{1-\alpha^2}}<z)\right)\right]\\ &< \binom{N}{N/2}\frac{1}{\alpha^3}\left\{E\left[\left(h\left(\frac{\alpha}{\sqrt{1-\alpha^2}}\right)+h\left(-\frac{\alpha}{\sqrt{1-\alpha^2}}\right)\right)\left(z^2-\frac{\alpha^2}{1-\alpha^2}\right)\right.\right.\\ &\quad\left.\times\left(0<z<\frac{\alpha}{\sqrt{1-\alpha^2}}\right)\right]\\ &\quad\left.+E\left[\left(h\left(\frac{\alpha}{\sqrt{1-\alpha^2}}\right)+h\left(-\frac{\alpha}{\sqrt{1-\alpha^2}}\right)\right)\left(z^2-\frac{\alpha^2}{1-\alpha^2}\right)\left(\frac{\alpha}{\sqrt{1-\alpha^2}}<z\right)\right]\right\}\\ &= \binom{N}{N/2}\frac{1}{\alpha^3}\left(h\left(\frac{\alpha}{\sqrt{1-\alpha^2}}\right)+h\left(-\frac{\alpha}{\sqrt{1-\alpha^2}}\right)\right)E\left[\left(z^2-\frac{\alpha^2}{1-\alpha^2}\right)\right]\\ &= 0.\end{aligned}$$

   Thus the centre of the distribution declines as α increases, i.e. $P\left[\frac{N}{2}\right]$ is a decreasing function of α.

2. For the case that $k = 0$, $\frac{\partial}{\partial z}(h(z)+h(-z)) = N\phi(z)\left(\Phi(z)^{N-1}-\Phi(-z)^{N-1}\right) > 0$, for $z > 0$. Therefore

$$\frac{\partial}{\partial\alpha}P[0] > \frac{1}{\alpha^3}\left(h\left(\frac{\alpha}{\sqrt{1-\alpha^2}}\right)+h\left(-\frac{\alpha}{\sqrt{1-\alpha^2}}\right)\right)E\left[\left(z^2-\frac{\alpha^2}{1-\alpha^2}\right)\right] = 0.$$

   Thus the left tail always goes up, i.e. $P[0]$ is an increasing function of α.

3. For the case that $k = N$, by symmetry, we have $P[N]$ is an increasing function of α. Thus we here reproved in, 2. and 3., proposition for the special case when $x_0 = 0$.

Now we turn our attention to the general case of k. Note first that $\frac{\partial}{\partial z}(h(z)+h(-z))|_{z=0} = h'(0) - h'(0) = 0$.

$$h''(z) = h'(z)\phi(z)\left(\frac{k-1}{\Phi(z)} - \frac{N-k-1}{\Phi(-z)} - \frac{1}{\frac{k}{N}-\Phi(z)} + \frac{\phi'(z)}{\phi^2(z)}\right).$$

Therefore, $h''(0) = 2^{5-N}\phi^2(0)\left(\frac{N-\sqrt{N}}{2}-k\right)\left(\frac{N+\sqrt{N}}{2}-k\right)$. For $0 < k < \frac{N-\sqrt{N}}{2}$ or $\frac{N+\sqrt{N}}{2} < k < N$, we have $h'(0) > c_{k,N} > 0$, for some positive constant $c_{k,N}$ dependent on k and N but not on α. We now denote $g(z) = h(z) + h(-z)$. Then $g''(0) = 2h''(0) > 2c_{k,N}$. By continuity of $g''(z)$, we have that, for some constant $z_{k,N}$ independent of α, $g''(z) > c_{k,N}$ for all $k \in [0, z_{k,N}]$. This implies that $g(z)$ is increasing in the interval $[0, z_{k,N}]$. Let α be such that $\frac{\alpha}{\sqrt{1-\alpha^2}} < z_{k,N}$. We then have, for $z \in \left[\frac{\alpha}{\sqrt{1-\alpha^2}}, z_{k,N}\right]$,

$$g(z) \geq g\left(\frac{\alpha}{\sqrt{1-\alpha^2}}\right) + c_{k,N}\left(z - \frac{\alpha}{\sqrt{1-\alpha^2}}\right)^2.$$

Given these properties of the function $g(z)$, we derive the following inequalities:

$$\begin{aligned}
&\frac{\alpha^3 \frac{\partial}{\partial\alpha}P[k]}{\binom{N}{k}} \\
= \;& E\left[g(z)\left(z^2 - \frac{\alpha^2}{1-\alpha^2}\right)(z>0)\right] \\
= \;& E\left[g(z)\left(z^2 - \frac{\alpha^2}{1-\alpha^2}\right)\left(\left(0 < z < \frac{\alpha}{\sqrt{1-\alpha^2}}\right) + \left(\frac{\alpha}{\sqrt{1-\alpha^2}} < z < z_{k,N}\right) + (z_{k,N} < z)\right)\right] \\
> \;& c_{k,N} E\left[\left(z - \frac{\alpha}{\sqrt{1-\alpha^2}}\right)^2\left(z^2 - \frac{\alpha^2}{1-\alpha^2}\right)(z>0)\right] \\
& + E\left[\left(h\left(\frac{\alpha}{\sqrt{1-\alpha^2}}\right) + c_{k,N}\left(z - \frac{\alpha}{\sqrt{1-\alpha^2}}\right)^2\right)\left(z^2 - \frac{\alpha^2}{1-\alpha^2}\right)(z_{k,N}, z)\right]
\end{aligned}$$

It is now clear that, for sufficiently small α, we have $\frac{\partial}{\partial\alpha}P[k] > 0$, for $0 < k < \frac{N-\sqrt{N}}{2}$ or $\frac{N+\sqrt{N}}{2} < k < N$. [1] Note that, when $\alpha = 1$, $P[k] = 0$ for all $0 < k < N$. Therefore, we have the hump shape for $P[k]$ as a function of α, when $0 < k < \frac{N-\sqrt{N}}{2}$ or $\frac{N+\sqrt{N}}{2} < k < N$.

□

D Proof of Proposition 4

$$\begin{aligned}
P[k] &= E[P[k|z]] \\
&= \int_{-\infty}^{\infty} P[k|z]\frac{\sqrt{1-\alpha^2}}{\sqrt{2\pi}\alpha}\exp\left[-\frac{\left(z - \frac{x_0}{\sqrt{1-\alpha^2}}\right)^2(1-\alpha^2)}{2\alpha^2}\right]dz.
\end{aligned}$$

[1]Similarly, we can prove that $\frac{\partial}{\partial\alpha}P[k] < 0$, for $\frac{N-\sqrt{N}}{2} < k < \frac{N+\sqrt{N}}{2}$.

$$\frac{\partial}{\partial \alpha} P[k] = \frac{1}{\alpha^3} E\left[P[k|z] \left(\left(\left(z - \frac{x_0}{\sqrt{1-\alpha^2}} \right)^2 - \frac{\alpha^2}{1-\alpha^2} \right) + \frac{\alpha^2 x_0}{\sqrt{1-\alpha^2}} \left(z - \frac{x_0}{\sqrt{1-\alpha^2}} \right) \right) \right]$$

Change variables such that $y = z - \frac{x_0}{\sqrt{1-\alpha^2}}$. Then $y \sim N\left[0, \frac{\alpha^2}{1-\alpha^2}\right]$.

$$P[k|y] = \binom{N}{k} \Phi\left(y + \frac{x_0}{\sqrt{1-\alpha^2}} \right)^k \Phi\left(-y - \frac{x_0}{\sqrt{1-\alpha^2}} \right)^{N-k} = \binom{N}{k} h\left(y + \frac{x_0}{\sqrt{1-\alpha^2}} \right)$$

$$\begin{aligned}
\frac{\partial}{\partial \alpha} P[k] &= \frac{1}{\alpha^3} E\left[P[k|y] \left(\left(y^2 - \frac{\alpha^2}{1-\alpha^2} \right) + \frac{\alpha^2 x_0 y}{\sqrt{1-\alpha^2}} \right) \right] \\
&= \frac{1}{\alpha^3} \binom{N}{k} E\left[\left\{ \left(h\left(y + \frac{x_0}{\sqrt{1-\alpha^2}} \right) + h\left(-y + \frac{x_0}{\sqrt{1-\alpha^2}} \right) \right) \left(y^2 - \frac{\alpha^2}{1-\alpha^2} \right) \right. \right. \\
&\quad \left. \left. + \left(h\left(y + \frac{x_0}{\sqrt{1-\alpha^2}} \right) - h\left(-y + \frac{x_0}{\sqrt{1-\alpha^2}} \right) \right) \frac{\alpha^2 x_0 y}{\sqrt{1-\alpha^2}} \right\} (y > 0) \right]
\end{aligned}$$

In the case $\frac{k}{N} < \Phi(x_0)$, we have $h'(x_0) = h(x_0) \frac{N}{\Phi(-x_0)} \left(\frac{k}{N} - \Phi(x_0) \right) < 0$. By continuity of $h'(x)$, we have that, for positive constants δ and c independent of α, $h'(x) < -c/2 < 0$ for all $x \in [x_0 - 2\delta, x_0 + 2\delta]$. We denote $g_1(y) = h\left(y + \frac{x_0}{\sqrt{1-\alpha^2}} \right) - h\left(-y + \frac{x_0}{\sqrt{1-\alpha^2}} \right)$. Then

$$\frac{\partial}{\partial y} g_1(y) = \left(h'\left(y + \frac{x_0}{\sqrt{1-\alpha^2}} \right) + h'\left(-y + \frac{x_0}{\sqrt{1-\alpha^2}} \right) \right)$$

It is then clear that for $\alpha < \frac{\delta}{\delta + |x_0|}$, we have $\frac{\partial}{\partial y} g_1(y) < -c < 0$ for all $y \in [0, \delta]$. Further more, $g_1(0) = 0$. Therefore,for all $y \in [0, \delta]$, $g_1(y) < -cy$. It is then clear that

$$E\left[g_1(y) \frac{\alpha^2 x_0 y}{\sqrt{1-\alpha^2}} (y > 0) \right] > 0$$

for sufficiently small α.
Now denote $g_2(y) = h\left(y + \frac{x_0}{\sqrt{1-\alpha^2}} \right) + h\left(-y + \frac{x_0}{\sqrt{1-\alpha^2}} \right)$. Obviously, $g_2'(0) = 0$.

$$\begin{aligned}
h''(z) &= \phi(z) \Phi(z)^{k-2} \Phi(-z)^{N-k-2} \left\{ -kz\Phi(z)\Phi(-z)^2 + (N-k)z\Phi(z)^2\Phi(-z) \right. \\
&\quad \left. \phi(z) \left(k(k-1)\Phi(-z)^2 - 2k(N-k)\Phi(z)\Phi(-z) + (N-k)(N-k-1)\Phi(z)^2 \right) \right\}
\end{aligned}$$

Substitute x_0 into the above equation, and fix k. Then, $h''(x_0)$ can be viewed as a quadratic form of N with the second order coefficient being $\phi(z)^2 \Phi(z)^k \Phi(-z)^{N-k-2} > 0$. Therefore, for sufficiently large N, we have $h''(x_0) > 0$. Notice that $g_2'(0) = 2h''(x_0) > 0$. Following a similar argument as in the proof of Proposition 3, we are then able to show that

$$E\left[g_2(y) \left(y^2 - \frac{\alpha^2}{1-\alpha^2} \right) (y > 0) \right] > 0$$

for small α. Thus we have shown that

$$\begin{aligned}\frac{\partial}{\partial\alpha}P[k] &= \frac{1}{\alpha^3}\binom{N}{k}\left\{E\left[g_1(y)\frac{\alpha^2 x_0 y}{\sqrt{1-\alpha^2}}(y>0)\right]\right. \\ &\quad \left. +E\left[g_2(y)\left(y^2-\frac{\alpha^2}{1-\alpha^2}\right)(y>0)\right]\right\} \\ &> 0,\end{aligned}$$

for sufficiently small α, and large N. This with the fact that when $\alpha = 1$, $P[k] = 0$ for all $0 < k < N$ finishes the proof of the first statement of this proposition.
To prove the second statement, note that when we substitute $N-k$ for k in the expression of $g_1'(0)$, we get $h'(x_0)|_{N-k} = h(x_0)\frac{N}{\Phi(-x_0)}\left(\frac{N-k}{N} - \Phi(x_0)\right) > 0$ for N large, which will make $E\left[g_1(y)\frac{\alpha^2 x_0 y}{\sqrt{1-\alpha^2}}(y>0)\right] < 0$ for small α. However, it is easy to see that $h'(x_0)|_{N-k} < N$. On the other hand, when we substitute $N-k$ for k in the expression of $g_2''(0)$, we again obtain a quadratic form of N with the second order coefficient being $2\phi(z)^2\Phi(z)^{k-2}\Phi(-z)^{N-k} > 0$. We still have

$$E\left[g_2(y)\left(y^2-\frac{\alpha^2}{1-\alpha^2}\right)(y>0)\right] > 0$$

for small α, and it is of second order of N, which makes it the dominating term comparing to $E\left[g_1(y)\frac{\alpha^2 x_0 y}{\sqrt{1-\alpha^2}}(y>0)\right]$. Therefore, we still have $\frac{\partial}{\partial\alpha}P[k] > 0$ for sufficiently small α, and large N. Thus the second statement of the proposition is proved.
The third and fourth statements of the proposition follow from symmetry.

□