Journal of Agricultural and Resource Economics 34(1):53–67

Techniques for Multivariate Simulation from Mixed Marginal Distributions with Application to Whole-Farm Revenue Simulation

John D. Anderson, Ardian Harri, and Keith H. Coble

Alternative techniques for representing dependencies among variables in multivariate simulation are discussed and compared in the context of rating a whole-farm insurance product. A procedure by Iman and Conover (IC) that is common in actuarial applications is compared to a new technique detailed by Phoon, Quek, and Huang (PQH). Results suggest that rates derived from the IC procedure may be inaccurate because the procedure produces biased estimates of correlation between simulated variables. This situation is improved with the PQH procedure.

Key words: correlation, crop insurance, multivariate simulation, whole-farm revenue

Introduction

Monte Carlo simulation represents an important analytical tool for evaluating a wide array of complex problems. Simulation of independent variables from parametric distributions or nonparametrically from empirical distributions is straightforward; however, modeling dependent variables from a joint distribution is considerably more difficult. This difficulty is compounded if the marginal distributions comprising the joint distribution are not consistent. As Biller and Nelson (2003) note, relatively few tools are readily available for simulating correlated variables from different marginal distributions (or different families of marginal distributions). Still, the number of applications requiring a flexible simulation procedure is substantial. For example, the calculation of rates for multi-crop revenue insurance requires the simulation of multiple prices and yields. So, too, does the analysis of the costs and benefits of whole-farm, revenue-based policy instruments—a topic which has taken on greater significance this year with a new Farm Bill that includes a choice between continued participation in the preexisting set of commodity programs or a new whole-farm, revenue-based countercyclical program. More generally, almost any risk analysis going beyond a single enterprise will encounter these problems. Thus, fundamental management issues such as diversification, portfolio analysis, and many investment decisions require this kind of modeling capability.

A number of techniques for simulating correlated random variables—including from mixed marginal distributions—do exist in the literature. For instance, Iman and Conover

John D. Anderson is extension professor, Ardian Harri is assistant professor, and Keith H. Coble is professor, all in the Department of Agricultural Economics, Mississippi State University.

Review coordinated by David K. Lambert.

(IC) (1982) present a procedure for correlating simulated variables from mixed marginal distributions in the multivariate case that has become standard in actuarial applications. More recently, Phoon, Quek, and Huang (PQH) (2004) describe a procedure for simulating correlated variables from mixed marginal distributions based on an eigen decomposition of the rank correlation matrix. While these procedures are both appropriate for the same applications, to date, no empirical work has been done comparing their performance.

Although the potential application of multivariate simulation techniques in the field of agricultural economics is essentially unlimited, the literature provides little detail on how to implement such procedures. With this fact in view, the objectives of our paper are threefold. First, the implementation of the PQH simulation procedures will be described in detail. Second, this procedure will be applied to a simulation of rates for a whole-farm revenue product. PQH results will be compared to results from the more familiar IC procedure. Within this context, the performance of both procedures in reproducing specified marginals and correlations will be compared. Finally, the robustness of each procedure to factors such as distribution parameters, alternative correlation matrices, and sample size will be investigated.

The exposition of a flexible multivariate simulation procedure will provide a source in the economics literature for a guide to programming the procedure. Moreover, results of the empirical application presented in this study will contribute useful information to aid practitioners in the selection of simulation methodologies for evaluating other specific applied problems in agricultural economics.

Review of Common Simulation Procedures

A number of techniques for simulating correlated random variables exist in the literature. Naylor et al. (1966) describe the programming of a methodology for simulating correlated variables from a multivariate normal distribution [Krzanowski (1988) provides a very accessible explanation of the theoretical underpinnings of this procedure]. Clements, Mapp, and Eidman (1971) adapt Naylor et al.'s procedure to farm revenue simulation, and their application of this technique is widely cited in the agricultural economics literature (e.g., Bailey and Richardson, 1985; Anderson and Zeuli, 2001). For the simulation of prices (for example, where the assumption of lognormality is not overly restrictive), this procedure is useful; however, if the assumption of normality is not well-supported (for example, if both prices and yields are being simulated), the restriction to the multivariate normal distribution is a major limitation.

Johnson and Tenenbein (1981) describe a procedure for simulating from a bivariate distribution with mixed marginals. Babcock and Hennessy (1996) illustrate the use of the Johnson-Tenenbein procedure in simulating correlated prices and yields from mixed marginal distributions (a beta distribution for yields and a lognormal distribution for prices) in the context of crop insurance rating. This procedure is quite flexible, accommodating simulation from either parametric or empirical distributions. While the ability to simulate from mixed marginals is attractive, the Johnson-Tenenbein procedure is not readily extended beyond the bivariate case. Hennessy, Babcock, and Hayes (1997) use a variation of the Johnson-Tenenbein procedure in a multivariate simulation of whole-farm revenue assurance (RA) rates. However, Hart, Hayes, and Babcock (2006) caution that the procedure is not capable of reliably maintaining correlations between more than two variables at a time.

As noted earlier, the Iman and Conover (IC) procedure for correlating simulated variables from mixed marginal distributions in the multivariate case has become standard in actuarial practice. In fact, Mildenhall (2005) opines that the procedure "should be part of every actuary's toolkit." The IC procedure involves simulating independent variables and then re-sorting them using information in the correlation matrix. This essentially brute-force approach is computationally intensive, particularly if the number of variables being simulated is large. Moreover, Ferson et al. (2004) note that the procedure has been criticized for its ad hoc nature. Nonetheless, the procedure has gained wide currency, owing largely to its intuitive appeal, its relative simplicity, and its suitability to a wide variety of programming environments.

Hart, Hayes, and Babcock (2006) present an application of the IC procedure to whole-farm crop insurance rating. Paulson and Babcock apply the procedure to the modeling of alternative grain contract structure (2007) as well as to an area revenue insurance instrument (2008). In the literature of many other disciplines, the IC procedure is pervasive, finding application in engineering (e.g., Vorechovsky and Chudoba, 2006), environmental sciences (e.g., Webster et al., 2003), and agronomy (e.g., Tattari et al., 2001), among many others.

The PQH procedure is similar to IC in that it is—to borrow the language of Iman and Conover—a "distribution-free" technique, allowing for the simulation of correlated variables from mixed marginal distributions, including empirical distributions. Their procedure consists of the simulation of correlated probabilities using information in the correlation matrix. These probabilities are used in an inverse transformation of the relevant marginal distribution to produce correlated variables from the simulation. To date, this procedure is not well-established in the literature. In one of the few applications of PQH outside of the engineering literature, Anderson, Coble, and Miller (2007) apply the procedure to the simulation of risk management strategies related to the countercyclical payment program in the 2002 Farm Bill.

Copula procedures represent another approach to modeling dependence among variables. A copula is simply a function (alternatively referred to as a dependence function) that defines the relationship of marginal distributions to their full multivariate distribution (Frees and Valdez, 1998). Clemens and Reilly (1999) present a good overview of copula procedures, including examples using empirical data. Simulation using the copula approach can be found in the agricultural economics literature. The Johnson-Tenenbein procedure referenced earlier is essentially a copula approach to bivariate simulation. More recently, Zhang et al. (2007) adopt a copula approach to simulate correlated prices and yields from a bivariate distribution using two different families of copulas (Gaussian and Frank) for comparison. Zhu, Ghosh, and Goodwin (2008) apply a copula approach in their development of a multi-crop insurance product design.

This work focuses on the PQH procedure because of the limited treatment it has received in the literature, because of its flexibility (suitability for both parametric and nonparametric applications), and because of its potential for application to very timely issues in agricultural economics research (e.g., multi-crop revenue and whole-farm policy instruments). Comparison with the IC procedure is presented since IC is the alternative technique most directly analogous to PQH and because of the familiarity with IC across a broad range of disciplines.

Methodology

The PQH procedure is, in general terms, a translation process, that is, the simulation of a non-Gaussian process based on a nonlinear transformation of an underlying Gaussian process (Phoon, Quek, and Huang, 2004). Fundamentally, the procedure involves using a Karhunen-Loeve (KL) expansion in the simulation of correlated normal deviates that are used as probabilities in an inverse transformation on the desired marginal distributions.

The KL expansion is a well-established approach to representing a variety of stochastic processes and is widely employed as a signal processing technique (e.g., see Hua, 1998). Van Trees (1968) presents the KL expansion of a Gaussian process, ω_k, with mean $\bar{\omega}_k$:

(1) $$\omega_k = \bar{\omega}_k + \sum_k \lambda_k f_k(x) f_k(x),$$

where λ_k and $f_k(x)$ are eigenvalues and eigenvectors, respectively, of the covariance function. Phoon, Huang, and Quek (2002) adapt this expression of the KL expansion to the simulation of a stochastic process using a vector of randomly generated independent standard normal variables, $\xi_k(\theta)$, as follows:

(2) $$\omega_k = \bar{\omega}_k + \sum_k \sqrt{\lambda_k}\,\xi_k(\theta) f_k(x).$$

PQH extend this technique to the simulation of multivariate non-Gaussian processes, noting that the fractile correlation (or other nonparametric measures of association such as Kendall's τ) of a nonnormal process will be the same as that of the underlying normal process since it is invariant to monotone transformations. The procedure described by PQH for simulating k random correlated variables starting with a rank correlation matrix and defined marginal distributions (including, as noted, the possibility of using empirical distributions) is implemented in five steps:

a. Convert rank correlation, $\mathbf{S}$, to Pearson correlation, $\boldsymbol{\rho}$, using

(3) $$\boldsymbol{\rho} = 2\sin[(\pi/6)\mathbf{S}].$$

b. Compute eigenvalues, λ_k, and eigenvector, $f_k(x)$, from $\boldsymbol{\rho}$ and confirm that $\boldsymbol{\rho}$ is nonnegative definite (i.e., that the minimum eigenvalue is equal to or greater than zero).

c. Use these eigen solutions in the KL expansion of a standard normal process to derive correlated standard normals, ω_k, as follows:

(4) $$\omega_k = \sum_k \sqrt{\lambda_k}\,\xi_k(\theta) f_k(x).$$

where $\xi_k(\theta)$ are independent standard normal variables, and other variables are as previously defined.

d. Determine probability associated with each of the correlated standard normal deviates via transformation on the standard normal cumulative distribution function (CDF).

Step (d) results in a vector of probability values that are correlated according to the rank correlation used at the beginning of the procedure. The final step of the PQH procedure is:

e. Translate correlated probabilities from step (d) into simulated outcomes by inverse transformation on the desired marginal distribution.

Numerical Comparison of IC and PQH Data Simulation

To evaluate the performance of the IC and PQH procedures, a simulation of six correlated variables with mixed marginal distributions was performed. The assumed correlation across variables is presented in table 1. Three of the variables were simulated from lognormal distributions differing in mean and volatility. The other three variables were simulated from beta distributions with different parameters a and b. These marginal distributions were chosen because lognormal and beta distributions are frequently used to model prices and yields, respectively, in the agricultural economics literature (e.g., see Babcock and Hennessy, 1996). Distributional parameters for this simulation are summarized in table 2.

For each procedure, the mean and standard deviation of each element of the correlation matrix are calculated across 1,000 samples of simulated data. Results are generated for sample sizes of 500, 1,000, 5,000, and 10,000; t-tests evaluate the difference between simulated correlation coefficients and the correlations prescribed in table 1.[1] Results of these tests are reported in table 3 for a sample size of 10,000. Note that the figures in this table are t-values comparing the simulated and actual correlations corresponding to the given position in the table. Clearly, t-tests suggest the PQH procedure provides a more accurate representation of the dependencies between variables than does the IC procedure. In fact, all of the correlation coefficients in the IC simulated data were significantly different from the prescribed value (at $p < 0.01$), while only one in the PQH simulated data was significantly different from the prescribed correlation value.[2]

To gain further insight into differences in the simulation outcomes from the two procedures, the mean squared error (MSE) was calculated across the 1,000 samples for three of the correlation coefficients: r_{P1Y1}, r_{P1P2}, and r_{Y1Y3}.[3] MSE values for r_{P1Y1} are shown in figure 1 and are representative of the results for the other correlation coefficients. The interesting feature of this graph is that the MSE of the correlation coefficient in the PQH simulation is relatively large in smaller samples but declines dramatically as sample size increases so that by a sample size of 5,000, MSE is smaller in the PQH simulated data than in the IC simulated data.

[1] A number of alternative correlation structures were investigated, and results of the comparison of the two procedures were very robust. Thus, our discussion of results will be confined to those associated with the correlation matrix in table 1.

[2] Note that the difference between the PQH and IC procedures is how they impose correlations on simulated data. The inverse transformation on marginal distributions is the same across both procedures. Hence, the discussion here focuses on simulated correlations (which may be expected to differ some across the two procedures) rather than on moments of the marginal distributions (which should not be expected to differ across the two procedures).

[3] In this context, the mean squared error is based on the difference between the sample mean for the correlation coefficient in question and the prescribed value for the correlation coefficient (see table 1).

Table 1. Rank Correlation Matrix Used in Simulating Multi-Crop Revenue Insurance Premium

	*P*1	*Y*1	*P*2	*Y*2	*P*3	*Y*3
*P*1	1.00	−0.35	0.50	−0.15	0.10	−0.05
*Y*1		1.00	−0.30	0.70	−0.08	0.30
*P*2			1.00	−0.42	0.28	−0.12
*Y*2				1.00	−0.07	0.25
*P*3					1.00	−0.20
*Y*3						1.00

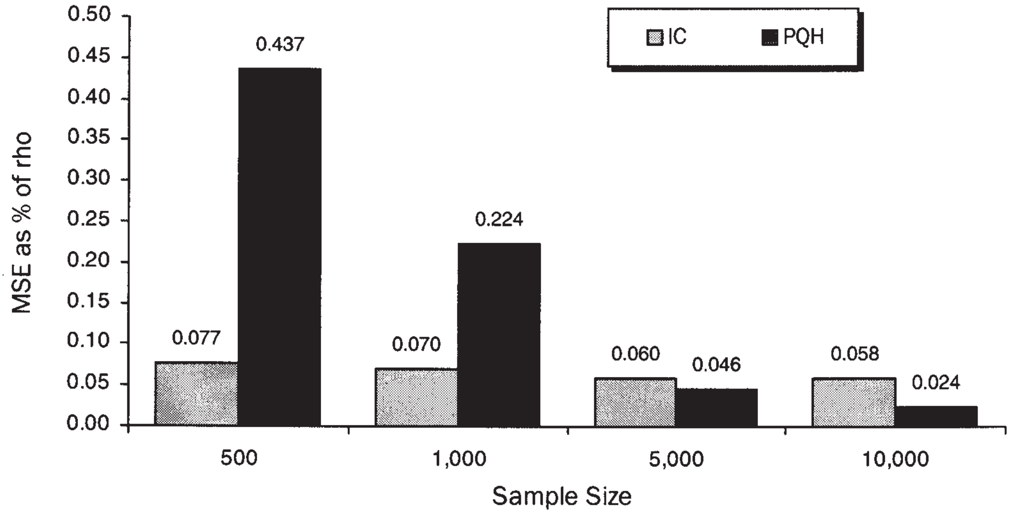
Table 2. Distribution Parameters Used in Multivariate Simulation to Compare Performance of IC and PQH Correlation Procedures

	Distribution Type and Parameters			
	Lognormal		Beta	
Variable	μ	σ	*a*	*b*
*P*1	3.00	0.10		
*P*2	8.00	0.30		
*P*3	5.00	0.20		
*Y*1			0.50	0.75
*Y*2			0.25	0.80
*Y*3			0.40	0.85

Table 3. *t*-Test Comparing Correlation Coefficients from PQH and IC Simulated Data to Correlation Coefficients Prescribed in Table 1

	*Y*1		*P*2		*Y*2		*P*3		*Y*3	
	PQH	IC	PQH	IC	PQH	IC	PQH	IC	PQH	IC
*P*1	−0.75	***−312.80***	−0.28	***300.81***	−0.12	***−88.93***	−0.36	***101.95***	0.48	***−45.87***
*Y*1			0.62	***−123.61***	−0.17	***264.99***	−0.79	***−61.05***	0.93	***201.00***
*P*2					0.95	***−182.57***	−0.82	***163.95***	***2.14***	***−69.72***
*Y*2							−0.21	***−37.93***	0.00	***124.24***
*P*3									−0.58	***−143.85***

Notes: *t*-values in bold/italics denote significance at $p < 0.01$; sample size = 10,000.



Note: Mean squared error presented as percentage of true correlation between $P1$ and $Y1$ (i.e., $\rho_{P1Y1} = -0.35$).

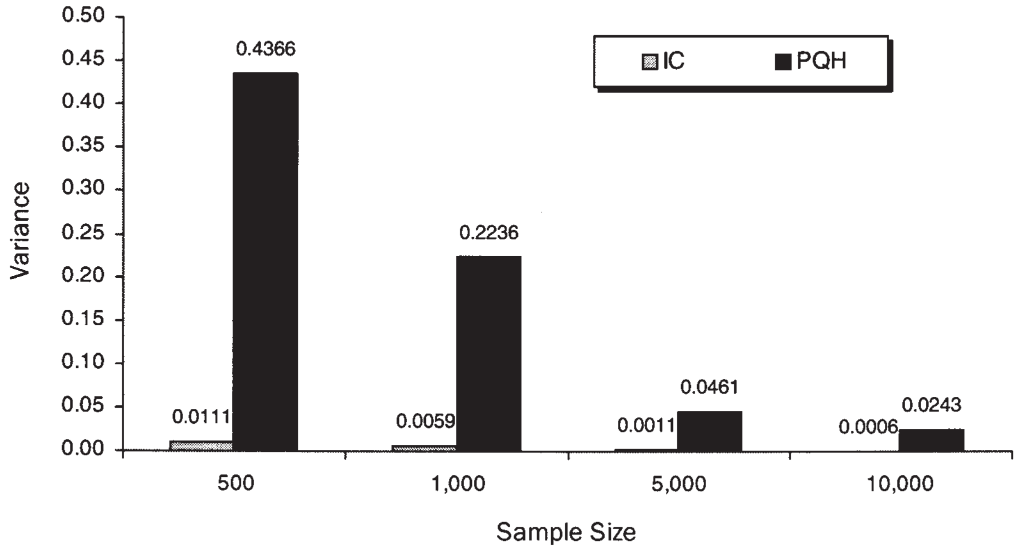
Figure 1. Mean squared error of simulated price/yield correlation coefficient using PQH and IC simulation procedures

Components of the MSE were decomposed for the r_{P1Y1} correlation coefficient for all sample sizes. Variance and bias of the correlation coefficient from both PQH and IC simulated data are shown in figures 2 and 3. The interesting point to note in these graphs is that, not surprisingly, the variance of the coefficient from the PQH simulation is quite large (in relative terms, at least) in the small sample but declines to a very small value in the larger samples. The variance associated with the coefficient from the IC simulated data is relatively small in all sample sizes. On the other hand, the coefficient from the IC simulated data displays a relatively large bias, and the magnitude of the bias is hardly mitigated by increasing sample size. By contrast, the bias associated with the coefficient from the PQH simulated data is relatively small, even in small sample sizes.

The bias of IC simulated correlations appears to be a general result, at least given the marginal distributions and covariance structure being modeled here. Figure 4 plots the bias of all 15 pairwise correlations for the six variables simulated here (for $N = 10{,}000$). Clearly, the IC simulated correlation coefficients display considerably more bias than the PQH simulated correlation coefficients—for every element of the correlation matrix. In most cases, the bias of the IC simulated correlation is several times greater than that of the PQH simulated correlations. Note that bias in figure 4 is expressed as a percentage of the actual correlation according to:

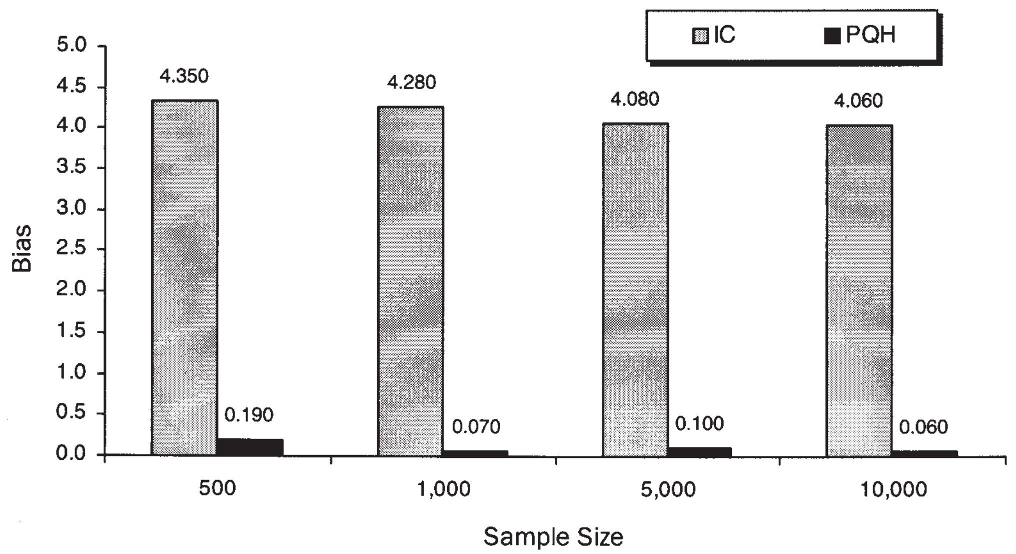
$$(5) \qquad Bias = \left(\rho_{actual} - \mathrm{E}\left[\rho_{simulated}\right]\right) / \rho_{actual}.$$

The fact that all of the IC bias values in figure 4 are positive is instructive. This finding indicates the IC procedure consistently simulates correlations that are smaller in absolute value than actual correlations.



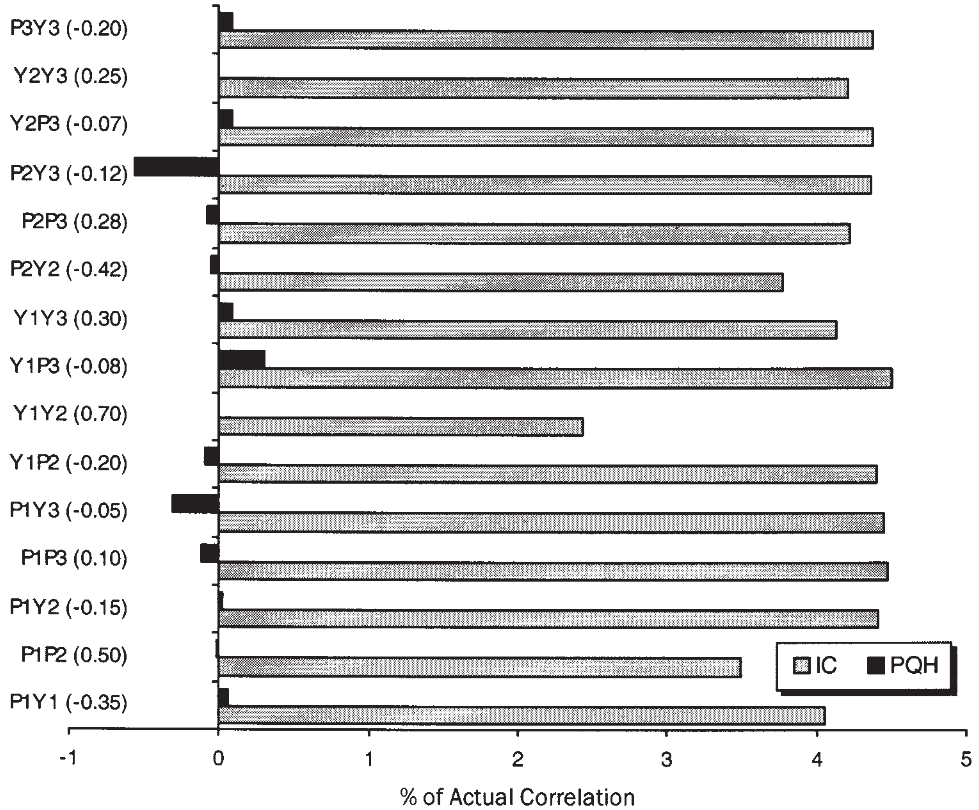
Note: Variance presented as percentage of true correlation between $P1$ and $Y1$ (i.e., $\rho_{P1Y1} = -0.35$).

Figure 2. Variance of simulated price/yield correlation coefficient using PQH and IC simulation procedures



Note: Bias presented as percentage of true correlation between $P1$ and $Y1$ (i.e., $\rho_{P1Y1} = -0.35$).

Figure 3. Bias of simulated price/yield correlation coefficient using PQH and IC simulation procedures

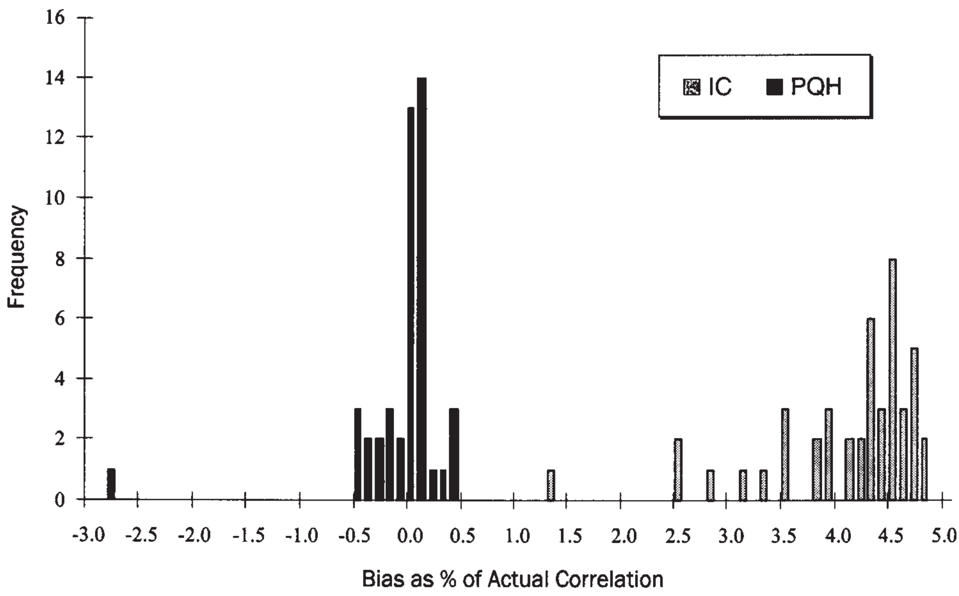


Note: Bias presented as percentage of true correlation coefficients (in parentheses next to category labels).

Figure 4. Bias of simulated correlation coefficients using PQH and IC simulation procedures

To further explore the bias of IC and PQH simulated correlations, two additional correlation matrices were simulated: one with smaller price/yield correlations and larger price/price correlations than those in table 1, and another with negative correlations between *P*3 and the other price variables. (This would be roughly consistent with a simulation of revenue for a farm producing both grain and livestock.) Figure 5 plots histograms for the bias values calculated for all 45 of the simulated correlations (15 pairwise correlations, three correlation matrices) from each procedure. The nature of the bias of the IC procedure noted above (i.e., that simulated correlations tend to be smaller in absolute value than actual correlations) appears to hold up when alternative correlation matrices are considered. Also of note is the fact that the PQH procedure seems more likely to produce outliers than the IC procedure. This result should be interpreted with a bit of caution, however. The largest bias values for either procedure were associated with very small correlation coefficients which, for all practical purposes, would likely be ignored. Considering only correlations greater than 0.10 in absolute value, bias values ranged from about −0.6% to +0.2% for the PQH procedure and from about +2.4% to +4.7% for the IC procedure.

There are a number of potential sources of this positive bias in the IC procedure. Iman and Conover caution that the distribution chosen for the "score" matrix (a set of



Note: Bias presented as percentage of true correlation coefficients.

Figure 5. Distribution of bias estimates for simulated correlation coefficients using PQH and IC simulation procedures

values used in the process of reordering the independently simulated values to achieve the desired correlation) can affect simulation results.[4] Though not specifically commenting on the issue of bias, Mildenhall (2005) notes that the IC procedure results in a more faithful reproduction of the desired correlation matrix when marginal distributions are symmetrical. The asymmetric nature of the distributions used in this study thus may be contributing to the bias in the IC results. It is possible, given the marginal distributions employed here, that a different specification of the score matrix could improve results. Evaluation of this issue is beyond the scope of this paper; however, this situation itself does represent one of the difficulties in implementing the IC procedure.

An Application to Whole-Farm Insurance Rating

To illustrate the practical implications of the differences in simulated data from the PQH and IC procedures, we model the rating of a whole-farm revenue insurance product consistent with the Revenue Assurance-Harvest Price Option (RA-HO). In this example, three crops are assumed to be produced. Since the primary purpose of this analysis is the evaluation of the simulation technique, hypothetical distributional parameters and correlations are assumed rather than derived from empirical data. This approach facilitates sensitivity analysis related to the parameters. The correlation matrix used in the simulation is reported in table 1. Parameters of the marginal distributions used in the analysis were changed from those reported in table 2 in order to be more consistent with

[4] In this study, following the preference expressed by Iman and Conover (1982), the standard normal distribution was used in generating the score matrix.

Table 4. Parameters of Beta (Yield) Distribution Used to Compare Performance of IC and PQH Correlation Procedures in Whole-Farm Rating Exercise

	Distribution Parameters			
Variable	*a*	*b*	Upper Bound	Lower Bound
Yield 1	2.67	1.78	300	50
Yield 2	2.08	1.03	90	15
Yield 3	2.93	3.47	95	12

actual distributions of prices and yields for corn, soybeans, and wheat. Specifically, the three lognormal distributions (price variables) were all assumed to have a volatility of 0.25. For the beta distributions (yield variables), parameters for the insurance rating simulation are summarized in table 4.

For the hypothetical multi-crop revenue coverage rated in this simulation (three possible crops), the per acre revenue guarantee, *TRG*, is defined as:

$$(6) \qquad TRG = CL \sum_{i=1}^{3} \tfrac{1}{3} \mathrm{E}(P_i)\mathrm{E}(Y_i),$$

where CL is the selected coverage level (75% in the analysis presented here), and $\mathrm{E}(P_i)$ and $\mathrm{E}(Y_i)$, respectively, are planting time expectations for price and yield for crop i. Since the revenue product is modeled with a harvest revenue option, the final per acre revenue guarantee, TRG^*, is defined as:

$$(7) \qquad TRG^* = CL \sum_{i=1}^{3} \tfrac{1}{3} \mathrm{Max}\left[\mathrm{E}(P_i), P_i\right]\mathrm{E}(Y_i),$$

where $\mathrm{E}(P_i)$ is the original planting time expectation of harvest price, P_i is the actual realized harvest time price, and other variables are as previously defined.

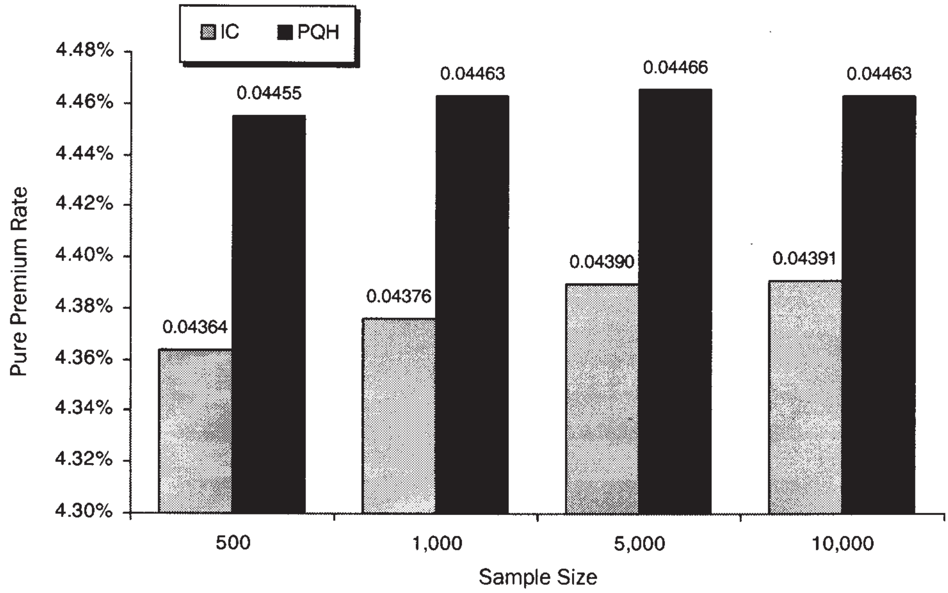
For this multi-crop product, a pure premium rate for coverage level c based on n simulated outcomes ($PPR_{c,n}$) is calculated as the average indemnity over n simulated outcomes[5] divided by the average initial revenue guarantee, or:

$$(8) \qquad PPR_{c,n} = \frac{\frac{1}{n}\sum_{j=1}^{n} \mathrm{Max}\left[0, TRG_j^* - TR_j\right]}{\mathrm{E}[TRG]},$$

where j denotes the simulated outcome, TR_j is the jth simulated total revenue outcome, and other variables are as previously defined.

Results of the PQH simulation are compared to results from an IC simulation using identical parameters. For each simulation, a set of 1,000 samples are drawn. For each procedure, the mean and standard deviation of the estimated premium rate and of each element of the correlation matrix are calculated across the 1,000 samples. Results are generated for sample sizes of 500, 1,000, 5,000, and 10,000; t-tests are used to test for statistically significant differences between the prescribed correlation and the average

[5] Technically, this applies only for large samples since the numerator of equation (8) approaches the population mean as n approaches infinity.



Note: Premium rates for all sample sizes are significantly different between IC and PQH procedures at $\alpha = 0.01$.

Figure 6. Simulated pure premium rates for multi-crop revenue insurance at the 75% coverage level using PQH and IC simulation procedures

correlation resulting from the two different simulation procedures. Simulated rates are also compared to evaluate the differences in premium rate resulting from the differences in simulation performance. Given the marginal distributions used in this application, it is not possible to derive a premium rate analytically. Thus, it is not possible to define categorically which simulation procedure produces results closest to the "true" rate. It is possible, however, to approximate the true rate numerically by simulating a large number of outcomes.

Whole-Farm Insurance Rating Results

Figure 6 shows simulated premium rates for the three-crop revenue derived from the PQH and IC procedures. Rates are very consistent across sample sizes for both procedures; however, rates are not necessarily consistent between procedures. The *t*-tests reveal a statistically significant difference ($p < 0.05$) between premium rates estimated from the two different procedures at all sample sizes. Specifically, as figure 6 makes clear, premium rates simulated from the PQH procedure are statistically larger than rates simulated from the IC procedure for all sample sizes. While statistically significant, these differences in rates are not necessarily economically significant. To put these statistical differences in context, using the expected prices and yields estimated here, each 0.1% change in premium rate amounts to a difference in premium paid of just under $0.35/acre at the 75% coverage level.

As noted, we cannot derive the true rate analytically. To approximate the true rate, we estimated the rates for 100,000 simulated outcomes. In this exercise, each procedure produced a rate that was virtually identical (the same to the fourth decimal place) to the rates shown in figure 6 for the 10,000 sample size (i.e., 0.0439 for the IC procedure and 0.0446 for the PQH procedure). In other words, the numerical approximation of the true rate differs significantly depending upon which simulation procedure is used. It seems reasonable to infer that the simulation resulting in the most accurate reproduction of the underlying data produces the more accurate rate.

The more accurate representation of the dependencies in the data from the PQH procedure in large samples suggests that the rates derived from this procedure are closer to the true rate than those derived from the IC simulation (though, as stated, there can be no analytical proof of this). The difference in rates estimated in the simulation of sample size 10,000 results in a total premium difference of around $0.25/acre at the 75% coverage level. On a total premium of about $15/acre based on the expected prices and yields used in this application, the difference in estimated premium is not economically significant, despite the high level of statistical significance. However, in a crop insurance context, rates must be generated for literally hundreds of empirical settings. Hence, a biased rating system is problematic in that adverse selection may be severe in certain contexts. The greater accuracy of the PQH procedure in reproducing the desired dependencies between variables, as well as the fact that it requires no subjective judgments to implement (such as the distribution of the score matrix with the IC procedure), suggests the PQH procedure is an attractive alternative to the widely used IC procedure for actuarial applications.

Summary and Conclusions

This study calculates rates for a hypothetical multi-crop revenue insurance policy using two alternative multivariate simulation techniques: the Iman and Conover (IC) (1982) procedure, which is standard in actuarial practice, and a newer procedure described by Phoon, Quek, and Huang (PQH) (2004). Results indicate significantly different rates can be obtained depending on which simulation procedure is used.

While it is not possible to calculate the true rate analytically from the price and yield distributions assumed here (i.e., lognormal prices and beta yields), it is possible to evaluate the accuracy of the simulations from each of the alternative procedures. Results of t-tests show that correlations in the PQH simulated data were not significantly different from the actual (assumed) correlation, while correlations in the IC simulated data were generally significantly different from actual values. More specifically, correlations in the IC simulated data are biased, being smaller in absolute value than the actual correlation. This bias was not corrected by increasing sample size (though the amount of the bias could be marginally reduced).

While perhaps not a major consideration, it is nonetheless worth noting that the programming of the PQH procedure is much more straightforward than for the IC procedure. As pointed out earlier, the IC procedure involves sorting data to achieve the desired correlation. This necessarily requires separating matrices into individual columns for individual sorting and then re-merging to produce a final simulated data set. This is a rather cumbersome process from a programming standpoint (and one that becomes more cumbersome as the number of simulated variables increases) compared to the PQH procedure, which for the most part involves intact matrix manipulation.[6]

[6] Sample code for implementing the PQH procedure is available from the authors upon request.

Results of this study suggest that more accurate rates for multi-crop insurance products (an area of growing interest) could be obtained by simulation with the PQH procedure, though the difference in rates is not economically significant in the case we examined. More generally, these results provide an effective and easily implemented alternative to existing multivariate simulation techniques that can be applied in a wide variety of analyses. For example, the 2008 Farm Bill includes both a revenue counter-cyclical program as well as a whole-farm disaster compensation program. Analysis of either of these policy instruments requires the type of multivariate simulation from mixed marginal distributions as described and evaluated here.

[*Received April 2008; final revision received February 2009.*]

References

Anderson, J. D., K. H. Coble, and J. C. Miller. "Hedging a Government Entitlement: The Case of Counter-cyclical Payments." *J. Agr. and Appl. Econ.* 39(2007):507–522.

Anderson, J. D., and K. A. Zeuli. "The Revenue Risk of Value-Based Pricing for Fed Cattle: A Simulation of Grid vs. Average Pricing." *Internat. Food and Agribus. Mgmt. Rev.* 4(2001):275–286.

Babcock, B. A., and D. A. Hennessy. "Input Demand Under Yield and Revenue Insurance." *Amer. J. Agr. Econ.* 78(1996):416–427.

Bailey, D. V., and J. W. Richardson. "Analysis of Selected Marketing Strategies: A Whole-Farm Simulation Approach." *Amer. J. Agr. Econ.* 67(1985):813–820.

Biller, B., and B. L. Nelson. "Modeling and Generating Multivariate Time-Series Input Processes Using a Vector Autoregressive Technique." *ACM Transactions on Modeling and Computer Simulation* 13(2003):211–237.

Clemens, R. T., and T. Reilly. "Correlations and Copulas for Decision and Risk Analysis." *Management Sci.* 45(1999):208–224.

Clements, A. M., H. P. Mapp, and V. R. Eidman. "A Procedure for Correlating Events in Farm Simulation Models." Exp. Sta. Bull. No. T-131, Oklahoma State University, Stillwater, 1971.

Ferson, S., R. B. Nelson, J. Hajagos, D. J. Berleant, J. Zhang, W. T. Tucker, L. R. Ginzburg, and W. L. Oberkampf. *Dependence in Probabilistic Modeling, Dempster-Shafer Theory, and Probability Bounds Analysis.* Pub. No. SAND2004-3072, Sandia National Laboratories, 2004.

Frees, E. W., and E. A. Valdez. "Understanding Relationships Using Copulas." *N. Amer. Actuarial J.* 2(1998):1–25.

Hart, C. E., D. J. Hayes, and B. A. Babcock. "Insuring Eggs in Baskets: Should the Government Insure Individual Risks?" *Can. J. Agr. Econ.* 54(2006):121–137.

Hennessy, D. A., B. A. Babcock, and D. J. Hayes. "The Budgetary and Producer Welfare Effects of Revenue Assurance." *Amer. J. Agr. Econ.* 79(1997):1024–1034.

Hua, Y. "Generalized Karhunen-Loeve Transform." *IEEE Signal Processing Letters* 5(1998):141–142.

Iman, R. L., and W. J. Conover. "A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables." *Communications in Statistics* B11(1982):311–334.

Johnson, M. E., and A. Tenenbein. "A Bivariate Distribution Family with Specified Marginals." *J. Amer. Statis. Assoc.* 76(1981):198–201.

Krzanowski, W. J. *Principles of Multivariate Analysis: A User's Perspective.* New York: Oxford University Press, 1988.

Mildenhall, S. J. "Correlation and Aggregate Loss Distributions with an Emphasis on the Iman-Conover Method." Report of the Research Working Party on Correlations and Dependencies Among All Risk Sources, Casualty Actuarial Society, 2005. Online. Available at http://www.casact.org/pubs/forum/06wforum/06w107.pdf.

Naylor, T. H., J. L. Balintfy, D. S. Burdick, and K. Chu. *Computer Simulation Techniques.* New York: John Wiley and Sons, Inc., 1966.

Paulson, N. D., and B. A. Babcock. "The Effects of Uncertainty and Contract Structure in Specialty Grain Markets." Paper presented at annual meetings of the American Agricultural Economics Association, Portland, OR, 29 July–1 August 2007.

———. "Get a GRIP: Should Area Revenue Coverage Be Offered Through the Farm Bill or as a Crop Insurance Program?" *J. Agr. and Resour. Econ.* 33,2(August 2008):137–153.

Phoon, K., H. Huang, and S. T. Quek. "Simulation of Second-Order Processes Using Karhunen-Loeve Expansion." *Computers and Structures* 8(2002):1049–1060.

Phoon, K., S. T. Quek, and H. Huang. "Simulation of Non-Gaussian Processes Using Fractile Correlation." *Probabilistic Engineering Mechanics* 19(2004):287–292.

Tattari, S., I. Barlund, S. Rekolainen, M. Posch, K. Siimes, H. R. Tuhkanen, and M. Yli-Halla. "Modeling Sediment Yield and Phosphorus Transport in Finnish Clayey Soils." *Transactions ASAE* 44(2001): 297–307

Van Trees, H. L. *Detection, Estimation, and Modulation Theory: Part I.* New York: John Wiley and Sons, Inc., 1968.

Vorechovsky, M., and R. Chudoba. "Stochastic Modeling of Multi-Filament Yarns: II. Random Properties over Length and Size Effect." *Internat. J. Solids and Structures* 43(2006):435–458.

Webster, M., C. Forest, J. Reilly, M. Babiker, D. Kicklighter, M. Mayer, R. Prinn, M. Sarofim, A. Sokolov, P. Stone, and C. Wang. "Uncertainty Analysis of Climate Change and Policy Response." *Climatic Change* 61(2003):295–320.

Zhang, R., J. E. Houston, D. Vedenov, and B. J. Barnett. "Hedging Downside Risk to Farm Income with Futures and Options: Effects of Government Payment Programs and Federal Crop Insurance Plans." Poster presented at annual meetings of the American Agricultural Economics Association, Portland, OR, 29 July–1 August 2007.

Zhu, Y., S. K. Ghosh, and B. K. Goodwin. "Modeling Dependence in the Design of Whole-Farm Insurance Contract: A Copula-Based Model Approach." Paper presented at annual meetings of the American Agricultural Economics Association, Orlando, FL, 27–29 July 2008.