Effects of Simulation Volume on Risk Metrics for Dynamo DFA Model

By William C. Scheel, Ph.D., DFA Technologies, LLC and
Gerald Kirschner, FCAS, MAAA, Deloitte Consulting LLP

Abstract: Of necessity, users of complex simulation models are faced with the question of "how many simulations should be run?" On one hand, the pragmatic consideration of shortening computer runtime with fewer simulation trials can preclude simulating enough of them to achieve precision. On the other hand, simulating many hundreds of thousands or millions of simulation trials can result in unacceptably long run times and/or require undesirable computer hardware expenditures to bring run times down to acceptable levels. Financial projection models for insurers, such as Dynamo, often have complex cellular logic and many random variables. Users of insurance company financial models often want to further complicate matters by considering correlations between different subsets of the model's random variables. Unfortunately, the runtime / accuracy tradeoff becomes even larger when considering correlations between variables.

Dynamo version 5, written for use in high performance computing (HPC)[1], as used for this paper, has in excess of 760 random variables, many of which are correlated. We have used this model to produce probability distribution and risk metrics such as Value at Risk (VaR), Tail Value at Risk (TVaR) and Expected Policyholder Deficit (EPD) for a variety of modeled variables. In order to construct many of the variables of interest, models such as Dynamo have cash flow overlays that enable the projection of financial statement accounting structures for the insurance entity being modeled. The logic of these types of models is enormously complex and even a single simulation is time consuming.

This paper begins by examining the effect that varying the number of simulations has on aggregate distributions of a series of seven right-tailed, correlated lognormal distributions. Not surprisingly, the values were found to be more dispersed for smaller sample sizes. What was surprising was finding that the values were also lower when using smaller sample sizes. Based on the simulations we performed, we conclude that a minimum of 100,000 trials is needed to produce stable aggregate results with sufficient observations in the extreme tails of the underlying distributions.

Similar conclusions are drawn for the modeled variables simulated with Dynamo 5. Sample sizes under 100,000 produce potentially misleading results for risk metrics associated with projected policyholders surplus. Based on the quantitative values produced by the HPC version of Dynamo 5 used in this article, we conclude that sample sizes in excess of 500,000 are warranted. The reason for the higher number of simulations in Dynamo 5 as compared to the seven variable example is the greater complexity of Dynamo, specifically the much larger number of random variables and the complexity of the correlated interactions between variables. As support for this, we observe that simulated metrics for Policyholders Surplus decreased by 2% to 3% when simulations were increased from 100,000 to 700,000. They decreased by 3% to 6% when simulations were increased from 10,000 to 700,000.

.

[1] High performance computing involving the parallelization of the Dynamo model so it would run in computer clusters offers a potential solution to the trade-off between precision and runtime. A small HPC cluster can reduce runtime by 1/3 for 100,000 trials, from about 1.5 hours to 33 minutes.

CONTENTS

INTRODUCTION

Dynamo is an open access dynamic financial analysis (DFA)[2] model built in Microsoft Excel. It is available on the Casualty Actuarial (CAS) web site.[3] The call paper program (herein after referred to as Call) encouraged model redesign but probably did not anticipate the reformation of the model to run in a high performance computing (HPC) environment.

> Participants are encouraged to develop any needed enhancements, such as add-on programs/macros to Dynamo 4.1. This call for papers is intended to foster the use of Dynamo 4.1 and to generate publicly available improvements to the model.

HPC Dynamo[4] still retains standalone properties, but it was redesigned to run with high-volume simulations in the hundreds of thousands[5] instead of a few thousand[6] simulations. The model was parallelized and runs in a services oriented architecture (SOA) wherein server computers simultaneously use multiple instances of Excel and the Dynamo model. Empirical probability distributions are built from the simulations being done in parallel across many computers. A pool of such computers is called an HPC cluster. Further, any single computer in the cluster may have many processing units or cores. So, where a cluster has 100 computers, each with four cores, it would be possible to run 400 instances of Excel in parallel.

In this fashion it is possible to run simulations with as many as 750,000 trials on a moderate-sized cluster in about 30 minutes.[7] The technology affords an interesting opportunity to examine the effects of sample size on various risk metrics being calculated in the Dynamo model.

To facilitate the evaluation of what we considered to be interesting and relevant metrics, we extended HPC Dynamo to calculate value at risk (VaR), tail value at risk (TVaR) and expected

---

[2] DFA involves simulation to obtain an empirical probability distribution for accounting metrics. As such, an accounting convention such as statutory or GAAP is required. Cash flows are generated for many dependent random variables, and these cash flows are evaluated within the accounting framework. Realizations of financial values from balance sheets or income statements obtained during the simulations are used to construct probability distributions for the financial values.

[3] Dynamo model, version 4.1 and documentation can be obtained at: http://www.casact.org/research/index.cfm?fa=padfam.

[4] HPC Dynamo version 5.x can be obtained at: http://www.casact.org/research/index.cfm?fa=dynamo. Please note there is a vast amount of both written material and video clips available on-line for version 5.x. This help documentation is directly accessible to users of Dynamo 5x from some new dialogs.

[5] HPC Dynamo must be run in Excel 2010 (Microsoft Office version 14).

[6] Dynamo 4, the model from which HPC Dynamo 5 was created, can, in theory, also generate several hundred thousand scenarios, but this may not be practical when it takes approximately three hours to run 5,000 simulations.

[7] The work done for this paper was generated on two clusters. One had about 240 cores and a smaller one had about 24 cores. There was a mixture of computer types involving both 64- and 32-bit computers. Two operating systems were used: Windows Server 2008 R2 and Windows 7. In our experience, neither of these platforms would be considered large HPC clusters. Each computer supporting the cluster had eight cores. As noted, HPC Dynamo also can be run on single instance of Excel 2010 without HPC functionality.

policyholder deficit (EPD) values for the DFA variables. We have extended HPC Dynamo in this manner in response to the direction that global insurance company solvency and financial regulations (i.e., Solvency II, IFRS) appear to be headed. Other standard statistics also are computed.

SECTION 1: COMPARISON OF SOLVENCY II AND OTHER RISK METRICS USING MULTIVARIATE SIMULATION OF LOGNORMAL DISTRIBUTIONS

Introduction

In this section we illustrate sampling phenomena for lognormal distributions that are correlated. This section is a simplification of the Dynamo 5 example that will be the focus of the next section. In this section we focus on a series of seven lognormally distributed variables. In the next section, we will work with the Dynamo model and its 760 random variables, of which only some are lognormally distributed.

We also use this occasion to review several risk metric constructs, including those being used for Solvency II (S II).

Solvency II Risk Aggregation

The Solvency II regime's standard formula is predicated on risk aggregation of different capital charges through an approach similar to classical portfolio theory, i.e., there is an assumed reduction in volatility arising from risk diversification. The derivation of the Basic Solvency Capital Requirement (BSCR)[8] uses a subjective correlation matrix similar to the one shown in Table 1 to capture this reduction in volatility, and it is calculated using (1).

[8] European Commission Internal Market and Services DG, Financial Institutions, Insurance and Pensions, "QIS4 Technical Specifications (MARKT/2505/08), Annex to Call for Advice from CEIOPS in QIS4(MARKT/2504/08)," pp. 286. This document is hereinafter referred to as QIS4. The CEIOPS Solvency II Directive is the globally operative document. It can be found, with highlights for "easy reading" in English, at http://www.solvency-ii-association.com/Solvency_ii_Directive_Text.html. The Committee of European Insurance and Occupational Pensions Supervisors (CEIOPS) web site has the latest rendering of the Solvency II Framework Directive. http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P6-TA-2009-0251+0+DOC+XML+V0//EN.

Table 1: QIS5 Correlation Matrix for BSCR[9]

$CorrSCR_{r,c}$	SCR_{market}	$SCR_{default}$	SCR_{life}	SCR_{health}	$SCR_{non\text{-}liife}$
SCR_{market}	1				
$SCR_{default}$	0.25	1			
SCR_{life}	0.25	0.25	1		
SCR_{health}	0.25	0.25	0.25	1	
$SCR_{non\text{-}life}$	0.25	0.5	0	0.25	1

$$BSCR = \sqrt{\sum_{rxc} CorrSCR_{r,c} \cdot SCR_r \cdot SCR_c} \tag{1}$$

where

$CorrSCR_{r,c}$ = the cells of the correlation matrix mandated by Solvency II[10]

SCR_r, SCR_c = Capital charges for the individual SCR risks according to the rows and columns of the correlation matrix CorrSCR

Portfolio Risk Aggregation

The Solvency II expression for BSCR is identical to the standard deviation of a portfolio of *equally weighted risks when the marginal standard deviations are the same as the capital charges*. This statement follows from the definition of the variance of a portfolio shown in (2).

$$\sigma_p^{\,2} = \sum_i \sum_j w_i w_j \sigma_i \sigma_j \rho_{ij} \tag{2}$$

Where

σ_i = standard deviation of the *i*-th risk component.

[9] QIS5 Correlation Matrix for BSCR, p. 96.
https://eiopa.europa.eu/fileadmin/tx_dam/files/consultations/QIS/QIS5/QIS5-technical_specifications_20100706.pdf

[10] The SCR_i shown in Table 1 are defined across broad risk categories identified within the S II. Each risk component is functionally related to a VaR metric. For example, $SCR_{non\text{-}life}$ is the non-life (i.e., property/casualty) component. It is a function of geographic and other risk attributes and is intended to calculate parameters of a lognormal distribution and VaR associated with that distribution. Other SCR components attempt to identify market (SCR_{market}), life (SCR_{life}), health (SCR_{health}) and operational risks ($SCR_{default}$) confronting insurers.

When the weights, w_i equal 1, equations (1) and (2) are identical. And, $\sigma_i = SCR_i$, when the *i*-th capital charge in SCR is the standard deviation of some random variable.

The limiting properties of large numbers of component risks may be thought to have the convergence properties of the Central Limit Theorem. Applying this assumption, a VaR measure for a portfolio of risks with mean, μ_p, and portfolio standard deviation, σ_p can then defined by (3).

$$VaR_\alpha = \mu_p + \Theta_\alpha * \sigma_p \tag{3}$$

Where,

Θ_α = standard normal value at a cumulative probability of α.

The assumption of a Gaussian process in (3) has rankled many observers. N.N. Taleb, for example, sees "Black Swans" showing up as extreme realizations in risk processes that are distinctly non-normal.[11] The chance-constrained metric in (3) for a portfolio of risks may understate the chance-constrained point derived without Gaussian assumptions. We believe Taleb would characterize marginal distributions for many insurance-related loss processes to be Black Swan candidates.

The aggregation method for BSCR indicated in is likely predicated on a methodology in which each component SCR can be thought of as a portfolio component standard deviation. This same approach is widely used among all of the SCR_x risk components throughout most S II capital charges.

A solvency capital charge can be a chance-constrained *portfolio* value such as a multiple of standard deviations as shown in (4).

$$SCR' = \Theta_\alpha * \sigma_p \tag{4}$$

But, the portfolio mean μ_p is defined by (5).

$$\mu_p = \sum_i w_i \mu_i \tag{5}$$

So, the *portfolio* capital charge, SCR', is given by (6) after substitution of and into and noting that the weights in equal 1.

$$SCR' = VaR_\alpha - \mu_p \tag{6}$$

And, as noted at the beginning of this section, the Solvency II expression for BSCR is the standard deviation of a portfolio of *equally weighted risks when the marginal standard deviations are the same*

[11] Black Swan theory explains high-impact, hard-to-predict, and rare events. They arise from non-normal, non-Gaussian expectations. N.N. Taleb, *The Black Swan: The Impact of the Highly Improbable*, ISBN-13: 9781400063512, 2007, 400 pp. Taleb is not without his critics. A summary of the more cogent ones is found at http://en.wikipedia.org/wiki/Taleb_distribution#Criticism_of_trading_strategies

as the capital charges.

Solvency II and Portfolio Aggregation

If we assume that SCR capital charges will, in practice, be larger than the marginal standard deviations of the SCR components, it means that the SCR, in equation (1) will be larger than σ_i in equation (2). This, in turn, would mean that the Solvency II standard formula approach to deriving a capital requirement would be inflated relative to the portfolio approach for defining a capital charge. The capital charges used in S II aggregation are typically more complex measurements than are illustrated in (7). Here the capital charge is a standard normal multiple, Θ_α, of the distribution's standard deviation.

$$SCR_i = (\mu_i + \Theta_\alpha \sigma_i) - \mu_i = \Theta_\alpha \sigma_i \tag{7}$$

We will examine this in the context of a portfolio of lognormal random variables with known parameters, $\{\mu_i, \sigma_i\}$. The values of these parameters appear in Table 2. Please note that the term, "Var *x*" means a lognormally distributed *variable* and does not mean value at risk or variance. The correlation matrix used both for the S II and portfolio approaches to developing capital charges is shown in Table 3.

Table 2: Parameters for Lognormal Distributions

Name	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6	Var 7
Risk Model	Log Normal	Log Normal	Log Normal	Log Normal	Log Normal	Log Normal	Log Normal
Mean	10000	50000	90000	130000	170000	210000	250000
Standard Deviation	5000	6000	7000	8000	9000	10000	11000

Table 3: Correlation Matrix for Lognormal Distributions

	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6	Var 7
Var 1	1.0000	0.1315	-0.0986	0.1972	0.3945	0.1972	-0.0723
Var 2	0.1315	1.0000	-0.1972	0.1315	0.3944	0.3945	0.0328
Var 3	-0.0986	-0.1972	1.0000	0.3287	0.1315	0.3945	0.1972
Var 4	0.1972	0.1315	0.3287	1.0000	0.0000	-0.0657	0.1315
Var 5	0.3945	0.3945	0.1315	0.0000	1.0000	0.0328	0.0131
Var 6	0.1972	0.3945	0.3945	-0.0657	0.0328	1.0000	0.5260
Var 7	-0.0723	0.0328	0.1972	0.1315	0.0132	0.5260	1.0000

In the next section we describe aggregation based on a third approach to a capital charge—the difference between VaR and the mean of the multivariate *aggregate* loss distribution for the lognormal marginal variates described in Table 2 and rank correlated by Table 3.

However, at this point it is instructive to present all three values for these aggregation approaches using this simplified seven variable model. The capital charges appear in Table 4. These capital charges reflect the range of outcomes achieved after 750,000 simulations and taking the .995 percentile of the resulting aggregate distribution.

Table 4: Capital Charges Under Solvency II, Portfolio, and Aggregate Loss Aggregation Methods

Method of Aggregation	Capital Charge
Aggregate Loss	81,268
Solvency II BSCR	85,654
Portfolio	77,597

The capital charge using the S II methodology exceeds the portfolio approach, and by a sizable margin. Of course, in actual application, this margin will depend on the underlying loss distributions and the correlation matrix.

Aggregate Loss Distribution Using the Iman-Conover Method of Inducing Correlations

The multivariate simulation methods we deploy use the Iman-Conover approach for inducing correlation into independent distributions.[12] The first step is to simulate values from each of the seven lognormal variables independently of one another to produce a table of *n* rows by seven columns, where each row represents one scenario in the overall simulation exercise. The second step is to reorder the rows by sorting them from low to high using the values in the first column as the sort field. The matrix being illustrated in Table 5 show the results of 10 scenarios after reordering them based on the simulated values for Var 1.[13] The matrix is then shuffled so that the rearrangement has the Spearman rank correlations shown in Table 3. The result of this Iman-Conover induction of correlation into independent distributions appears in Table 6. This approach is particularly useful when correlation is subjective, and the loss processes are developed and parameterized by independent groups of actuaries. It is especially useful for multivariate simulation. Each row of Table 6 contains an *n*-tuple from a multivariate distribution with Spearman correlations shown in Table 3. The rows are realizations for the seven variables that may be used for different trials in a simulation.

[12] The Iman-Conover method is described in the report of the Casualty Actuarial Society's Working Party on Correlations and Dependencies Among All Risk Sources found at http://www.casact.org/pubs/forum/06wforum/06w107.pdf. Also see Kirschner, Gerald S., Colin Kerley, and Belinda Isaacs, "Two Approaches to Calculating Correlated Reserve Indications Across Multiple Lines of Business," *Variance* 2:1, 2008, pp. 15-38.

[13] The table shows the first ten rows of 25,000 used with the Iman-Conover method.

Table 5: Lognormal Variates Before Induction of Rank Correlation

Var 1	Var 2	Var 3	Var 4	Var 5	Var 6	Var 7
1,261	56,244	86,464	139,679	184,313	207,222	220,879
1,350	41,798	82,325	125,670	177,085	201,510	260,059
1,404	53,743	91,548	119,955	167,478	233,410	246,224
1,525	44,553	80,663	142,115	158,827	208,627	235,541
1,620	47,549	77,273	127,529	157,531	197,469	247,044
1,671	47,671	82,639	125,521	183,718	208,845	237,500
1,721	54,840	86,908	132,476	173,432	224,805	265,150
1,734	61,729	83,191	122,804	176,781	201,987	249,738
1,743	55,287	91,678	130,586	169,779	207,816	251,302
1,808	52,759	97,670	133,850	181,218	206,202	266,005

Table 6: Lognormal Variates After Induction of Rank Correlation

Var 1	Var 2	Var 3	Var 4	Var 5	Var 6	Var 7	Aggregate
1,261	48,005	97,603	128,608	151,988	212,276	261,124	900,866
1,350	53,071	103,625	121,782	160,720	224,589	259,898	925,035
1,404	40,877	87,598	125,819	153,027	189,504	249,399	847,627
1,525	51,668	96,151	135,583	169,766	191,346	238,395	884,434
1,620	50,617	91,693	127,287	161,166	202,703	255,470	890,555
1,671	50,021	90,844	123,747	149,731	208,077	244,267	868,359
1,721	58,834	83,220	122,938	147,002	219,338	261,585	894,638
1,734	39,731	102,377	129,503	153,077	200,082	244,553	871,057
1,743	38,745	85,285	121,838	143,780	192,489	243,678	827,558
1,808	44,030	89,124	130,296	153,196	190,989	240,596	850,038

Each of the variables in a row of Table 6 is added to produce an observation in the aggregate loss distribution as shown in the final column of each row. This is a multivariate empirical distribution, but there is no available multivariate probability distribution that defines it. That is, the aggregate loss distribution is not constructed with a variance/covariance matrix, and it does not use Pearsonian correlation. Nevertheless, it is an aggregate distribution based on independently derived probability distributions that are observed to have pairwise Spearman rank correlations. It is multivariate in that sense.

We note that this empirical probability distribution is not directly used in Dynamo. Instead, the

multivariate Iman-Conover trials are available for use in Dynamo. The multivariate variables may be used in dependent cells so that a simulation in Dynamo is using random variates that are correlated. It is possible to have many clusters of such correlated variables where each is used for different cell dependencies.[14] For example, new business growth among lines of business could be a function of random variables within a pod or cluster that are correlated.[15] DFA variables dependent on them will be generated with the underlying correlation structure of the pod or cluster.

[14] The technique is very useful when the underlying correlation structure of a cluster of variables is subjective. It is important to remember, however, that subjective correlations must be reckoned as *rank* correlations.
[15] The term "pod" and "cluster" are used interchangeably in this paper. Each refers to a collection of variables with a correlation structure and multivariate properties defined within the Iman-Conover methodology.

Sensitivity to Sample Size

We begin our discussion of simulation volume effects, or sample size effects, with the example in Table 2. Except for small sample sizes, both the S II and portfolio methodologies should be relatively insensitive to sampling error because they depend on first and second moments of distributions and sampling error will rapidly diminish with simulation volume. But, because the underlying distributions are lognormal, we would expect sampling error to have a more profound impact on the variables with the highest second moments, i.e., Var 6 and Var 7. This expectation is confirmed in Table 7.

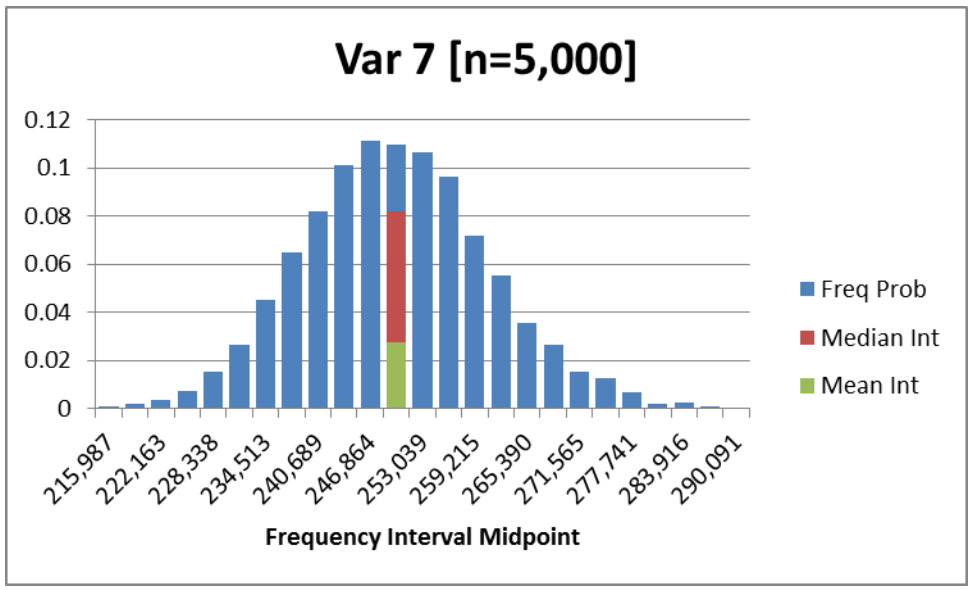
Table 7: Capital Charges for Different Trial Volumes

				Capital Charge Methods			
Name	Mean	Standard Deviation	VaR	Aggregate Loss Method	Solvency II BSCR Method	Port-folio Method	Trials
Var 1	9,843	4,918	27,280	17,436			1,000
Var 1	9,995	4,993	31,076	21,081			5,000
Var 1	9,928	5,042	31,394	21,467			10,000
Var 1	9,952	5,013	30,702	20,750			25,000
Var 1	9,987	5,051	30,802	20,815			50,000
Var 1	9,990	5,045	30,663	20,673			100,000
Var 1	10,005	5,017	30,261	20,256			250,000
Var 1	10,005	5,006	30,207	20,202			500,000
Var 1	10,004	5,006	30,233	20,228			750,000
⋮	⋮	⋮	⋮	⋮			
Var 7	250,062	11,305	278,418	28,356			1,000
Var 7	250,128	11,065	279,004	28,877			5,000
Var 7	250,023	10,926	279,204	29,181			10,000
Var 7	250,028	10,937	279,340	29,312			25,000
Var 7	250,000	10,977	279,778	29,778			50,000
Var 7	249,976	10,980	279,963	29,987			100,000
Var 7	249,985	10,997	279,866	29,881			250,000
Var 7	249,986	11,002	279,776	29,790			500,000
Var 7	249,992	11,005	279,722	29,730			750,000
Aggregate	909,344	30,092	987,239	77,895	82,780	78,377	1,000
Aggregate	910,016	29,796	990,112	80,096	85,222	77,357	5,000
Aggregate	909,866	30,135	992,724	82,858	85,856	77,526	10,000
Aggregate	909,972	30,079	990,519	80,547	85,238	77,419	25,000
Aggregate	910,065	30,122	991,457	81,392	86,266	77,683	50,000
Aggregate	909,925	30,063	991,227	81,301	86,184	77,604	100,000
Aggregate	909,979	30,008	991,009	81,030	85,773	77,549	250,000
Aggregate	910,005	30,062	991,133	81,128	85,584	77,588	500,000
Aggregate	910,004	30,045	991,272	81,268	85,654	77,597	750,000

The aggregate distribution capital charge is also affected by sample size as can be seen at the bottom box of Table 7. Visual comparison of the two segments of this box show aggregate capital charges (left column of box) to be both lower and more dispersed for smaller sample sizes. (For example, the average of the Variables and Aggregate column that aggregates for between 1,000 and 50,000 trials is 80,558 as compared to an average of 81,182 for the simulations' runs that used between 100,000 and 750,000 trials.)

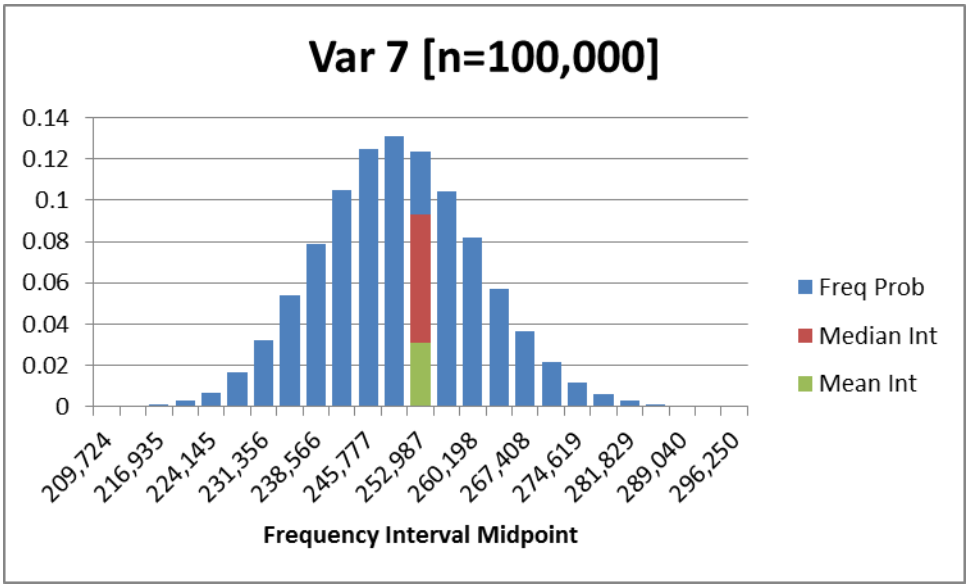
Higher sample sizes for the lognormal distributions result in more observations in the extreme tails. This effect is clearly evident by examining the tail areas of Table 7 where more extreme observations occur with the 750,000 sample size relative to a sample size of only 5,000. The increase in sample size from 100,000 to 750,000 (charts B and C) illustrates how significant shifts in distribution statistics can unfold even when increasing from a comparatively high sample size of 100,000 to extreme sampling sizes such as 750,000. This impact is documented in Table 7 for Var 7. The mean increases from 249,976 to 249,992. However, VaR declines from 279,963 to 279,722.

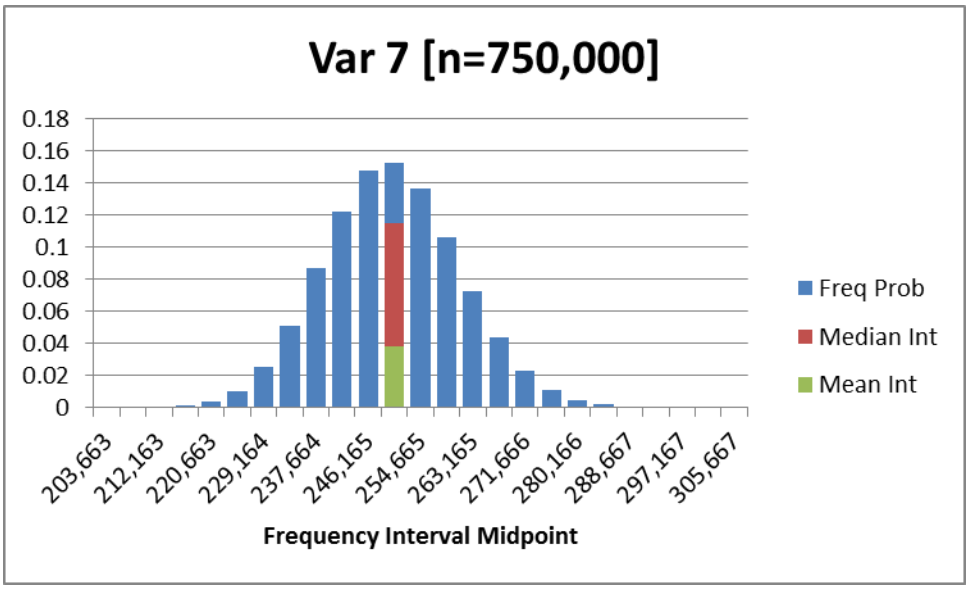
Figure 1A: High-Variance Lognormal Distribution for Different Sample Sizes[16]



[16] The graphics used in this paper are produced by Dynamo 5 for any simulated variable. The term "Int" in the legend refers to interval. The mean and median intervals are overlaid in their frequency intervals as visualization of where these central tendency measures fall. This information is not particularly useful for this paper, but can be a useful for heavily skewed distributions.

Figures 1B and 1C: High-Variance Lognormal Distribution for Different Sample Sizes[17]

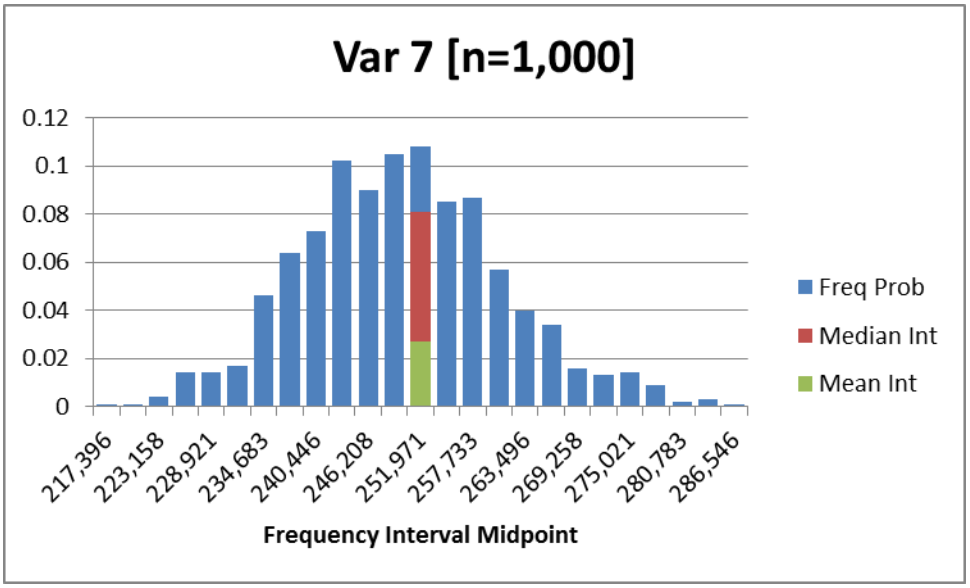




Because prior versions of Dynamo were formulated for sample sizes of only 1,000, the frequency distribution graph for this 1,000 sample size appears in Figure 2. The effects of low sample size are clearly evident both in fewer extreme values and discontinuities in shape of the frequency distribution as compared to the higher sample volumes shown in Figures 1A, 1B, and 1C.

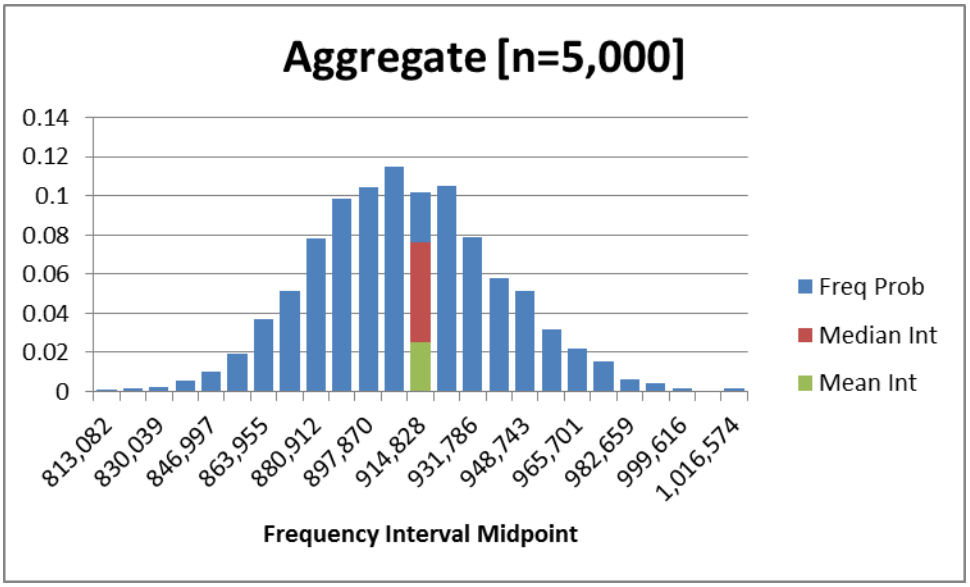
[17] The graphics used in this paper are produced by Dynamo 5 for any simulated variable. The term "Int" in the legend refers to interval. The mean and median intervals are overlaid in their frequency intervals as visualization of where these central tendency measures fall. This information is not particularly useful for this paper, but can be a useful for heavily skewed distributions.

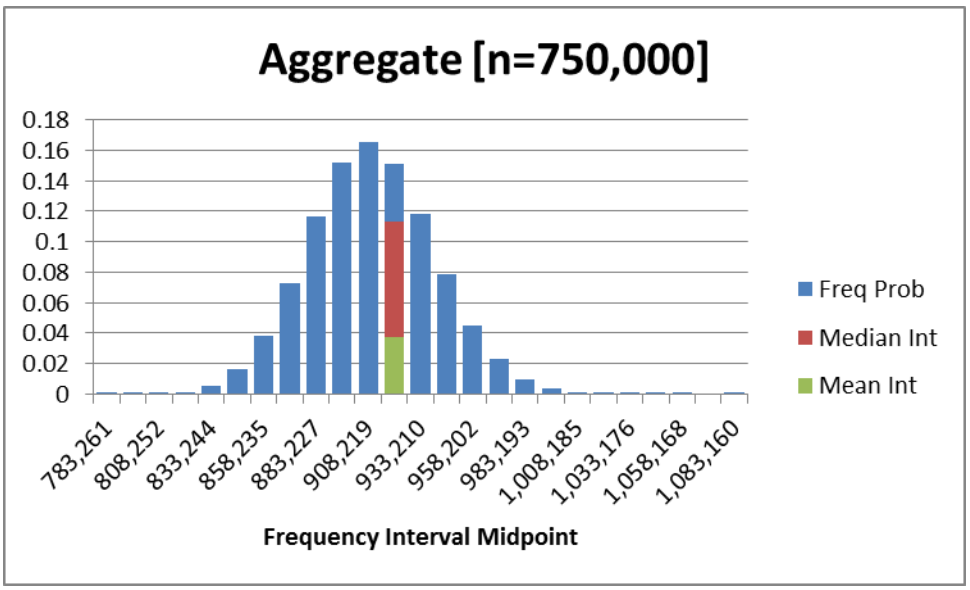
Figure 2: High-Variance Lognormal Original Dynamo 1K Sample Size



The impact of sample size also occurs for the aggregate loss distribution. Here, too, more extreme values emerge with the 750,000 sample size. The .995 VaR for the aggregate loss distribution with a sample size of 5,000 is 990,112 as compared to 991,272 for the 750,000 sample trial. But, this leads into the question of how many simulations is enough? A comparison of Figures 3A and 3B illustrates visually the effects of the central limit theorem. Highly skewed lognormal distributions when aggregated will, with sufficient sample sizes, produce a normally distributed sum. As we move from a clearly insufficient sample size of 1,000 shown in Figure 2 to 750,000 shown in Figure 3B we find an unfolding of increasing precision throughout the probability distribution. Sample size matters. The added precision obtained by using Excel in an HPC cluster is valuable, but at the same time there is an asymptotic collapse of sampling error. At some point, enough is enough. If insurance company modeling were as simple as the seven variable example being used in this section, one might be tempted to conclude that the time and effort and expense required to increase the number of trials from 1,000 to 750,000 does not justify the 0.1% increase in the .995 VaR. However, insurance company modeling is not this simple. We now turn to the analysis of sample size on Dynamo DFA variables to examine a more complex modeling situation.

Figures 3A and 3B: Effect of Sampling Size on Aggregate Loss Distribution





SECTION 2: EFFECT OF SAMPLE SIZE ON DYNAMO DFA VARIABLES DISTRIBUTIONS

Introduction

Because Excel is used for Dynamo, it can be relatively easy to model complex interactions for a large number of different DFA variables. Business operations can be modeled with complex cash flow and accounting dependencies using many random variables. Given a set of random variates (Dynamo has in excess of 760 inverse probability functions), a single calculation of the Dynamo workbook produces an empirical realization for the DFA variables being monitored. The parallelization of this process results in these realizations being calculated simultaneously in a computer cluster. Hence, HPC Dynamo can produce probability distributions with 500,000 or more observations in a short time relative to what time would be required were these observations to be done serially in a single instance of Excel. We have seen in the previous section the effects of sample size in the context of a portfolio of lognormal variables, and we now turn to similar experiments for DFA variables.

High-Volume Sampling Illuminates Extremities in Both Tails of a Distribution

Often we are more concerned about the extreme tail that represents adverse experience. High-volume observations enabled by parallelization of the simulation produces enhanced precision *throughout* the probability distribution. We have more observations at both extremities and, of course, a bevy of additional results that are largely unnecessary in the interior of the distribution. At some point, sampling error affecting moments of the distribution decays to a materially insignificant amount. More simulations do not necessarily produce a better answer. Error in estimating extreme percentiles or even moments required for solvency measurement is materially changed at simulation volumes that might be considered exceedingly large if attempted in a stand-alone computing environment.[18]

Consider the 0.995 value at risk (VaR) column in Table 8. This table contains various statistics and risk metrics for the fifth year projection of policyholders' surplus. This variable is the result of a complex set of cell dependencies in Dynamo. All of the DFA variables that can be assembled using

[18] Recall that the original Dynamo only simulated 1,000 observations. And, the results reported by Burkett et al. 2010 were based on 5,000 simulations using Dynamo version 4.1. Version 4.1 required about 2.0078 sec/simulation on a fast desktop computer. It took about three hours to produce 5,000 trials. At that rate, over 16 days would be needed to create 700,000 simulations. In addition to the use of HPC, there have been substantial improvements in Dynamo VBA coding, all of which enhance performance. In a small HPC cluster running 29 simultaneous instances of HPC Dynamo (a core resource allocation one of several types for HPC jobs), three hours is reduced to 1.25 minutes for 5,000 trials. A single trial takes about .015 seconds compared to over two seconds. And, this calculation involves multivariate simulation not available in Dynamo 4.1.

Dynamo have this property. There is no closed form solution for measuring statutory or GAAP variables that are based on cash flows which, by themselves have no closed solution. Simulation is the only viable approach to deriving probability distributions on these DFA variates.

The rows of Table 8 contain results for increasing simulation volumes. Although measurements are shown for samples sizes under 10,000, these small sample sizes have 0.995 VaRs that are heavily affected by the algorithm used to extrapolate this extreme percentile. The number of observations is smaller than the precision sought for that extremity. This algorithmic effect can be seen in the bowing of the VaRs between 1,000 and 10,000 observations. Beginning at 10,000 observations, however, a secular decline in VaR values occurs with increased simulation counts. The VaR for the 10,000 trial simulation is 12,898. By the time the 700,000 trial simulation is run, the VaR has reduced to 12,130, i.e., a 6% reduction. This 6% reduction is very likely to be considered material when considering minimum capital requirements. Similarly, one observes a 1% reduction in the VaR when moving from 500,000 to 700,000—this change may, too, be considered material.

Statistics relating to central tendency, such as the mean and median, also change, and change materially when moving from 100,000 to 700,000 trials. Both the mean and median are reduced by 2.3%.

The effect of moving from 10,000 to 700,000 trials is large. And, it is larger for extreme percentiles...profoundly so. VaR is reduced by about 6%. The mean is reduced by about 2%. The benefit of increased trial counts is higher at distribution tails than for central tendency.

Table 8: Effects of Sample Size on Policyholders Surplus[19]

Observations	Mean	Standard Deviation	Coef of Variation	Minimum	Maximum	.010 Percentile	Median	.990 Percentile	EPD[20]	TVaR[21]	VaR[22]
1,000	21,762	3,277	0.151	11,609	30,682	13,550	21,856	28,604	9,290	3,401	12,188
5,000	21,678	3,238	0.149	9,311	33,723	13,824	21,719	28,827	9,391	3,415	12,891
10,000	21,607	3,222	0.149	6,975	33,723	13,769	21,697	28,667	9,203	3,399	12,898
20,000	21,547	3,231	0.150	6,975	33,723	13,644	21,643	28,511	9,171	3,382	12,749
30,000	21,519	3,237	0.150	-1,368	33,723	13,592	21,599	28,597	9,224	3,379	12,653
40,000	21,496	3,242	0.151	-1,368	33,723	13,522	21,572	28,619	9,201	3,373	12,515
50,000	21,478	3,248	0.151	-1,368	33,918	13,531	21,547	28,628	9,204	3,368	12,545
60,000	21,463	3,252	0.152	-1,621	33,918	13,507	21,533	28,634	9,194	3,365	12,572
70,000	21,445	3,259	0.152	-1,621	33,918	13,471	21,521	28,662	9,180	3,359	12,545
80,000	21,437	3,262	0.152	-1,621	33,918	13,448	21,520	28,651	9,149	3,356	12,503
90,000	21,429	3,262	0.152	-1,621	33,918	13,427	21,513	28,645	9,142	3,354	12,515
100,000	21,419	3,263	0.152	-1,621	33,918	13,409	21,504	28,632	9,131	3,351	12,494
200,000	21,340	3,278	0.154	-15,412	34,368	13,318	21,432	28,584	9,071	3,331	12,405
250,000	21,312	3,283	0.154	-15,412	34,447	13,264	21,399	28,559	9,063	3,324	12,359
300,000	21,287	3,285	0.154	-15,412	34,447	13,260	21,375	28,537	9,049	3,318	12,347
400,000	21,242	3,289	0.155	-15,412	34,643	13,215	21,332	28,512	9,026	3,308	12,317
500,000	21,201	3,295	0.155	-15,412	34,643	13,150	21,291	28,480	9,003	3,298	12,257
600,000	21,161	3,301	0.156	-31,483	34,643	13,097	21,251	28,456	8,984	3,289	12,195
700,000	21,116	3,308	0.157	-31,483	34,643	13,043	21,207	28,427	8,960	3,278	12,130

[19] Table 8 illustrates the type of statistics available for all Dynamo-simulated variables. Statutory policyholders surplus for a company with two multi-peril and workers compensation lines of business is illustrated in the open source version of Dynamo 5. This is the source of Table 8, and the lognormal distributions used in Section 1 are among the Dynamo distributions used for variates leading to policyholders surplus. It is available on Casualty Actuarial Society web site http://www.casact.org/research/index.cfm?fa=dynamo.

[20] Expected Policyholder Deficit for area bounded between 0.5 and 0.8.

[21] Tail Value at Risk for tail above 0.995.

[22] Value at Risk for 0.995.

How Many Simulation Trials Are Enough?

The results for various statistics and risk metrics shown in Table 8 are clearly impacted by simulation volume. The importance of a high performance computational environment becomes apparent when attempting to pragmatically answer the question of how many trials is enough. In all of metrics in Table 8, we believe a minimum of 100,000 trials is essential to reduce sampling error to an acceptable minimum level. A strong argument can be made for 700,000 trials. The precision obtained when increasing trial count from 100,000 to 700,000 is a difference of 1.88%, 2.20%, and 2.90%, respectively for expected policyholder deficit, tail value at risk and value at risk. There is no risk metric that is immune from a reduction in sampling error achieved with high-volume simulations.

An HPC approach is highly desirable when simulation volumes reach a range of 100,000 and a necessity when they reach 700,000. A single machine just cannot run fast enough to produce this volume of trials. Precision is achieved in a reasonable time frame only by using high-performance computing.

Performance Benchmarks

The runtimes shown in Table 9 reflect calculation overhead relating to calculation of multivariate pods and statistics/risk metrics. The former occurs at the beginning of each HPC job whereas the latter is incurred at job conclusion. Both of them are done on the client computer. The simulations are done on cluster compute nodes, and they involve primarily the generation of random variables, including the lookup of pre-calculated multivariate simulations that were done by the client when the Excel workbook is prepared for upload to the HPC cluster. In order to improve cluster performance, the simulations received by the client from the compute nodes is written to disk rather than inserted immediately into the client worksheet. When the simulations are complete, this file is read and, at that time, the results are written to the simulation output worksheet. For trial counts in excess of 100,000, the insertion of new rows of data into this output area is a slow operation in Excel. This transfer and the subsequent derivation of statistics add time to the end of the job.

The calculation of multivariate simulation variates, particularly for large simulation counts can be relatively slow. The setup of multivariate random variables using the Iman-Conover methodology requires a Choleski decomposition and a potentially large matrix inversion. When the trial count approaches 100,000, this process is relatively slow because it has not been converted yet into compiled code in HPC Dynamo 5. The Iman-Conover code implementation relies on VBA code. Counts over 100,000 are commensurately slower. Similarly, when the trial counts are large the development of statistics and risk metrics after simulations are complete is also relatively slow. The effects of this overhead are apparent in Table 9. The simulations per second decline somewhat with

increased simulation count.

Table 9: Runtimes for HPC Dynamo 5[23]

Trials	Small HPC Cluster Runtime (27 core allocations over 5 nodes)[24]	Approximate Runtime Standalone (2 cores single node)
10000	1.15 mins	9 mins
25000	3.43 mins	22.5 mins
50000	9.53 mins	45 mins
100000	33.27 mins	90 mins
500000	407.34 mins	???
700000	946.56 mins	???

The potential power of parallelization and use of a computer cluster can be seen in Table 9. The runtimes using HPC are faster than running on a single computer and, for the larger sample sizes most appropriate for risk metrics the improvement is dramatically so. HPC cluster performance is never linear, and this is evident in Table 9. A substantial overhead occurs in loading an instance of Excel for each core and when the Dynamo workbook is opened by each core instance. There is an additional overhead for higher simulation volumes because of additional system activity in scheduling those simulations across the cluster. When a simulated array of DFA variables is completed by a cluster computer, it must be inserted into the client Excel instance of Dynamo. This too is an additional and significant source of overhead directly related to simulation volume. Pragmatically, even if the small cluster were only two times faster than a single computer for very high simulation volumes, 0.66 days for 700,000 simulations of 74 DFA variables is better than an estimated 1.31 days it might otherwise take for a single computer.

[23] The runtimes are for the simulation of 74 DFA variables and two multivariate pods. The time includes preparation of statistical and risk metric output for these variables. When simulation counts are large, the derivation of multivariate deviates takes more time because of sorting requirements involved in the Iman-Conover method. The runtimes are for the complete setup of multivariate values, simulations and derivation of statistics and risk metrics for all DFA variables.
[24] This is a very small HPC cluster. The performance gains over a single actuarial workstation are even more impressive given that they are derived from a modest extension of the workstation from 1 (standalone) to 5 nodes (computers). However, several of the additional computers are multiple-core servers.

CONCLUSION

This paper has used HPC Dynamo to identify the effects of sample size on DFA variable probability distributions. The impact of sampling error is so profound that a dilemma occurs. The number of trials needed to reduce the material impact of sampling error on risk metrics exceeds 100,000 trials. On a single computer the runtime becomes prohibitively large. The parallelization of simulations and their calculation on many simultaneous instances of Excel necessitates added expenditure for the cluster computers and multiple copies of Excel required for each of the node computers in the cluster. And, of course, each cluster computer must have an operating system. HPC Excel requires at least one server computer. The dilemma arises in that a reduction in sampling error to materially insignificant levels requires more trials that only can be achieved for increased costs.[25]

We have set out to answer the question of "How many simulations is enough." It is unlikely that any analysis of DFA variables involving less than 100,000-500,000 trials should be used, particularly when these variables are used to measure the effects of capital attribution or are used as proxies for risk-bearing measurements.

In the first part of this paper, the effects of sample size were examined within the context of aggregate probability distributions for correlated lognormal variables. This measurement was done using an aggregate loss distribution. We showed material impacts of sample size on the aggregate loss distribution and risk metrics such as Solvency II-styled calculations that rely on the properties of the aggregate loss distribution. The same observation applies across both parts of this paper—simulation volumes must be large and will require the use of high-performance computing. DFA variables in Dynamo can be constructed from any statutory, GAAP or cash flow variable. The probability distributions for these variables are highly sensitive to the number of simulation trials used in their estimation. Expected policyholder deficit, tail value at risk and value at risk decreased by 2% to 3% when simulations were increased from 100,000 to 700,000. They decrease by three to 6% when simulations were increased from 10,000 to 700,000. Variables such as VaR that are used in solvency compliance metrics have extreme sensitivity to simulation volume.

End Notes

In their 2009 paper, "A Holistic Approach to Setting Risk Limits," Burkett et al. observed that Dynamo 4.1 contained some inaccurate reconciliations among balance sheet, income statement, and cash flow statement values. Those inconsistencies remain in the Dynamo 5 model that has been

[25] At the time of this writing, HPC Excel running in Azure is only possible using an on-premise head node. The head node is a server computer. This computer is required if Azure is used and deployed in the VM Node role required for HPC Excel.

used for this paper. In the authors' views, these inconsistencies do not change the conclusions we have reached in this paper, but we do recommend that any user of Dynamo consider the potential effect of these inconsistencies on the results being produced and the usage of the results by their organization.

Acknowledgement

About Deloitte
Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a U.K. private company limited by guarantee, and its network of member firms, each of which is a legally separate and independent entity. Please see www.deloitte.com/about for a detailed description of the legal structure of Deloitte Touche Tohmatsu Limited and its member firms. Please see www.deloitte.com/us/about for a detailed description of the legal structure of Deloitte LLP and its subsidiaries. Certain services may not be available to attest clients under the rules and regulations of public accounting.
This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor.
Deloitte, its affiliates, and related entities shall not be responsible for any loss sustained by any person who relies on this publication