

# **The Importance (or Non-Importance) of Distributional Assumptions in Monte Carlo Models of Saving**

**James P. Dow, Jr.**

Department of Finance, Real Estate and Insurance  
California State University, Northridge

*Working Draft*  
*Not for quotation without author's permission*

January 6, 2009

Key Words: Monte Carlo, Saving

JEL Classification:

**ABSTRACT:** Monte Carlo models of saving use a wide variety of processes to generate their random returns. This paper examines in detail several different methods in the context of a typical savings problem. The paper finds that there isn't a dramatic difference between using a bootstrap procedure and drawing returns from a normal distribution. Whether the Great Depression is included or not in the bootstrap also doesn't matter significantly. What does matter is whether serial correlation is incorporated, particularly if bonds are a significant part of the portfolio. Unfortunately, each of the bootstrap methods that deal with serial correlation has significant flaws. Drawing random numbers from an autoregressive process avoids some of these flaws but introduces several estimation issues.

\*Department of Finance, Real Estate and Insurance, California State University, Northridge, CA 91330.  
(818) 677-4539. Email: james.dow@csun.edu.

## I. Introduction

Increasingly, Monte Carlo methods have been used to asses the results of saving plans (e.g Schleef and Eisinger, 2007, Cooley, Hubbard and Walz, 2003). Typically in these problems, an individual is planning on saving a certain number of dollars each year with a fixed asset allocation between stocks and bonds. A sequence of random asset returns is generated repeatedly to construct a distribution of final wealth. In practice, there are a number of different ways to generate the random returns. This paper investigates how much the assumption about the return generating process matters. Specifically, it investigates four questions:

*1) How many draws are needed to get a stable sample?*

Call one sequence of random returns and the corresponding final wealth a “run”. A simulation consists of a large number of runs, with the assumption the distribution of final wealth will not be affected by the particular draws. It turns out that even in simulations with a large number of runs ( $N=100,000$ ) there will be variation in means and standard deviations of final wealth from simulation to simulation. However, this variation does not seem to be economically meaningful; that is, the “margin of error” is smaller than the effect of changing the asset allocation between stocks and bonds by one percentage point.

*2) How much does the Great Depression and WWII matter?*

The years from 1924-1945 represent a period of unusual turmoil in the US economy as it went through the Great Depression and World War II. During this period, stock and bond returns showed increased volatility. It has been argued that this period should be left out of the sample due to its exceptionalness. Whether one should or not depends on the likelihood of events like this happening again (and the experiences of 2008 may settle *that* question). However, this paper asks a narrower question: does it make a material difference if you include this period? It turns out that the return differences are not large enough to materially affect the results.

*3) Does it matter if you use bootstrapping or draw returns from a normal distribution?*

While it is common to draw random returns from a normal distribution, there is good evidence that stock returns are not normally distributed. An alternative approach is to draw the random returns from the data set itself, i.e. bootstrapping. Distributions of final wealth are calculated using both a normal distribution and a bootstrapped distribution and then compared. While they are not exactly the same, the differences are not substantial.

*4) Does serial correlation matter?*

It has been argued that stock returns show both short-run momentum and long-run mean reversion. If so, the existence of serial correlation could affect the distribution of savings. Prior studies have tried to capture this effect using two different procedures. Cooley, Hubbard and Walz (2003) use overlapping periods, starting with the first year of the data sample and continuing until the run that ends with the last year in the data. Schleef and Eisinger (2007) use a random overlapping process. For each run, a random starting point is chosen and then followed sequentially until the last year in the data set is reached in which case a new point is chosen randomly and the process continues. Both procedures are biased. For the overlapping procedure, middle years are more likely to be chosen than end years, while in the random overlapping procedure, middle and late years are more likely to be chosen than early years. An alternate approach is to begin with each year sequentially and then if the last year is reached, continue with the first year. This is unbiased, in that all years appear equally, but introduces a connection between the starting and ending years that doesn't exist in the data. It turns out that the inclusion of serial correlation does significantly change the results, but each of the ways of dealing with it has their own serious flaws.

An alternate approach is to draw returns from a distribution of random numbers that incorporates serial correlation. Simulations were run with AR(2) processes for stock and bond returns. The results lie between the bootstrap procedures with serial correlation and the procedures drawing from normal distributions without serial correlation.

The following sections of the paper go through each of the four questions in more detail.

## **II. How many draws are needed to get a stable sample?**

Typically, simulations will take the average of a large number of runs to eliminate the effect of variation in the randomly drawn stock and bond returns. This raises the question, how large is large? To examine this, we will repeat this simulation process several times to evaluate the stability of the final distribution.

The economic problem is an individual saving for their retirement. The individual begins with no wealth and will save \$10,000 each year for the next 30 years. The asset allocation is assumed fixed over time (and for this set of simulations, equal to 50% stocks and 50% bonds). Returns are assumed to be normally distributed. The means, standard deviations and correlations were calculated from data from Ibbotson & Associates (Large-Cap Stocks and Corporate Bonds) for the years 1926-2005. All returns are adjusted for inflation.

Each simulation consists of a number of “runs”, where a run is one realization of lifetime saving and so one value of final wealth. The output of the simulation is the mean and standard deviation of final wealth across the runs. To evaluate the effect of the number of runs on the stability of the distribution, 20 separate simulations were performed, each simulation consisting of a large number of runs (for example, 10,000 runs each). Then the mean and standard deviation of the mean level of final wealth across the simulations was determined. For the process to be stable, the standard deviation of the mean results should be small across the 20 simulations.

Table 1 reports the results for runs equal to 10,000, 25,000 and 100,000. When calculating the mean level of final wealth using simulations with 10,000 runs, the results ranged from \$854,122 to \$871,633 with a standard deviation of \$4,376. Even though this might seem like a large number of runs, the dollar amount of final wealth does not converge to a fixed number. This test is repeated for N=25,000 and N=100,000. The range of results that you get and the standard deviation get smaller as N increases although the variation doesn’t disappear entirely. An N=100,000 the standard deviation is a little over 0.1% of the mean value.

Table 1. The Number of Runs and the Distribution of the Results (20 Simulations Each)

Runs	Mean	Standard Dev	Range
N=10,000	862,000	4,367	871,633-854,122
N=25,000	863,056	2,481	867,547-859,238
N=100,000	863,109	1,235	864,882-860,905

Is this too much? One way to ask the question is to see if the variation is economically meaningful. If we are looking at the expected dollar amount at retirement, being off by 1% due to calculation error is probably within the tolerance for error, given the general difficulties of calculation, and these numbers are certainly under that limit.

Another way to ask the question is to see if calculation error would affect the decision over asset allocation. To test this, we ran *one* simulation (of 100,000 runs) for each of a number of asset allocations. The share of stock went from 55% to 45%. The results are reported on Table 2. The mean of final wealth is reported for each asset allocation.

The difference between the results should be large compared with the random variance across simulations (equal to \$1,235 in Table 1). Since the differences in final wealth are about \$8,000 per percentage point change in asset allocation, it is unlikely that that random variation would produce an incorrect result in terms of the asset allocation decision. In other words, the error is not economically meaningful for this kind of problem. Because of this, we will use simulations of 100,000 runs for the remainder of the paper.

Table 2. Mean Final Wealth as a Function of the Share of Stock

Share Stock	Mean
55	909,954
54	901,283
53	888,742
52	881,772
51	872,326
50	860,964
49	852,902
48	845,281
47	835,750
46	826,532
45	817,995

### III. How Much Does the Great Depression and WWII Matter?

It has been argued that the Great Depression and World War II were atypical times for the economy and so should not be included when determining the distributions for stock and bond returns (although recent events may show it to be not so atypical). Kim, Nelson and Startz (1988) has argued that the finding of long-term mean reversion in stock returns is very sensitive to the inclusion of this period. Whether or not to include the Great Depression would certainly depend on subjective estimates of the likelihood of wild swings in stock prices in the future. However, we can ask the narrower question, does it matter whether we do so or not?

Table 6 reports means and averages of stock and bond returns for the time up through the Great Depression and WWII (1924-1947), the time after the WWII (1948-2005) and entire sample period.

Table 3. Bond and Stock Returns by Time Period.

	1924-1947		1948-2005		1924-2005	
	Stocks	Bonds	Stocks	Bonds	Stocks	Bonds
Mean	8.46	4.03	9.31	2.94	9.07	3.24
Stan. Dev.	27.16	8.43	17.28	10.20	20.29	9.70
Rho	.23		.29		.26	

Excluding the Great Depression period from the sample increases stock returns slightly and decreases bond returns slightly but does not make a dramatic difference. There is a slight decrease in stock volatility and an increase in bond volatility. Excluding the Great Depression will shift the optimal allocation towards stocks, but the differences in results would not be significant enough to outweigh other factors for determining inclusion. Unless one has *a priori* reasons for excluding this period, and conservatism and recent events argues otherwise, these years should be kept in the data. A separate issue is whether the Great Depression affects the long-run mean reversion properties of the sample. Issues related to serial correlation will be discussed in sections V and VI.

#### **IV. Does it matter if you use bootstrapping or draw from a random distribution?**

It has been argued that stock returns deviate enough from normal that instead of generating numbers randomly from a normal distribution it is better to use a bootstrapping process. To evaluate this, a bootstrap was used to generate a distribution of final wealth and this was compared to the previous results. There are 80 years of data in this sample, so for each year of the simulation, a random number between 1 and 80 was drawn and the stock and bond returns for that data year were used. This process preserves the cross correlation between bonds and stocks but does not take into account serial correlation (which is the same as with the simulations using the normal distribution). Table 4,A-C, reports the average mean and standard deviation of final wealth for one run with allocations of 100% stock, 50% stock and 0% stock. The first rows of the tables show the results when using the Normal and Bootstrap distributions. As can be seen, the results are extremely close. There is no practical advantage to using the Bootstrap procedure.

#### **V. Does serial correlation matter (using actual time series)?**

It has been argued that the existence of mean reversion in stock prices implies that investors should hold a relatively greater amount of stock when they are young, as the variability of stocks relative to bonds is relatively lower at longer investing horizons (Cochrane). Several papers have used a bootstrap process to do this (Schleef and Eisinger, 2007, Cooley, Hubbard and Walz, 2003)

We will evaluate three separate procedures for including serial correlation. The first procedure (Overlap) uses all 30-year contiguous periods starting in 1926. The last starting period is 1975, so that the 30-year stretch ends in the last data year of 2005. There is nothing stochastic in this approach, so the results are the average from 51 runs, one for each starting period. The advantage of this approach is that it does not add any correlation structure to the data that is not already there. The disadvantage is that it is biased since years in the middle of the sample period appear in the calculations more often than years at the ends of the sample.

A second approach (Wraparound) is a variation on this. When the run hits 2005 it “wraps around” and continues from 1926, so that every year in the data period is used as a starting point exactly once. The results are unbiased since each year is used an equal number of times, but it adds structure to the serial correlation properties of the data by assuming that the behavior right after 1926 follows that leading up through 2005.

A third approach was suggested by Schleef and Eisinger (2007) which I will call “Random Wrap”. When the simulation hits 2005 it continues at a new random starting point with each year from 1926-2005 being equally likely. While it also adds structure to the data, by randomizing the new starting point, it prevents the results from being driven by just one added connection. Unfortunately, the procedure is biased since years early on in the sample will be visited less often than those later. Since this procedure is random, the simulation is repeated for 100,000 runs.

The results are reported on Tables 4A-C for asset allocations of 100% stock, 50% stock and 0% stock. The three simulation methods that incorporate serial correlation are compared with two methods (Normal and Bootstrap) that do not. With 100% stock, the inclusion of serial correlation results in a moderate reduction in the average level of final wealth along with a dramatic reduction in standard deviation. The latter is consistent with the hypothesis of long-run mean reversion. Interestingly, there is not much difference between the three methods that incorporate serial correlation.

For 50% stock, we again see a reduction in the mean and standard deviations, but here the results across the serial-correlation simulations differ significantly, particularly with the Overlap process. With 0% stock, there are significant differences, with the Wraparound method showing a dramatic increase in variation. Obviously, the properties of bond returns are driving this.

Table 4A. Distributions of Final Wealth.

<b>Stock=100%</b>	<b>Mean</b>	<b>Standard Dev</b>
Normal	1,506,355	1,389,558
Bootstrap	1,506,896	1,399,566
Overlap	1,142,267	554,920
Random Wrap	1,258,069	560,528
Wraparound	1,203,474	542,875
AR2	1,377,026	1,074,361

Table 4B. Distributions of Final Wealth.

<b>Stock=50%</b>	<b>Mean</b>	<b>Standard Dev</b>
Normal	863,415	417,678
Bootstrap	864,692	413,754
Overlap	693,517	111,378
Random Wrap	789,690	296,096
Wraparound	838,006	318,886
AR2	863,361	405,410

Table 4C. Distributions of Final Wealth.

<b>Stock=0%</b>	<b>Mean</b>	<b>Standard Dev</b>
Normal	511,311	179,992
Bootstrap	511,473	177,989
Overlap	434,377	167,202
Random Wrap	502,384	250,228
Wraparound	571,182	304,951
AR2	544,052	257,378

To see why this is, we can compare bond and stock returns across the early, middle and later parts of the data periods on Table 5. While there are small differences in stock returns, there are dramatic

differences in bond returns. The Overlap process and the Random Wrap process are biased, which gives the middle years (with low bond returns) excessive importance. All three processes also take into account the effect of serial correlation in bond returns. Investors who invested 100% in bonds in the middle years would have seen dramatically different results from those who invested later. These effects will not show up in the Normal or Bootstrap procedures because of the mixing of periods. These kinds of streaks happen by chance, but only with relatively small probabilities.

Table 5. Mean Asset Returns Grouped by Years

	Mean Stock Return	Mean Bond Return
First 25 Years	9.37	3.68
Middle 30 Years	8.04	-1.07
Last 25 Years	10.00	7.98

The implication of these results is that the serial correlation properties of the data do matter substantially. Unfortunately, each of the ways of incorporating serial correlation has significant flaws. Alternate ways of including serial correlation, such as drawing returns from a random *serially-correlated* distribution, need to be considered. Furthermore, equal attention needs to be paid to bond prices since they have serial correlation properties that dramatically affect the results.

## VI. Does serial correlation matter (using draws from a distribution of random numbers)?

Given the various problems with using historical series of data, there may be some advantage to estimating a stochastic process than incorporates serial dependence and then use that to generate random returns. There is a well-established literature on long-run mean reversion in stock prices ( e.g. Poterba and Summers, 1988, and Fama and French, 1988, arguing for; Kim and Startz, 1988, arguing against) and the conclusion seems to be that there is some evidence for mean reversion, but that it's not estimated with

any precision and so we cannot conclusively reject either the hypothesis of mean reversion or a random walk.

This paper does not address the question of how best to test for mean reversion, rather it starts with a mean-reversive process and determines its affect on the simulation. An AR(2) process is separately estimated for bond and stock returns (where the autoregressive coefficients are for the deviation from the mean).

Table 6. Coefficients from AR(2) Regressions

	Constant	AR1coef.	AR2 coef.	Sigma <sup>2</sup>
Stocks	0.0903	0.0224	-0.1550	0.03966
Bonds	0.0328	0.1680	0.1054	0.00886

The correlation of the fitted residuals across the two equations was used as the correlation of the innovations for the simulation. As can be seen, stocks show mean reversion in the second year, while bonds show persistence in returns.

The results of simulations using the AR2 processes to draw returns are reported on Table 4, A-C (AR2). When compared with the Normal (or Bootstrap) simulations, the AR2 simulations show more volatility when wealth is primarily invested in bonds and less volatility when primarily invested in stock. For the 50/50 allocation, the effects pretty much cancel out. When compared with the wraparound simulation, the results are less extreme. There is less effect from mean reversion in stock and less effect from persistence in bonds.

There are obvious drawbacks to the AR2 process implemented in this paper. First, the actual stochastic generating process is likely to be much more sophisticated than a simple AR2 (and the evidence from the Wraparound runs shows that this matters). Second, the equations need to be estimated jointly and not independently. Finally, the simulations started each run at the mean return rather than drawing a starting point from a distribution of returns (that is, an investor may begin saving when the stock market is “out of equilibrium”). This raises another issue. If we are assuming that the stock market can be out of equilibrium, then investors should use that knowledge and adjust their asset allocation accordingly.

## **VII. Conclusion.**

Generally, the results of this paper are reassuring. As long as a large enough number of runs are used, it doesn't matter significantly how the returns are drawn. The exception to this is if one wishes to include serial correlation. Since the existence of mean reversion in stock returns is quite controversial, and it has a significant effect on Monte Carlo simulation of savings, it represents an important area for future research.

## VIII. References

- Boudoukh and Richardson, 1994, “The statistics of long-horizon regressions revisited” *Mathematical Finance* 4, 103-119.
- Cooley, Phillip, Carl Hubbard and Daniel Walz, 2003 A comparative analysis of retirement portfolio success rates: simulation versus overlapping periods. *Financial Services Review*, 12, 115-128.
- Kim, Nelson and Startz, 1988, “Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence”, *Review of Economic Studies*.
- Lo, 1991, “Long Term Memory in Stock Market Prices”, *Econometrica* 59, 1279-1313
- Poterba and Summers, 1988, “Mean Reversion in Stock Returns: Evidence and Implications”, *Journal of Financial Economics* 76, 1142-1151.
- Fama and French, 1988, “Permanent and Temporary Components of Stock Prices”, *Journal of Political Economy* 96, 246-273.
- Richardson, 1993, “Temporary Components of Stock Prices: A Skeptic’s View”, *Journal of Business and Economic Statistics*, 11, 199-207.
- Richardson and Smith, 1991, “Tests of Financial Models with the Presence of Overlapping Observations”, *Review of Financial Studies*, 4, 227-253.
- Richardson and Stock, 1989, “Drawing Inferences From Statistics Based on Multi-Year Asset Returns”, *Journal of Financial Economics* 25, 323-348.
- Schleef, Harold and Robert Eisinger, 2007, Hitting or missing the retirement target: comparing contribution and asset allocation schemes of simulated portfolios *Financial Services Review*, 16, 229-243.