Searching For Lost Decades in U.S and Global Equities

Blake LeBaron *
International Business School
Brandeis University

June 2010, Revised: June 2012

Abstract

This paper estimates the probability of a "lost decade" where equity investments lose value over a ten year period. The findings are a reminder that equity investments are risky even over longer time periods, and investors should take this into consideration when making portfolio choices. It also introduces a simple method to allow the reader to combine beliefs about long run stock returns along with computer simulated return distributions. Finally, the results for the U.S. are augmented with international data which strengthen the case for long horizon risk.

*International Business School, Brandeis University, 415 South Street, Mailstop 32, Waltham, MA 02453 - 2728, blebaron@brandeis.edu, www.brandeis.edu/~blebaron. The author is also a research associate at the National Bureau of Economic Research.

1 Introduction

Financial time series generate many interesting psychological milestones which cause great excitement to investors. In the U.S. the level of the Dow Industrial index is often compared with $10,000$, and previously, $1,000$ as a benchmark. At the end of 2009 we observed another milestone, the occurrence of a "lost decade." Over the previous 10 years the return of a diversified portfolio of U.S. stocks lost value in nominal terms. Even before adjusting for positive inflation over the period, the investor putting a dollar in the market in 2000, reinvesting dividends, and compounding until 2009 would have ended up with less than a dollar. This paper examines the probability of a U.S. equity investment losing value over a decade. This is done while imposing minimal assumptions on long range returns data. It also allows readers to use their own beliefs about future equity returns.

The perspective of this paper is to use this long history to analyze U.S. long term equity performance. The historical return record is used to impose minimal assumptions on the distribution of returns in a bootstrap based methodology. There is a vast literature exploring the long range properties of financial market data. The exploration of the equity premium is the most extensive.[1] This paper considers equity returns alone, and only a single feature of their distribution. There are several reasons for doing this. First, this stays with the "lost decades" theme already discussed. Second, it avoids having to estimate or proxy for risk free rates of return in early periods when this might be difficult. Finally, it keeps the bootstrap methodology relatively simple and assumption free.

This paper's U.S. centered data perspective will be augmented using the cross sectional country return series used in Dimson et al. (2002) and Dimson, Marsh & Staunton (2008). This data set contains a comprehensive asset return cross section from 18 industrialized countries starting at the beginning of the twentieth century. These series are used as outside information that should influence investors beliefs about the U.S. empirical experience.

Section 2 introduces and summarizes the historical return series. Section 3 performs some initial monte-carlo runs which will justify much of the bootstrapping methodology in the paper. Section 4 presents the main estimates of decade losses assuming independent annual returns. Section 5 extends the bootstrap to handle dependent returns. Section 6 compares the U.S. results to an international data set, and section 7 concludes.

[1] There are many collections of classic papers on this topic. These include Goetzmann & Ibbotson (2006), Mehra (2008), and most recently P. Brett Hammond, Leibowitz & Siegel (2011). Many books have covered the topic using extensive long range data sets both U.S. and international. These include Cornell (1999), Dimson, Marsh & Staunton (2002) and Siegel (2002).

2 Return Summary

The long return history is built by merging two stock return data sets. The first is the monthly returns series described in Schwert (1990) which extends back to 1802. The second is the annual series, beginning in 1871, constructed by Shiller, and used in Shiller (2000).[2] Shiller's data set also includes inflation series from 1872 on. This is augmented with inflation series obtained from "Measuring Worth."[3] From 1918 on the stock returns are from the S&P composite portfolio. From 1872-1917 the stock market information is from the indices created in Cowles & Associates (1939) to track aggregate stock market movements. Earlier returns were assembled by Schwert to best track aggregate market movements. They involve several different sources, and become relatively narrow indices as one moves back in time.[4] In the earliest period they are mostly bank stocks, and in later periods they include railroad stocks. There are obvious survivorship biases in these indices. Also, many are constructed as monthly averages from bid and ask prices, making precise time series analysis at higher frequencies difficult.[5]

Figure 1 gives an overall picture of lost decades. The y-axis plots the total decade return including dividends for the 10 years ending on the year given by the x-axis. Both nominal, and inflation adjusted real returns are plotted. For the decade ending in 2009 both real and nominal returns are negative indicating that investors would have lost value on their equity investments. The appearance of lost decades in this figure is visually relatively rare. The decades ending in 1858, 1939, and 1940 are the only other decades with negative nominal equity returns. If one considers real returns, then several others appear. The most recent of these would be the decades ending in the late 1970's and early 1980's where large U.S. inflation adds a heavy cost to a relatively flat market. This picture is interesting, but not informative as to how likely these events are. The analysis below addresses this question.

The first two rows of table 1 provide summary statistics for the annual holding period returns used to construct the previous figure. The mean nominal and real annual returns of 9.1 and 7.8 percent, respectively, represent this 209 year U.S. history of equity returns. The table also reports the standard deviation of about 17 percent per year for both series. This estimate will not be surprising to most investors familiar with the properties of long range return series. The table also presents skewness and kurtosis levels for these series. These are quick tests for whether a normal distribution would be a reasonable approxi-

[2]Both of these data sets are available at the authors' websites.
[3]See http://www.measuringworth.com/ for full information on the methodology behind the early inflation estimates.
[4]Also, dividend yields can only be approximated for the earliest samples from the first part of the 19th century.
[5]See Schwert (1990) for detailed discussions.

mation. Skewness is near zero, and Kurtosis near 4, which is larger than its value of 3 for a normal distribution. The last column in the table presents a simple test for normality, the Jarque-Bera test. The values given are the p-values corresponding to the normal null hypothesis with unknown mean and variance.[6] The first two rows indicate a weak rejection of normality for the nominal returns, and a borderline p-value of 0.09 for the real returns.

Long horizon returns should not be expected to follow a normal distribution since they are compounded short horizon returns,

$$r_{t,A} = \prod_{j=0}^{11}(1 + r_{t-j}) - 1 \tag{1}$$

where r_t are monthly returns. The logged annual returns would then be

$$\log(1 + r_{t,A}) = \sum_{j=0}^{11} \log(1 + r_{t-j}). \tag{2}$$

If these annual logged returns are independent with finite second moments, then geometric, or log returns at longer horizons must approach normality. Lines 3 and 4 of table 1 report the same set of summary statistics for the logged annual nominal and real returns. Results are similar to those for the holding period returns, except for a moderate increase in kurtosis. This shows up in the Jarque-Bera statistics yielding a p-value of zero in both cases. A visual test for normality is presented in figure 2 which shows the density for the nominal holding period and log returns superimposed with a normal density. The figures suggest that the practical deviations from normality in these series may be quite small. This will be checked in the later simulations.

The last four rows in table 1 present subsample estimates using 1871 as a break date. This is the date when the Schwert series ends, and the Shiller series begins. Also, a plot of the full return time series is presented in figure 3. The figure shows no obvious visual changes in the series. The table shows that the earlier subsample contains a reduced mean and standard deviation as compared to the full sample. It also displays much larger kurtosis. In the latter periods kurtosis falls enough so that the Jarque-Bera test is unable to reject normality for both real and nominal returns.[7]

[6] Since small sample properties of the Jarque-Bera test are questionable, the p-values report results from a 100,000 length monte-carlo.

[7] One might initially suspect that there are more tail events driving the larger kurtosis in the first subsample, but this does not seem to be the case. In looking at figure 3 one sees similar tail behavior, but a higher concentration of smaller return years (in absolute value) in the earlier part of the data set. This might be driven by the fact that this is a relatively narrow index, and some years may not have

3 Monte-carlo Experiments

The initial results on the series suggest caution in terms of assumptions about stock returns. Normality and log normality appear plausible, but there is enough evidence against normality, that bootstrap simulations using the empirical distributions may be necessary. Table 2 performs a test of the bootstrap methodology that will be used throughout the paper. The procedure will assume returns are independent over time. Drawing a new decade of returns involves drawing a new sample, often much longer than the original. A sample of decade length returns is then constructed from this longer series.[8]

The first two lines in table 2 begin with a simple monte-carlo drawing returns of a given sample size from a log normal distribution. The mean and standard deviation of this distribution are set to the original sample estimates. For each monte-carlo run the ten year portfolio return is calculated, and the simulation records whether it is a gain or loss over a decade. The table reports the root mean squared error (RMSE) of the estimated loss probabilities against the true value.[9]

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{p}_i - p)^2}. \tag{3}$$

The first row uses non-overlapping decade periods to estimate the decade loss. The second row uses overlapping decades which utilize all possible decade combinations for the generated time series. MSE estimates fall as the sample sizes increase, and as expected the overlapping data provides more precise estimates.

If one is prepared to accept the hypothesis that annual returns are well approximated by a log normal distribution, then for each given sample, the estimate of long range loss probabilities should use the normal cumulative density function (CDF) on estimates of decade returns and volatility. The third row reports simulations using the same null distribution, but each sample is used to estimate the mean and standard deviation for one year log returns. This is then adjusted to the correct distribution for decade length log returns by multiplying the mean and standard deviation by 10, and $\sqrt{10}$, respectively. The probability of a decade loss, is then estimated from a normal CDF. These values are reported in the third row of table 2. Given that the assumption of log normality is true

many large information events in the industries covered.

[8]For example, one can construct a series of length $T = 100,000$ years from the 209 years of returns. There is obviously huge overlap in reusing data. However, it is important to understand how far one can push this experiment in terms of generating improved information on long range returns.

[9]One can also estimate the bias and variance of these estimates. For thŔse estimators thŔ bias turned out to the small, and therefore the variance is approximately equal to the MSE.

in the monte-carlo, the MSE of this estimator is going to be smaller than the estimates from the empirical quantiles as is shown by the reduction in MSE from the first two rows. Beyond sample sizes of 1000 this reduction is modest, especially as compared to the case using overlapping data.

The final two rows of table 2 report monte-carlo simulations of the bootstrap procedure that will be used in the paper. For each monte-carlo iteration a 200 year return series is drawn following the log normal distribution set to the sample mean and variance from the merged historical returns. To estimate loss probabilities new series are redrawn with replacement with length T ranging from 100 to 50000. The column labeled, 5000, uses 5000 returns drawn from the original 200 return series to estimate the decade loss probability.Moving along either of the last two rows shows that there is improvement in precision as the bootstrap sampling length is increased beyond the original sample size. The improvement is somewhat faster for the overlapping case, but both cases reach a minimum by sample size 5000. Their improvement over equivalent monte-carlo runs is interesting. At a sample size of 5000 the bootstrap shows a MSE of 0.048 for both the overlapping and non overlapping cases which compares favorably with values of 0.072 and 0.057 for the 200 sample monte-carlo cases in rows one and two. This also compares favorably with the estimated MSE for the normal CDF of 0.046 from row three. It is interesting that the bootstrap estimator does not give up much in terms of precision as compared to the normal CDF case, which suggests that it is a powerful robust estimator for these probabilities.

4 Independent Returns

Long return series will be generated from the 209 annual return series by sampling with replacement. The length of the sample will be fixed at $T = 250000$ which goes well beyond the sample lengths used in table 2. This is done to drive the variability from the bootstrap sample close to zero using a very large monte-carlo. Obviously, data sampling variability is still important, since there are still only 209 data points. The results in table 2 give an idea of what our precision should be using a bootstrap on the original sample.

Figure 4 uses these simulated decade returns to generate a histogram of ending wealth levels for an investor starting with a 1 dollar investment at the beginning of a decade. Two important features should be noted. First, the area to the left of the black line indicates the lost decades, or periods when the investment lost money. Although small, the area is not insignificant. The well known near log normality of this distribution drives the strong right skew which is evident in the figure. At decade lengths it is interesting to note that probabilities are not insignificant for wealth doubling, or even increasing six fold.

Table 3 presents estimates of decade loss probabilities from the long time series sampling exercise with $T = 250,000$ years, using overlapping decades. The first column and row correspond to the simulation presented in the previous figure. This is a bootstrap using the actual returns in the data, and yields a probability of 0.072.[10] It is generated from returns with a nominal mean return of 9.1 percent per year. The second column repeats this estimation for real returns where the mean return falls to 7.7 percent per year in the sample. The probability of a lost decade in this case rises to 0.12. The second line in the table reports bootstrap standard errors for each of the probability estimates. These values are calculated by drawing a new set of 209 annual returns with replacement, and then using these as a sample for the expansion to the long T sampling methodology for decade return estimation as is done for the actual returns. The bootstrap procedure is repeated $1,000$ times, and the standard deviation across this simulations is shown in parenthesis. For the first two return assumptions standard errors of 0.034 and 0.047 show that the estimated probabilities still contains a large amount of uncertainty. This is a reminder that with 200 years of data we will not be able to make strong statements about the tails of decade return distributions.

The sample mean returns from the data may not be the best estimate of expected returns for future decades. This paper does not take a strong stand on what future returns should be, but other return assumptions can easily be applied to the simulations.[11] The last three columns in table 3 estimate loss probabilities assuming annual mean returns of $10, 6$, and 4 percent. The sample mean is subtracted from the historical series, and one of the three expected return assumptions is added to the returns. It is clear that the estimated loss probabilities change dramatically as assumptions about returns change. For example, if investors are only expecting a 6 percent return, then they should expect a loss over a decade arriving with probability 0.186. The optimists, at 10 percent, should expect a decade loss only about 5 percent of the time.

Figure 5 shows the impact of various return assumptions on the expected decade loss probabilities. Readers can quickly put their own return assumption on the x-axis to as-

[10] The standard error for this estimate will be approximately $\sqrt{p(1-p)/n}$ using the asymptotic approximation to the binomial trial. Results in table 2 show that the nonoverlapping, and overlapping cases are similar, so for $T = 250,000$, the non overlapping decade estimates correspond to $n = 25,000$. For $p = 0.07$ this yields an estimate of 0.0016 which was confirmed with a slightly lower value of 0.0015 from a monte-carlo experiment which used the longer sample, and overlapping decades. The variability from the long length sampling is small which is the reason for choosing the large value of T.

[11] Investors may not agree with either of these long run expected return levels. See P. Brett Hammond et al. (2011) for many different perspectives on this. Also, Welch has conducted survey results that are reported on his website http://research.ivo-welch.info/equpdate-results2009.html.

sess their long term chances of a loss. The probabilities are calculated for a discrete set of expected returns using the same bootstrap long range sampling methods used in the table. For example, at an expected annual return of 8 percent, the decade loss probability is approximately 10 percent. This falls to 2.5 percent for assumed long range returns of 12 percent.

The next two rows in table 3 check whether the results are driven by returns coming from the earliest part of the sample. The sample is restricted to the period from $1872-2010$, and the same loss probabilities are estimated. The decade loss probabilities of 0.067 and 0.122 for the real and nominal returns show little change from the full sample. The mean adjusted returns are again used in the last three columns. For these simulations the means will be exactly the same as in the full sample runs, but other aspects of the distributions will be different.

The results presented so far have made minimal assumptions for the return distributions. The final results in table 3 assume that long horizon returns are log normal. In this case estimating the loss probability involves only estimating the mean and standard deviation for log returns, expanding these to decade length, and then using a normal CDF. Results of this estimation are given in the row labeled "log normal." The numbers do not change substantially from the corresponding bootstrap values. For example, for real returns, the probability goes from 0.120 for the data to 0.118 under the log normal assumption. The numbers in parenthesis correspond to bootstrap standard errors estimated by redrawing the annual returns, and using the new series for mean and standard deviation estimation. In other words, it recreates the estimation error on the two moments which go into the log normal probability estimates.

One possible reason for the relatively large probabilities in table 3 is that the ten year horizon is too short. Figure 6 displays loss probability estimates over a range of time horizons. These values are estimated using the methods from table 3 with both the bootstrap and log normal assumptions. As the horizon increases, the probabilities fall as expected. However, even at a 20 year horizon, the point estimate of the probability of a loss is still near 5 percent which is not zero. One has to move out to nearly the 40 and 50 year horizons to consider the loss probabilities as negligible. Beyond the very short horizons, the impact of the normality assumptions is negligible.

5 Dependent Returns

Up to this point returns have been assumed to be independent over time. If long range returns were dependent this could change the estimated risk of decade returns. Figure 7 presents the estimated autocorrelations for the real and nominal returns with asymptotic 95 percent confidence bands around the uncorrelated null hypothesis. Evidence for any correlation in these series is very weak. Combining the first 5 autocorrelations into a Ljung/Box test yields a p-value of 0.07 for the nominal returns, and 0.16 for the real returns. Figure 8 looks at another measure of long range dependence, the variance ratio. Assuming σ^2 is the variance of annual log returns, independence gives m period return variances of $m\sigma^2$. Deviations from this are a direct measure how risk is increasing with horizon lengths.[12] The figure shows the variance ratio declining from one, indicating some long range mean reversion, but the asymptotic confidence bands remind us that the sample length is still too short to say anything significant about these values.[13]

The results on dependence show some weak indications of mean reversion. A cautious tester should consider some form of dependent bootstrapping to explore the possible impact of this. Two different methods for generating dependence in the annual returns series will be used here. First, a parametric bootstrap is performed, using an estimated autoregressive model with 5 lags, an AR(5), for the returns process. The model and residuals are estimated, and the residuals are then redrawn with replacement and used along with the parameter estimates to generate a new simulated time series of length $T = 250,000$.[14][15] The independent loss probabilities are repeated in the first row of table 6 for comparison. The second row presents the estimated loss probabilities using the AR(5) null model, using overlapping decades on the long sample as in table 3. There is some reduction for both the real returns, and the nominal returns series indicating the data has some weak information about long run mean reversion. However, the probabilities are still not trivial.

A second simulation experiment uses the stationary bootstrap to replicate the dependence in a nonparametric fashion.[16] This form of bootstrap draws returns in contiguous blocks where block length is controlled by a random variable, X_t which is 1 with probability λ, and 0 with probability $1 - \lambda$. Assume a new time series is being constructed at the

[12]See Poterba & Summers (1988) and Lo & MacKinlay (1988) for early examples of this.

[13]The 95 percent confidence bands are generated as in Lo & MacKinlay (1988).

[14]See Efron & Tibshirani (1993) for a basic description of the parametric bootstrap, and Maddala & Li (1996) for financial applications. Killian & Berkowitz (2000) is a useful survey which covers many issues on modeling dependence in time series.

[15]Given the reported autocorrelations, the AR(5) would appear to be an over fit model. Kilian (2001) shows that when bootstrapping standard errors, the dangers of under parameterization outweigh those of over parameterization.

[16]See Politis & Romano (1994) for the original derivations. Also, see Sullivan, Timmerman & White (1999) for a financial application.

current point t, and is drawn from point τ in the original series. The next point, $t+1$, will come from $\tau + 1$, if $X_t = 0$, and will come from a new point $\hat{\tau}$, if $X_t = 1$. This generates a series containing blocks of varying lengths from the old series, where the lengths are controlled by the behavior of X_t.[17] Results for this simulation are given in the second line of the table. For these runs, $\lambda = 0.2$ which gives an average block length of 5 years.[18] The values are close to those from the AR(5) simulation, and indicate that these two methods may be replicating a similar amount of dependence for the long range returns.

Figure 9 explores the possibility that weaker long range dependence might yield evidence for lower decade loss probabilities. The parametric AR simulation is run for lags of 1 through 25, and the loss probabilities, estimated using the previous methodology, are plotted. This figure uses the merged data set real returns as the starting data series. There is an early sharp drop off in probabilities as the lags increase from 1 to 5. However, at this point the probabilities stabilize near $0.07 - 0.08$ which is consistent with table 6. Adding further lags appears unlikely to dramatically reduce estimated long range losses. This is consistent with the evidence that long range mean reversion is weak.

6 International Cross Section

Up to this point the analysis has concentrated on only long range series built from U.S. returns and inflation series. Series used in Dimson et al. (2002) provide a useful long range cross section for comparison.[19] The series are annual, and extend from 1900 though 2010 for 111 years of annual data. Only real equity returns will be used here.

Table 5 presents the results for the international returns. It includes all the countries in the data set along with value and equal weighted portfolios. The first two columns record the annual mean and standard deviation for the logged returns. Bootstrap repeats the long resampling procedure taking each series out to $250,000$ observations to estimate the decade loss probability. The column labeled normal uses the independent log normal return assumption, and uses the sample means and standard deviations to estimate the 10 year loss probabilities.

The results from table 3 reported a decade loss probability for real returns in the U.S. of 0.12. Comparison of this number with values in table 5 shows that relative to the rest of the world, the U.S. is a safer country than most when it comes to long term tail risk. Using the bootstrap estimator, countries vary from a high of 0.396 for Italy to a low of

[17]The varying block lengths follow a geometric distribution with mean $1/\lambda$.

[18]Values for 10 and 20 years have also been tried generating similar results.

[19] These series are available from Morningstar.

0.108 for Australia. Similar to the previous results, the normal approximations do not have a large impact on the results, which continues to support the idea that normality is not a bad assumption at long horizons.

A graphical summary of this table is given in figure 10 which displays a histogram of the bootstrapped loss probabilities from table 5. This gives a clear picture of where the U.S. lies in terms of long run risks. Also, it shows that if investors are going use this data to adjust their beliefs about risk in the U.S., they should increase their risk assessment. Finally, the last two lines in table 5 present results for both value and equal weighted global portfolios. These results show surprisingly small reductions in risk from either form of diversification. The equal weighted portfolio reduces the the decade loss probability to 0.10 which is lower than the individual countries, as it has to be, but the gain is small. Furthermore, given various frictions to international investing over the early parts of this sample, the feasibility of achieving these returns should be viewed with some skepticism.

Table 6 performs some additional experiments, exploring the cross sectional dimensions of the real return data, and how it impacts investors. All the experiments are bootstrap simulations done with $10,000$ replications. Also, for calculating the loss probabilities each simulation uses the normal approximation using estimated means and standard deviations.

The first experiment assumes a random draw of new returns data which is structured by country. First, one of the countries is drawn at random from the pool. Then its returns are redrawn with replacement giving a new sample which is used to estimate the mean and standard deviation. The investor here is viewing the countries as different, but could potentially face data that looks like any one of them going into the future. The columns report the $0.1, 0.5, 0.9$ quantiles for the decade loss probability distribution. The median value of 0.245 is consistent with the cross sectional results, and the graphical information in figure 10, all of which show that the probability of a lost decade is large in the international returns series. The bootstrap runs generate a large amount of dispersion with the quantiles ranging from 0.099 to 0.445.

The second row of the table pools the entire data set into one set of returns, and then draws country samples from this pooled population. This implements a null hypothesis that all country returns come from the same population. Pooling all the data reduces the dispersion from the separate country sampling method as is seen in the narrowing of the extreme quantiles to 0.136 and 0.415. However, the median value of 0.260 is does not change much by moving to the pooled sample.

The next two rows labeled, "Mean dispersion," gets the standard deviation from the pooled return series which is fixed across all simulations. The mean is estimated by again drawing a random country, and redrawing its returns as before. The purpose of this is to explore the variability in sample means alone, while imposing the hypothesis of a common variance. Results in this case do not vary much from those in the first two simulations. The next line, labeled, "Std. dispersion," reverses this experiment to using the pooled mean estimate, but standard deviations coming from each randomly drawn country. In this experiment the dispersion of the results is greatly reduced with quantile levels at 0.198 and 0.315. This suggests that much of the sample variability is coming from changes in sample means rather than variances.

The last two rows test the impact of dependence on the results. None of the international returns series show much evidence for return autocorrelation, but the importance of this is tested by repeating the parametric bootstrap that was used before with the U.S. returns. Two models, an autoregressive model of order 2, AR(2), and an AR(5) are used as a simple dependent null hypothesis. These are estimated once on each country, and then simulated for $T = 250,000$ years to estimate the decade loss probability. The quantiles represent the cross section across the countries. The results again show little change from the earlier results that assumed independence in the returns. The median loss probabilities for the two dependent cases are 0.252 and 0.233 for the AR(2), and AR(5) respectively. Evidence from the international returns again suggest that long range return dependence should not reduce investors' beliefs about the riskiness of decade returns.

7 Summary

Lost decades are often treated as a kind of "black swan" event that are almost impossible. Results in this paper show that while they are a tail event, they may not be as far out in the tail as the popular press would have us think. Allowing the data to speak directly in an independent bootstrap, with two centuries of return time series, the estimate of a portfolio loss over a decade is about seven percent. A life long investor facing 6 decades of investments should consider a probability $0.35 = 1 - (0.93)^6$ of seeing at least one lost decade in their lifetime.[20] For the investor concerned with real returns the results are more depressing with decade loss probabilities of twelve percent. The bootstrap methodology is not dependent on distributional assumptions about annual returns. However, in the reported estimates, normality assumptions for annual returns do not have a major impact

[20] The probability having 6 decades with positive returns would be $0.93^6 = (1 - 0.07)^6$. One less this probability is the chance of seeing at least one losing decade.

on decade length results.

The estimated loss probabilities are checked for robustness in two ways. First, the independent null hypothesis is weakened by using several methods for simulated dependent long range returns. In all the tests there is only a small reduction in the decade loss probability, which is consistent with the very weak mean reversion present in aggregate long range returns. Second, the U.S. experience is compared with international equity data using several different tests. Consistent with other research, global data should not give U.S. investors any increased confidence in terms of risk. On the contrary, long run results across the globe consistently appear riskier in terms of decade losses in real equity returns.

The simple message here is that stock markets are volatile. Even in the long run volatility is still important. These results emphasize that 10 year periods where an equity portfolio loses value in either real or nominal terms should be an event on which investors put some weight in making their investment decisions.

References

Cornell, B. (1999), *The Equity Risk Premium: The Long-Run Future of the Stock Market*, Wiley, New York, NY.

Cowles, III, A. & Associates (1939), *Common Stock Indexes*, number 3 *in* 'Cowles Commision Monograph', 2nd edn, Principia Press, Bloomington, Indiana.

Dimson, E., Marsh, P. & Staunton, M. (2002), *Triumph of the Optimists: 101 Years of Global Investment Returns*, Princeton University Press, Princeton, New Jersey.

Dimson, E., Marsh, P. & Staunton, M. (2008), The worldwide equity premium: A smaller puzzle, *in* R. Mehra, ed., 'The Handbook of the Equity Risk Premium', Elsevier, Amsterdam, Netherlands.

Efron, B. & Tibshirani, R. (1993), *An Introduction to the Bootstrap*, Chapman and Hall, New York.

Goetzmann, W. N. & Ibbotson, R. G. (2006), *The Equity Risk Premium*, Oxford University Press, Oxford, UK.

Kilian, L. (2001), 'Impulse response analysis in vector autoregressions with unknown lag order', *Journal of Forecasting* **20**, 161–179.

Killian, L. & Berkowitz, J. (2000), 'Recent developments in boottrapping time series', *Econometric Reviews* **19**(1), 1–48.

Lo, A. W. & MacKinlay, A. C. (1988), 'Stock prices do not follow random walks: Evidence from a simple specification test', *Review of Financial Studies* **1**, 41–66.

Maddala, G. S. & Li, H. (1996), Bootstrap based tests in financial models, *in* G. S. Maddala & C. R. Rao, eds, 'Handbook of Statistics', Vol. 14, North-Holland, Amsterdam, pp. 463–488.

Mehra, R., ed. (2008), *Handbook of the Equity Risk Premium*, Elsevier, Amsterdam, the Netherlands.

P. Brett Hammond, J., Leibowitz, M. L. & Siegel, L. B. (2011), *Rethinking the Equity Risk Premium*, CFA Institute.

Politis, D. & Romano, J. (1994), 'The stationary bootstrap', *Journal of the American Statistical Association* **89**, 1303–1313.

Poterba, J. M. & Summers, L. H. (1988), 'Mean reversion in stock prices: Evidence and implications', *Journal of Financial Economics* **22**, 27–59.

Schwert, G. W. (1990), 'Indexes of United States stock prices from 1802-1987', *Journal of Business* **63**, 399–426.

Shiller, R. J. (2000), *Irrational Exuberance*, Princeton University Press, Princeton, NJ.

Siegel, J. J. (2002), *Stocks for the Long Run: The Definitive Guide to Financial Market Returns and Long-Term Investment Strategies*, 3rd edn, McGraw-Hill, New York, NY.

Sullivan, R., Timmerman, A. & White, H. (1999), 'Data-snooping, technical trading rule performance and the bootstrap', *Journal of Finance* **54**, 1647–1691.

Table 1: *Summary Statistics*

Series	Mean	Std.	Skewness	Kurtosis	Normal (p-value)
Nominal Returns	9.1	16.8	0.18	3.79	0.04
Real Returns	7.8	17.1	0.10	3.68	0.09
Nominal Log	7.5	15.9	-0.53	4.55	0.00
Real Log	6.2	16.4	-0.53	3.91	0.00
Nominal (1802-1871)	6.9	14.0	1.18	7.34	0.00
Nominal (1872-2010)	10.2	18.1	-0.13	3.13	0.78
Real (1802-1871)	7.1	15.6	0.64	5.90	0.00
Real (1872-2010)	8.1	17.8	-0.10	2.99	0.89

Annual arithmetic and geometric returns. Sample length is 209 years covering returns from years ending 1802-2010. All returns include dividends. Mean and Std. estimates are in annual percentages. Kurtosis estimates the kurtosis in the return distribution. This value would be 3 for a normal distribution. Normal refers to a Jarque-Bera test for normality. The value reported is the p-value for the normal null hypothesis. The period 1802-1871 uses the data from the Schwert (1990), series. The period 1872-2010 uses the dataset contained in the Shiller (2000), annual series.

Table 2: *Root Mean Squared Error Estimates*

	Sample Size							
Monte-carlo	100	200	500	1000	5000	10000	25000	50000
Log normal: no overlap	0.102	0.072	0.045	0.032	0.014	0.010	0.006	0.005
Log normal: overlap	0.083	0.057	0.036	0.025	0.012	0.008	0.005	0.004
Normal CDF	0.066	0.046	0.029	0.021	0.009	0.006	0.004	0.003
IID Bootstrap(200): no overlap	0.111	0.086	0.065	0.057	0.048	0.048	0.047	0.046
IID Bootstrap(200): overlap	0.096	0.075	0.058	0.053	0.048	0.047	0.046	0.048

Root mean squared error of decade loss estimates using simulated log normal annual returns. Simulation cross section is $10,000$ runs for all experiments. **Log normal: no overlap** simulates non overlapping decades of returns for the given sample sizes. **Log normal: overlap** repeats the previous simulation for the same sizes, but uses overlapping decade returns. **Normal CDF** estimates the mean and standard deviation for each set of simulated annual data in the monte-carlo, transforms these to decade values by multiplying by 10, and $\sqrt{10}$, respectively, and then uses a normal CDF to estimate the loss probability. The final two rows perform bootstrap experiments that replicate procedures that will be used with the actual series. In each monte-carlo run, a 200 year sample is drawn with normal returns, and these are then redrawn with replacement to generate the sample sizes given by the column headings.

Table 3: *Probability of Decade Loss: Independent Returns*

	Annual Return				
Simulation	9.1% (Nominal)	7.7% (Real)	10%	6%	4%
Bootstrap	0.072	0.120	0.052	0.186	0.300
	(0.034)	(0.047)	(0.028)	(0.061)	(0.075)
Bootstrap: 1872-2010	0.067	0.122	0.073	0.218	0.332
	(0.042)	(0.060)	(0.045)	(0.082)	(0.098)
Log Normal	0.069	0.118	0.047	0.191	0.313
	(0.036)	(0.048)	(0.028)	(0.062)	(0.078)

Probabilities of a loss in equity portfolios compounded with dividends over decades. All mean returns are adjusted to the the given annual mean return shown in the column heading. Normal uses log normally distributed returns adjusted to estimated standard deviations. "Bootstrap" uses actual returns with readjusted means. Returns are drawn independently from the set of all annual returns.

Table 4: *Probability of Decade Loss: Dependent Returns*

Simulation	Real(7.7%)	Nominal(9.1%)
Independent	0.120	0.072
Autoregressive: AR(5)	0.081	0.042
Stationary	0.106	0.051

Dependent bootstraps. Autoregressive estimates an autoregressive model with 5 lags on the return series, and uses the estimated parameters and resampled residuals to generate dependent data. The stationary bootstrap draws random blocks from the original time series.

Table 5: *Probability of Decade Loss: Country real equity returns*

Country	Mean	Std	Bootstrap	Normal
Australia	0.072	0.179	0.108	0.103
Belgium	0.025	0.228	0.360	0.364
Canada	0.057	0.166	0.136	0.139
Denmark	0.050	0.187	0.194	0.200
Finland	0.052	0.276	0.268	0.274
France	0.030	0.226	0.329	0.337
Germany	0.030	0.345	0.326	0.391
Ireland	0.037	0.233	0.287	0.309
Italy	0.020	0.291	0.396	0.415
Japan	0.037	0.326	0.306	0.358
Netherlands	0.049	0.203	0.220	0.225
Norway	0.041	0.238	0.285	0.292
New Zealand	0.057	0.184	0.157	0.166
South Africa	0.071	0.204	0.132	0.137
Spain	0.035	0.209	0.291	0.299
Sweden	0.061	0.215	0.179	0.184
Switzerland	0.041	0.189	0.238	0.244
United Kingdom	0.052	0.193	0.186	0.197
US	0.061	0.199	0.166	0.167
World	0.053	0.173	0.165	0.165
Equal weighted	0.061	0.148	0.100	0.097

Summary statistics and decade loss probabilities for developed country returns. Bootstrap repeats methods from table 3, and Normal assumes a log normal return distribution.

Table 6: *Probability of Decade Loss: Cross section experiments*

Experiment	$q_{0.10}$	$q_{0.50}$	$q_{0.90}$
Country draw	0.099	0.245	0.445
Full sample	0.136	0.260	0.415
Mean dispersion	0.145	0.262	0.442
Std. dispersion	0.198	0.249	0.315
AR(2)	0.127	0.252	0.347
AR(5)	0.102	0.233	0.350

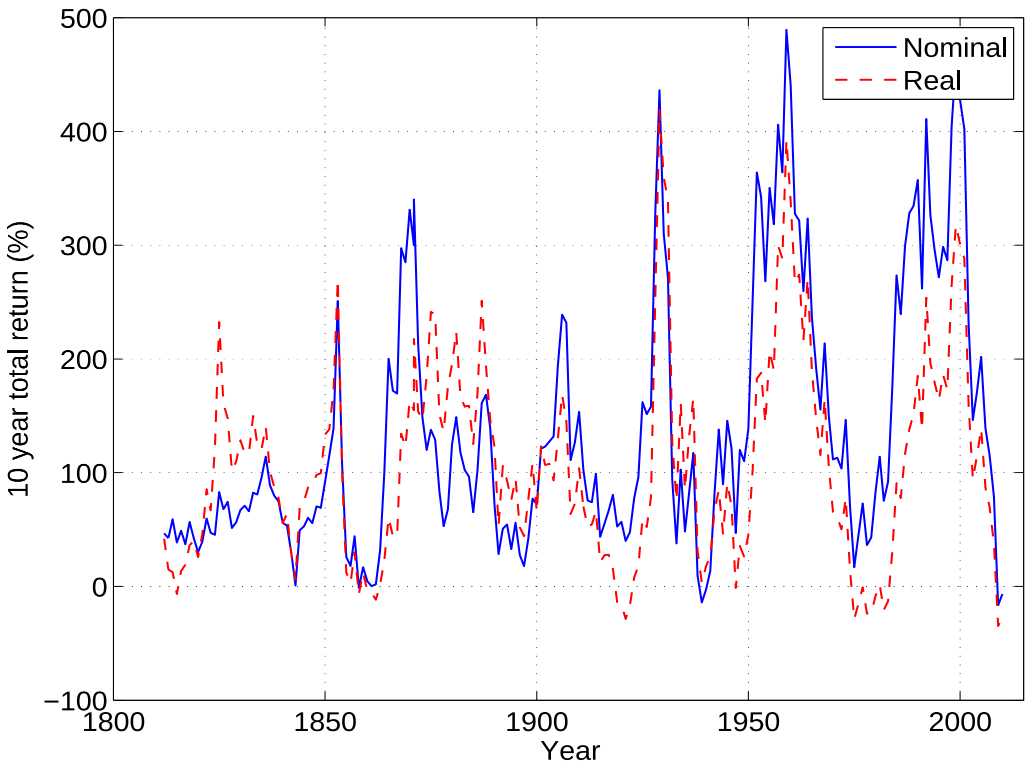

Figure 1: **Preceding Decade Returns**

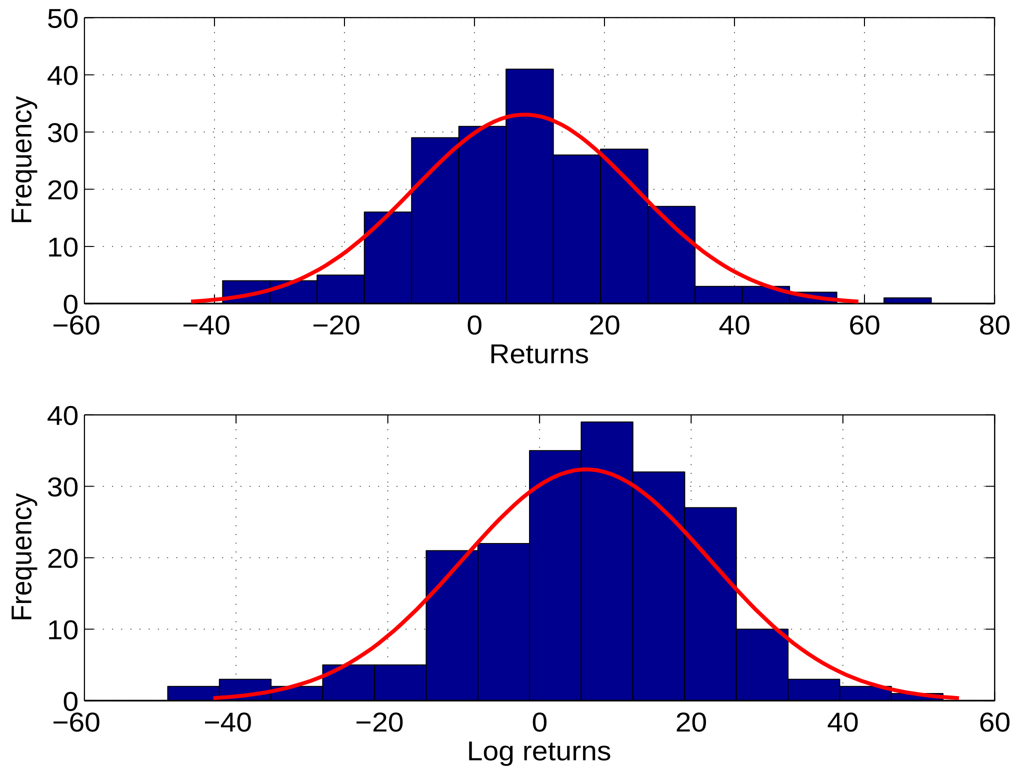

Figure 2: **Annual Return Densities**

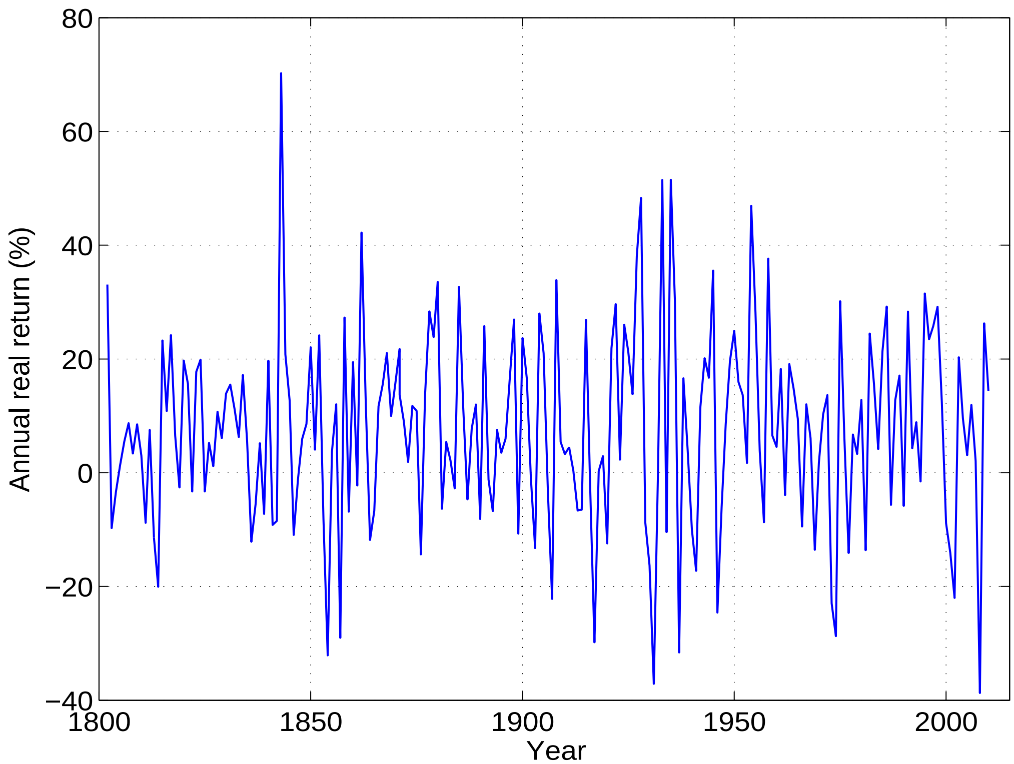

Figure 3: **Annual Returns**

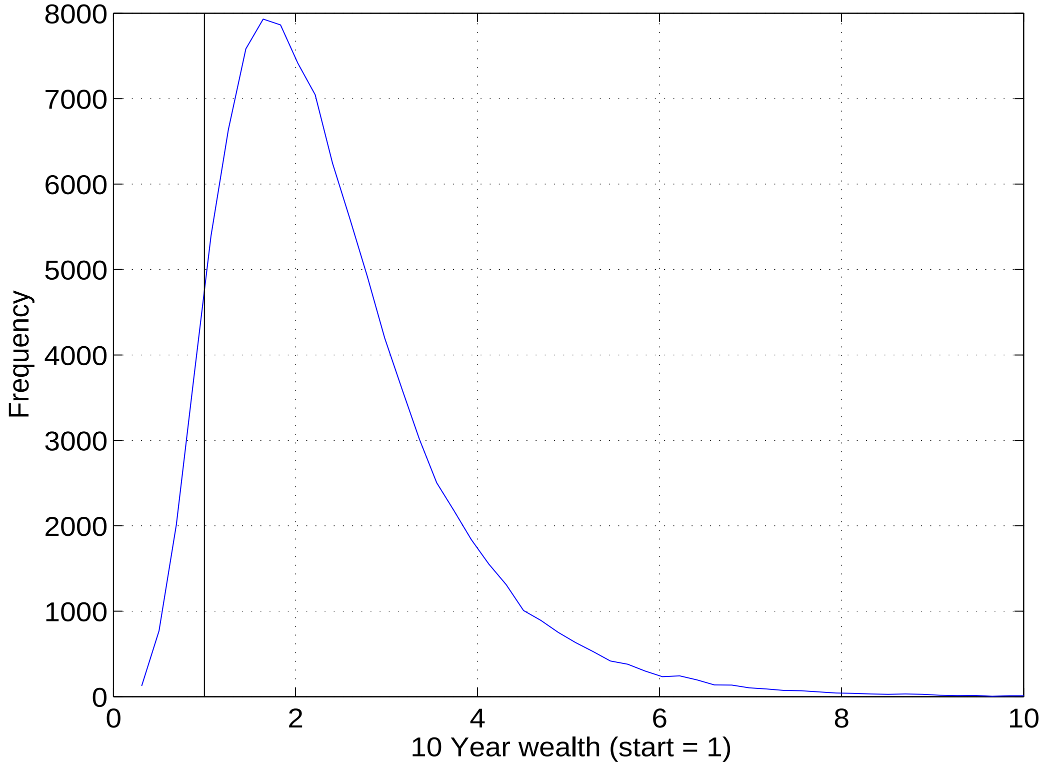

Figure 4: **10 Year Portfolio Values**

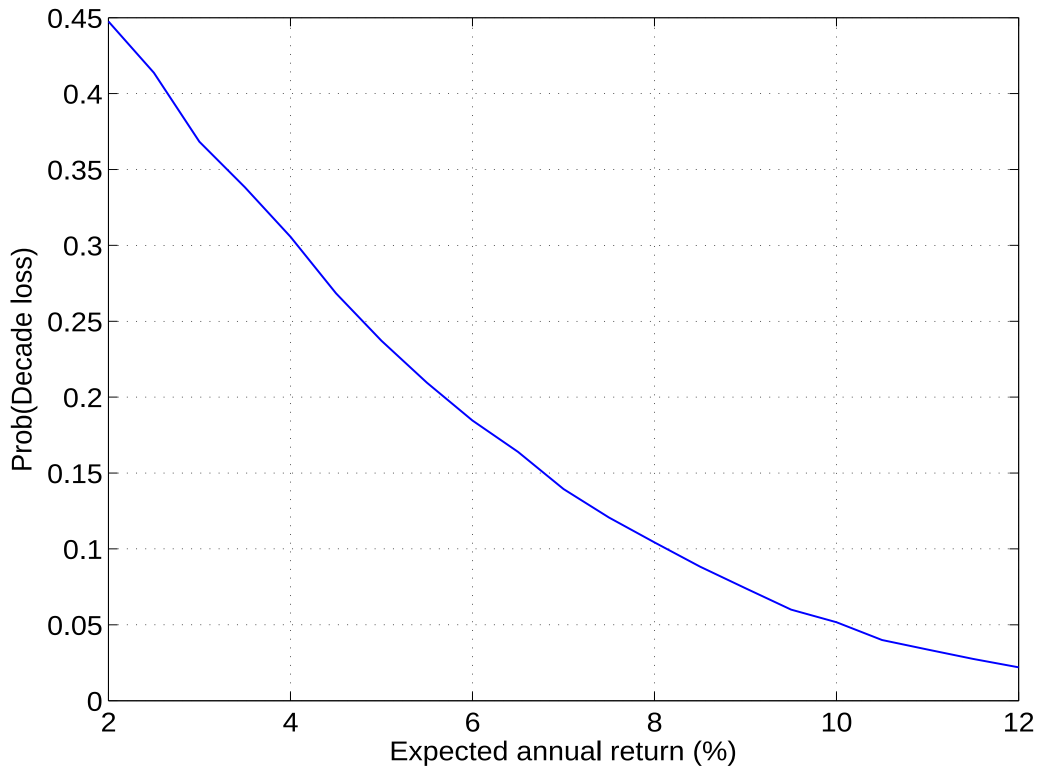

Figure 5: **Decade Loss Probabilities**

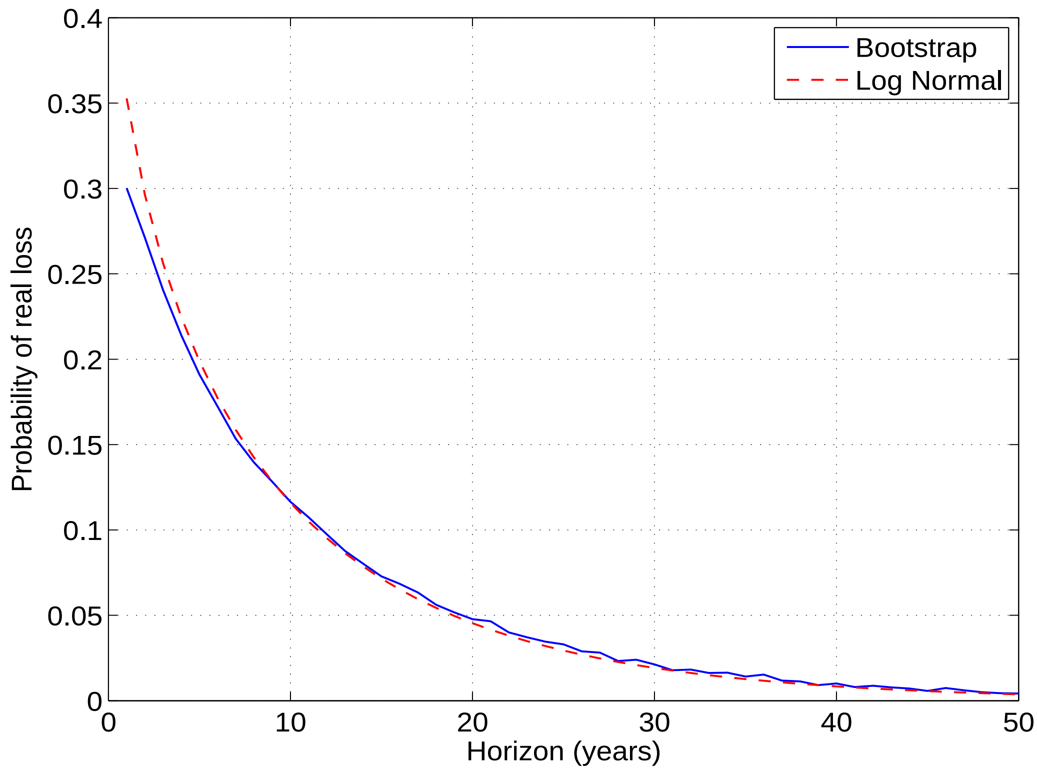

Figure 6: **Loss probabilities for increasing horizons**

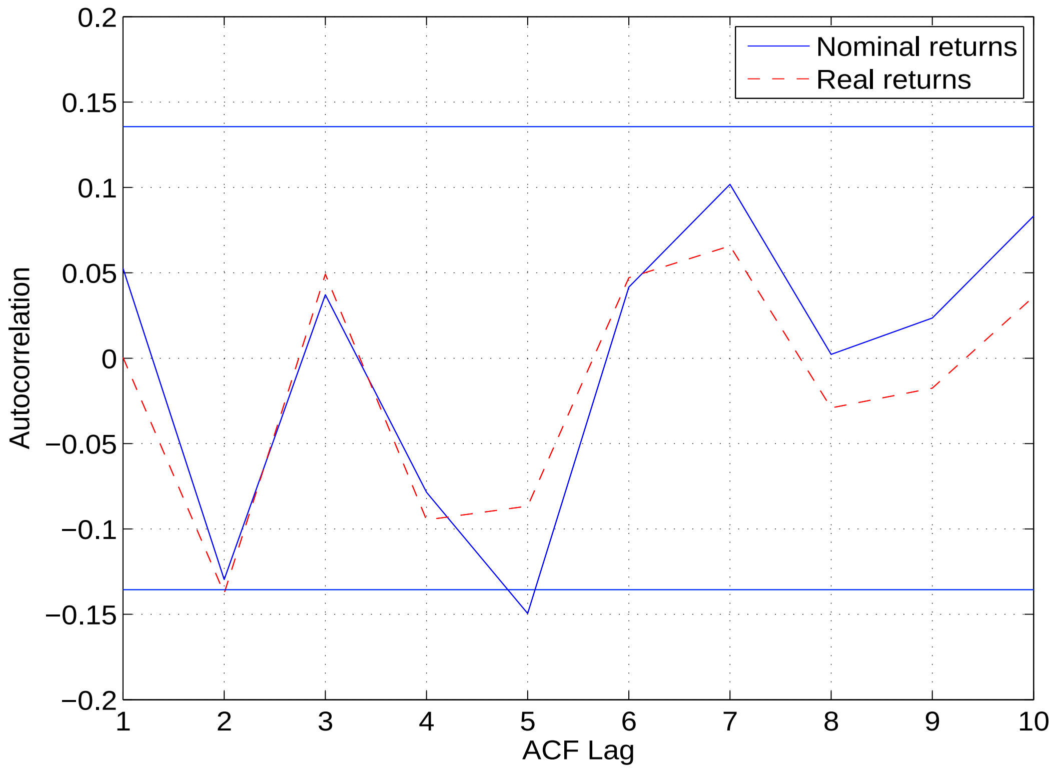

Figure 7: **Annual Return Autocorrelations**

Autocorrelations and 95 percent confidence bounds around uncorrelated null (zero).

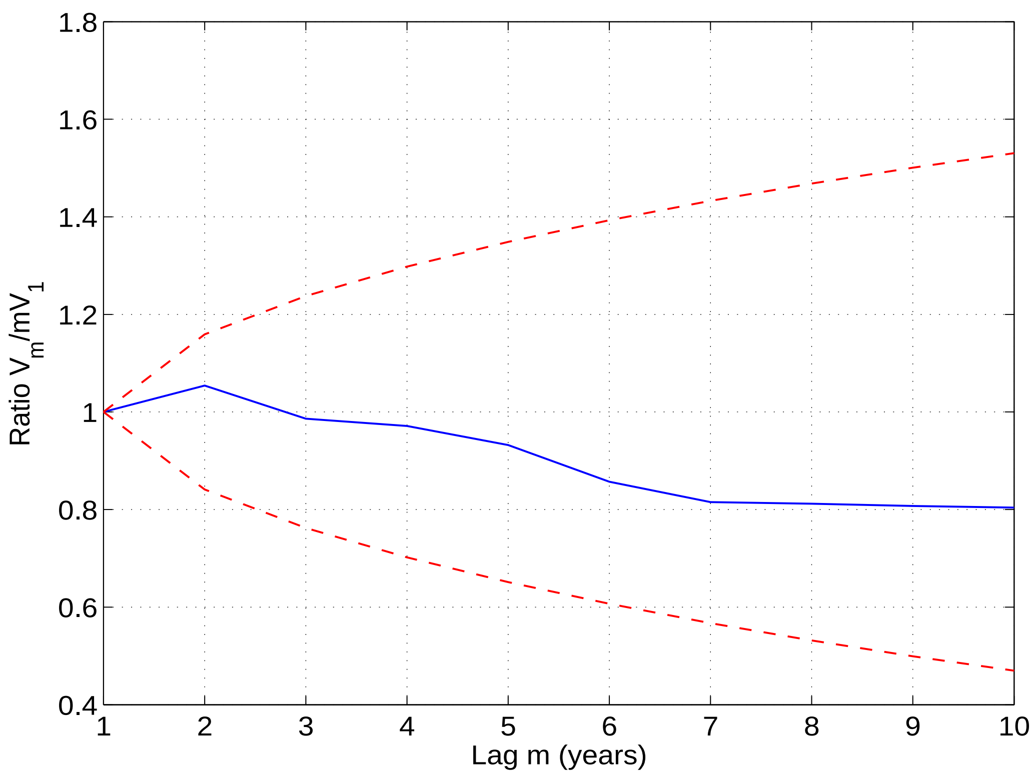

Figure 8: **Variance Ratios**

Estimated variance ratios $\frac{V_m}{mV_1}$. Estimate and asymptotic 95 percent confidence band under IID null (ratio = 1).

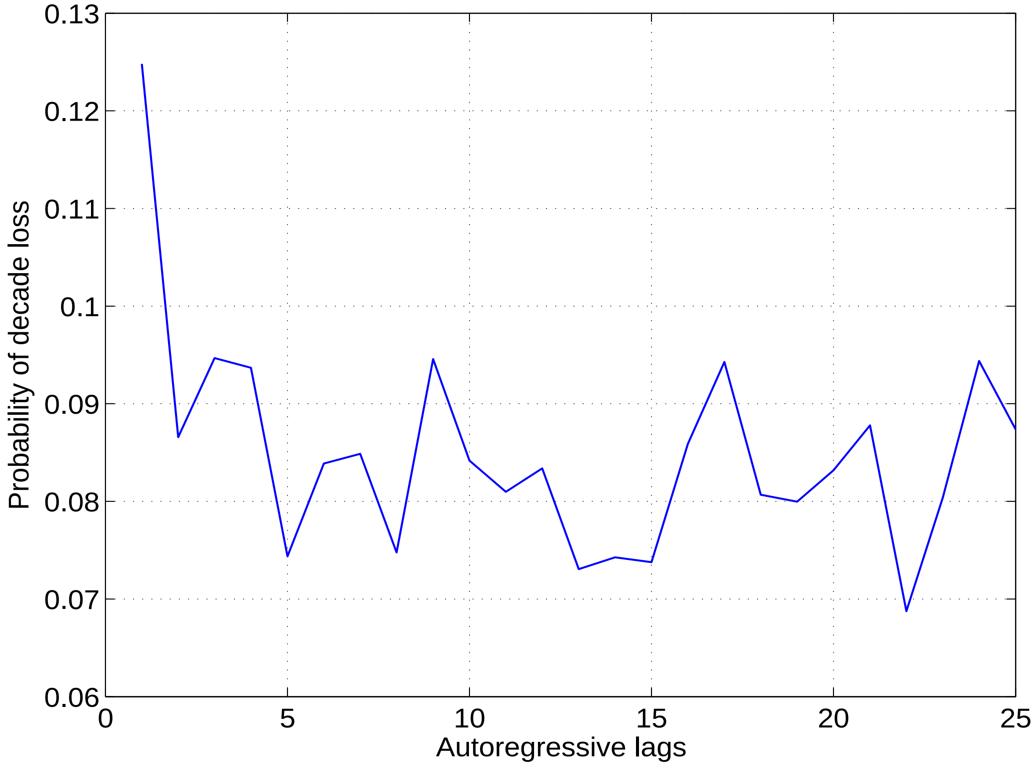

Figure 9: **Decade losses over varying autoregressive models**

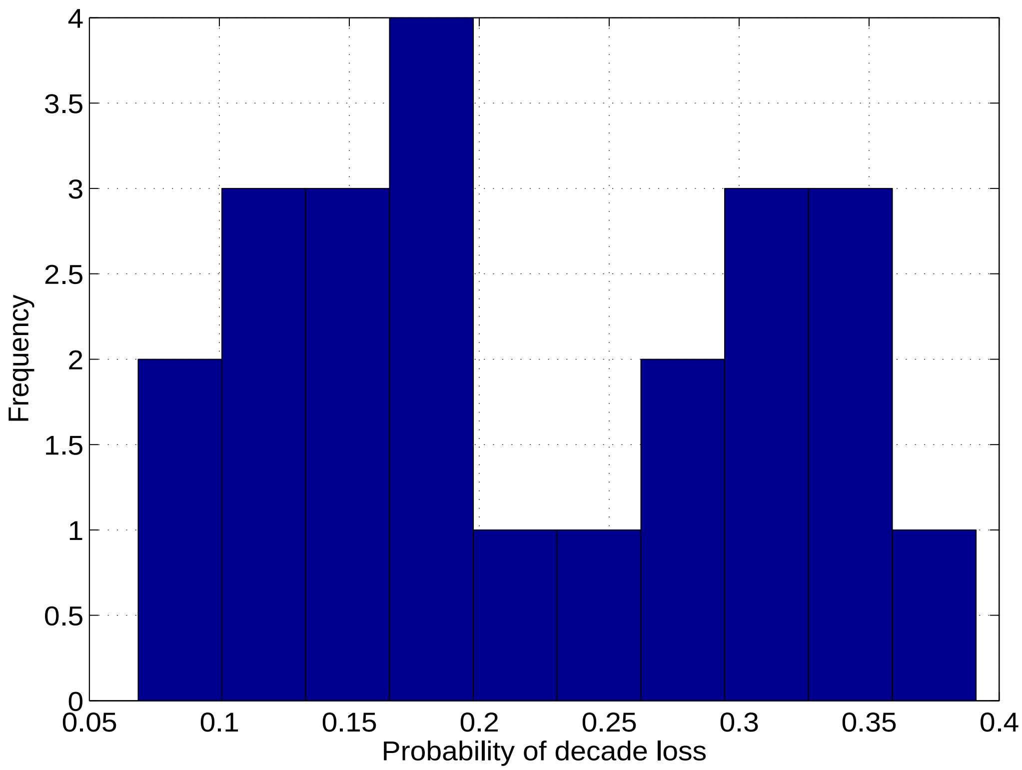

Figure 10: **Country Loss histogram**