

Simulating Alternative Endowment Benchmarks with 150 years of history.

100% Equities...again? A Discussion Paper.

Art Steinmetz

2023-07-16

Abstract

This discussion paper examines perspectives on time diversification with the goal of demonstrating that holding defensive assets in a perpetual capital portfolio, such as a college endowment, is unnecessary even with routine cash draws on the fund. This is well-researched ground, but it is worth revisiting. In this paper, we introduce two novelties. First, we look at the historical record for a much longer time period than is typical, going back to 1871 up to the present (2023), including the awful 2022 returns in both stocks and fixed income. Second, we suggest the notion of investment horizon as a declared policy parameter for defining risk acceptance by policy committees and the performance of endowment managers. We do not discuss the role of “alternatives,” but simply show that typical passive benchmarks are suboptimal..

Introduction

Way back in 1994, Thaler and Williamson (Thaler and Williamson 1994) asked “why not 100% equities?” for an endowment portfolio. The main point being that asset diversification is unimportant when the investor time horizon is sufficiently long, perpetual in the case of a college endowment. Cliff Asness (Asness 1996) critiqued this view in 1996, pointing out that the decision of which risky assets to hold and how much risk to bear are two separate decisions. Asness demonstrated that, while 100% equities beats the traditional 60/40 portfolio in absolute return over time, this allocation is far from optimal. Asness argues that simply buying the broad stock market is sub-optimal. It can be outperformed by adding diversifying asset classes *that themselves outperform*. Further, constructing a levered 60/40 portfolio will outperform a 100% stock portfolio in absolute terms while delivering lower volatility (see Appendix). In fairness to Thaler and Williamson, they don’t argue FOR the broad market as the sole asset class as

much as they argue AGAINST a large defensive position to mitigate volatility for perpetual-life endowments. Asness asks us to consider what “large” means. Also, Thaler, *et.al.* would certainly not argue against diversification of assets that enhance returns, like private equity, just against assets that are dilutive of returns, like cash, for no good purpose. Now, nearly 30 years later it is worth asking the question again, “Why not 100% equities?”

Discussions of the appropriate level of risk are often framed in terms of meeting the cash needs of an endowment sponsor during periods of steep bear markets. Certainly, there need to be enough liquid assets to meet those needs, but public equities can meet those needs. In that circumstance, there may be some aversion to the notion of being forced to sell stocks in a down market. I argue that this should not be a concern. Selling “low” feels wrong because it is a “dead weight” loss that cannot be recovered. However, this ignores the opportunity cost of not being invested in stocks in the previous period, however long you want to define the period. This is a behavioral finance problem, not an investment problem. Certainly, needed liquidity must be provided for with liquid assets. Public stock markets do this while creating more total wealth despite adverse periods.

This fear is analogous to the fallacy of “dollar-cost averaging,” where a small investor is concerned about the market declining right after their money is put to work. They are often advised to invest a portion over time. This eases the “pain of regret,” which we know is a greater motivator than the pleasure of gain, even though it will likely leave the investor with less wealth over time. (Hayley 2012)

Consider the following illustration from the “[Wall Street Journal](#)”¹. Jeremy Grantham of GMO, a respected investor, has been warning of a stock market selloff since 2015 when the S&P 500 was around 2,000. He’s been right twice in the intervening years, yet the S&P 500 is over 4,400 today, eight years later. He may be right again. Was keeping “dry powder” appropriate for a long-term investor? Ben Iker, GMO’s co-head of asset allocation, walks back his boss’ warnings: “There is this weird problem with the investment industry: Most of the reasons why people invest are long-term in nature. Retirement, pension plans, university endowments—they have very long-term goals. Yet, so many investment decisions are made in the short term.”

An endowment’s tolerance for risk should determine how much it is willing to stretch for returns beyond liquid public markets, but should not guide it to low-return investments. Alternatively, a balance between “stock return plus” and “stock return minus” assets makes sense only if the overall blended expected return does not drop below the passive benchmark. This could provide lower volatility while meeting the required return, but, again, here we argue for time diversification as a better way to think about the issue.

¹Grantham Warns AI Boom Won’t Prevent Market Bubble From Bursting, Wall Street Journal July 01, 2023 <https://www.wsj.com/articles/grantham-warns-ai-boom-wont-prevent-market-bubble-from-bursting-fd7f5a98>

This line of thinking calls into question popular endowment asset classes like hedge funds which may be dilutive of returns. The notion that hedge funds are good because they have “less risk than long-only equities” is not a good reason to hold them since they have earned a lower return. Neither is “generating alpha” unless the endowment is willing to engage in a risk parity approach which adds a substantial amount of leverage. This is a “bridge too far” for most endowments. Absolute expected return should be the starting point to evaluate a hedge fund.

How Does a “Fully Invested” Portfolio Compare?

It is not controversial to say that, over the long run, stocks have outperformed bonds. The question for endowment managers is how long they should expect to wait for adverse results to be recouped. To demonstrate the merits of a “fully invested” portfolio we look two different constructions.

1. 100% U.S. stocks.
2. 70/30% U.S. stocks and 10-year Treasury Bonds.

We choose a “70/30” portfolio as more representative of a typical endowment benchmark than the “60/40” often recommended for retail investors.² The data we use is from Robert Shiller at Yale, who has built very long time series of stock, bond, and cash returns³ going back to 1871 to support his work on cyclically-adjusted P/E ratios. This very long history includes more periods of adverse stock returns, such as the 1929 crash, than most analyses, so is more conservative in supporting a “defensive” allocation.

We will compare these two portfolios over the full history, a rolling 10-year history and the most adverse period for stocks around the crash of 1929. Then we will use a Monte-Carlo approach to sample many possible 10-year outcomes, including hypothetical endowment cash draw of both a fixed amount and fixed percentage of the endowment balance.

Assumptions for the Analysis

1. While an endowment has an “infinite” duration, the people who manage the endowment do not, so I look at rolling 10-year results as practical “long-term” horizon.
2. Quarterly rebalancing of the 70/30 portfolio.

²In the NACUBO survey of large endowment funds, the effective equity exposure, for the 11 years ending in 2019, was 72%. <https://richardmennis.com/blog/endowment-performance> (2020)

³<http://www.econ.yale.edu/~shiller/data.htm>

i Note

In the charts that follow “stock_ret” is the broad U.S. stock market, “bond_ret” is the 10-year Treasury, “bm_ret” is the 70/30 benchmark, blending stocks and 10-year Treasury notes, and “cash” is the 1-year Treasury sampled annually until 1934, when it switches to 3-month T-Bills sampled quarterly. This is how the Shiller data set is constructed.

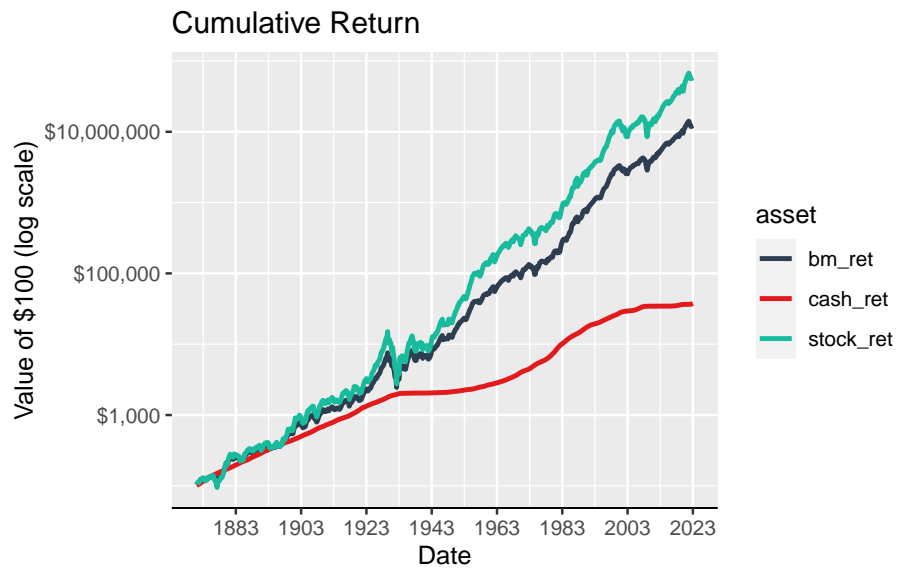
Summary Statistics 1871 Q1 to 2023 Q1

Statistic	cash_ret	bond_ret	stock_ret	cpi_growth
Avg. Annl. Return	3.96	4.78	10.67	2.17
Volatility	1.28	5.25	17.73	2.84

Correlation Matrix 1871 Q1 to 2023 Q1

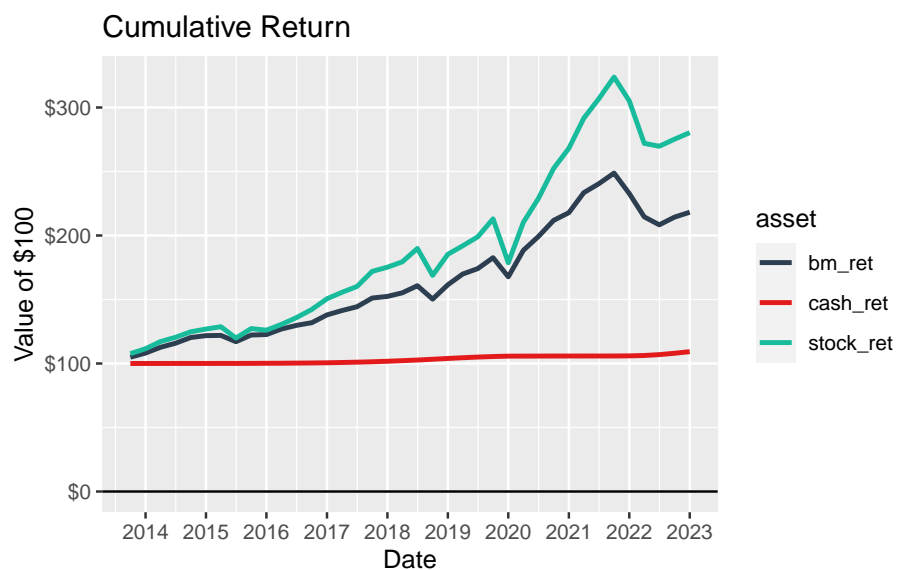
Asset	cash_ret	bond_ret	stock_ret	cpi_growth
cash_ret	1.00	0.16	-0.07	0.13
bond_ret	0.16	1.00	-0.06	-0.02
stock_ret	-0.07	-0.06	1.00	-0.05
cpi_growth	0.13	-0.02	-0.05	1.00

Looking at the full history, there are no surprises. An all stock portfolio wins, as it should. If we believe the U.S. is a going concern, if we believe the country will continue to grow and if our universities believe in their mission of turning out the next generation of people to make the country a success, then a broad investment in the equity of the nation is a good investment.

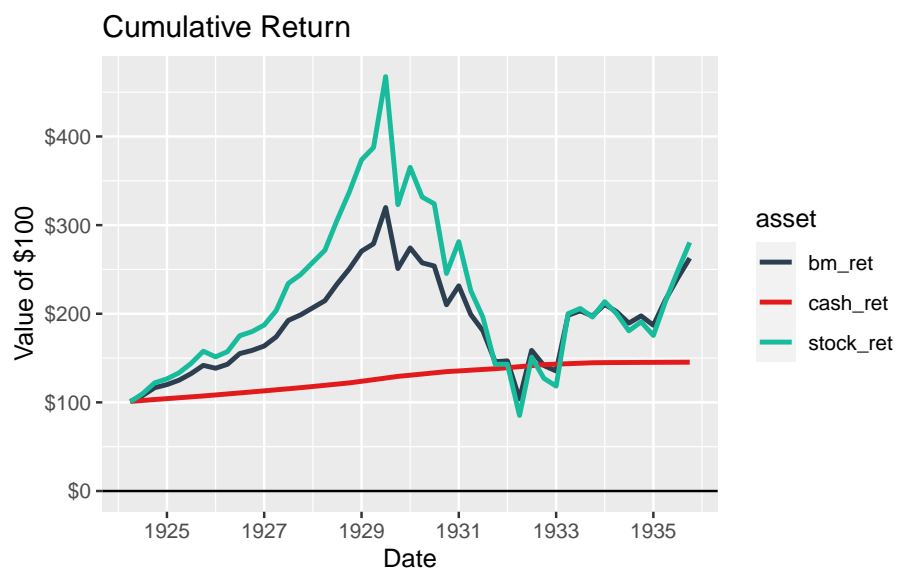


However, as J.P. Morgan blithely noted, “Markets will fluctuate.” We are more interested in how an endowment might weather the storms over shorter periods. How much courage will the managers need to face them?

2022 was a terrible year for stocks but how does the last decade look? The opportunity cost of holding defensive assets was huge. Consider the arrival of COVID. It panicked many retail investors into selling but it was a small blip on the radar even on a short horizon. The onset of Fed tightening was widely telegraphed and quite consequential for stocks, but also for bonds, which provided not much shelter. Still, an 18-month horizon, from the middle of 2021 to the end of 2022, was sufficient to make an investor indifferent among all the assets shown below, including cash.



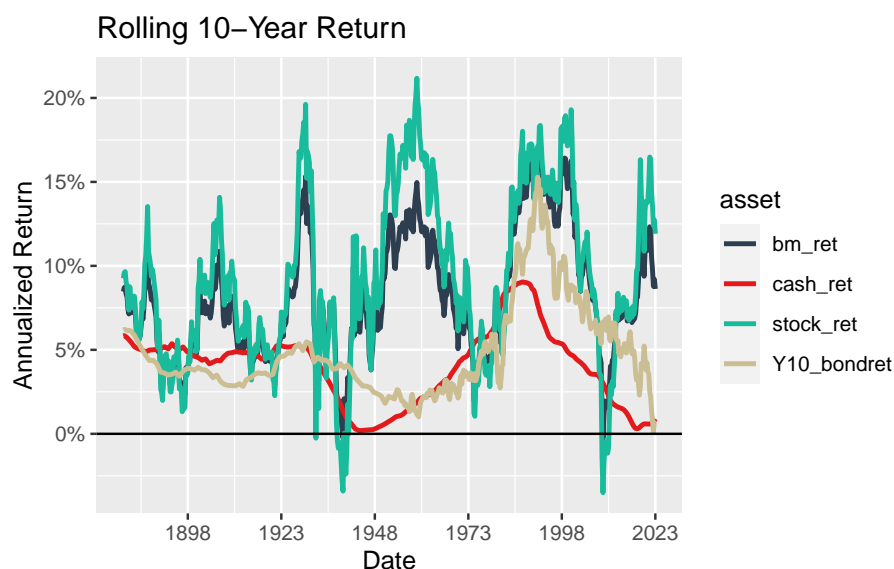
Let us now turn to a seemingly more adverse environment that still captures our imagination. We will focus on the period around the stock market crash of 1929, a terrible event for all portfolios. However, would a policy of 70/30 in stocks and bonds, instead of all stocks, have been better? Would cash have been better?



In a world of perfect foresight, cash would have been the best choice in the moment. However, over an appropriately long horizon, were defensive assets

warranted around the crash? No. Would the trustees of an endowment have been able to provide the necessary political capital to avoid firing the managers? Maybe not. But trustees who spent the decade from 1925 to 1935 in a deep sleep, like Rip Van Winkle, would have been perfectly satisfied with the results of the portfolio.

There have certainly been 10-year windows when a fully invested portfolio was not the best choice. How often and how long do they last? The chart below gives you a sense of the range of actual historical rolling 10-year returns for each portfolio. We can make a few observations. It is rare, but more often than not, for stocks to have a negative return for a 10-year horizon. There has never been a negative 10-year stretch for the blended benchmark. As an aside, the trailing 10-year return for bonds in 2022 Q3 was the worst in our history, but still slightly positive.



In essence, there have been three periods where a stocks-only portfolio materially underperformed a 70/30 portfolio on a 10-year basis: during the Great Depression, the Great Recession, and the worst episode, the policy-induced recession of 1937-1938. In the first two instances, the rolling return soon favored the all-stock portfolio again, even against cash.

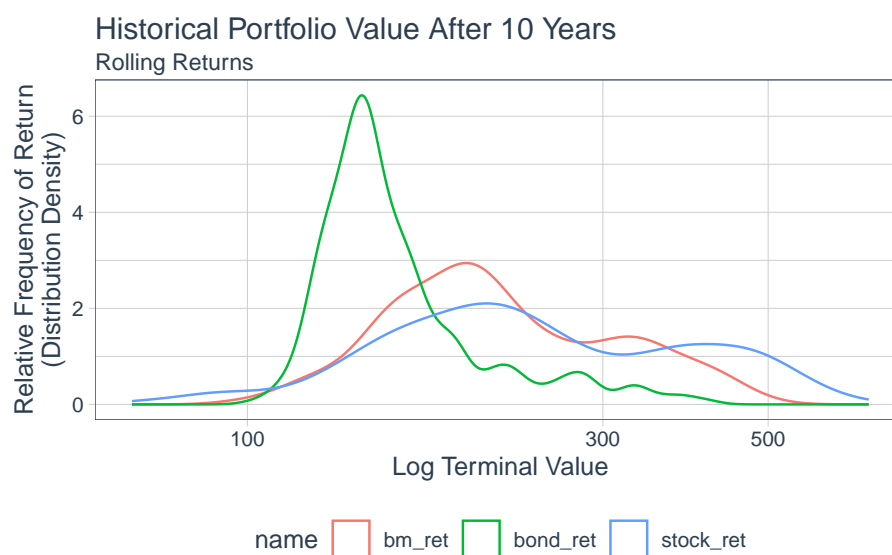
In the last instance, it took five years before a manager could say that omitting a defensive component was the better choice for the prior ten years. It is now widely agreed that the “[recession within the depression](#)” (Waiwood, n.d.) was caused by an unfortunate confluence of three policy changes; the Fed tightened policy to sop up “excess” bank reserves despite shaky growth, the Treasury decided to sterilize the gold inflows Europe was sending to buy arms and the Social Security payroll tax was commenced (Bordo and Haubrich 2012). Hope-

fully, policy makers have learned some lessons from that episode.

A more sobering view considers the cumulative value of a portfolio over a rolling 10-year window. The range of outcomes has been quite large.

Rolling 10-Year Portfolio Value
Value of \$100 After 10 Years

statistic	cash_ret	stock_ret	bond_ret	bm_ret
Min.	102	70	101	97
1st Qu.	124	174	135	169
Median	156	227	146	205
Mean	151	268	165	232
3rd Qu.	165	360	170	281
Max.	238	682	415	500



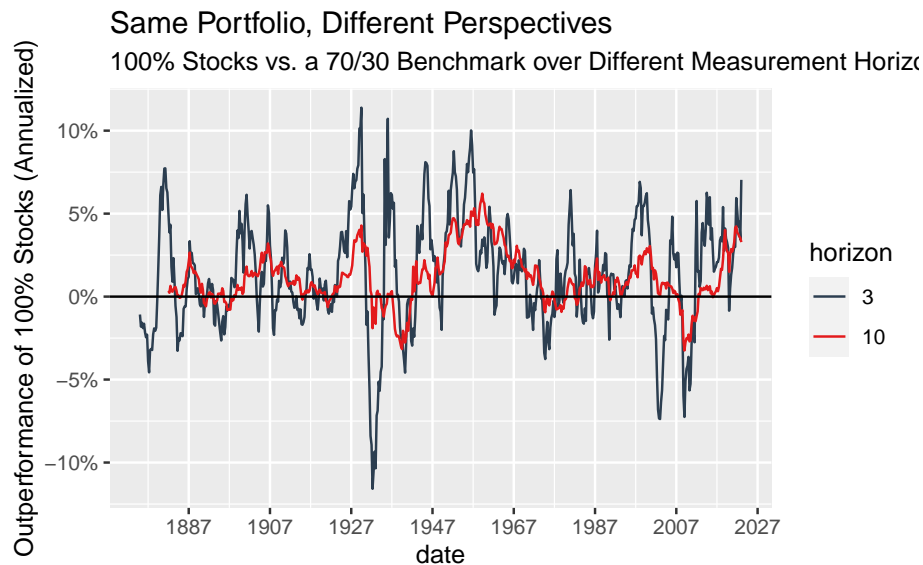
Nevertheless, the stock-only portfolio outperformed the 70/30 benchmark 81% of the time over a rolling 10-year horizon over our history.

Two managers walk into a bar...

One says to the other, “I beat my 70/30 benchmark 65% of the time.” The other says “I beat my 70/30 benchmark 95% of the time.” Then they compare holdings over time. They are identical. How can this be?

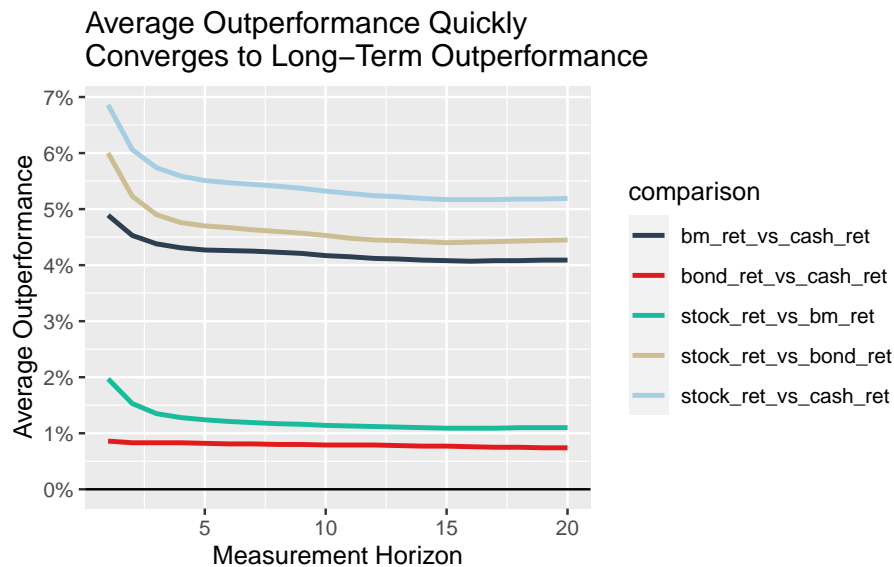
For an asset that trends upwards, we can be more certain of a positive outcome on a longer horizon. On a shorter horizon, there will be more negative return periods. Upon reflection, this is obvious but it can color our thinking about risk. When we casually speak of market “volatility” we often mean “loss.” Nobody complains about upside risk. In that framework, diversification usually means adding assets to blunt potential loss. Alternatively, we can think about risk as a consequence of time, rather than just drawdown, where time itself is a risk mitigant. If you accept *as a given* that public stock markets always recoup losses eventually, then time is the variable that should define risk tolerance within any given portfolio. The literature on time diversification is extensive and uncontroversial so it is paradoxical that “defensive” assets continue to be held in endowments with essentially perpetual horizons. (William Reichenstein and Dorset 1995) The fact that continual cash drawdowns are required does not change this. Suppose the average endowment draws 5% of the principal annually. That implies a horizon of 20 years, making the likelihood that defensive assets enhance the wealth of the endowment very small.

To examine the consequences of different horizons we will compare the annualized return differences across multiple horizons among pairs of portfolios, stocks, bonds, the 70/30 benchmark and cash. Below we show one pair, stocks vs. the 70/30 benchmark, for three-year and ten-year horizons. Note the increased frequency of underperformance of pure stocks for the three-year horizon with a larger magnitude of deviation. We are interested in measuring the average excess return for stocks and the fraction of time stocks are the better investment for the multiple horizons to understand how we perceive performance.



Below we can see that average outperformance converges to the long horizon result after about three years, so horizon doesn’t have a big effect on AVERAGE

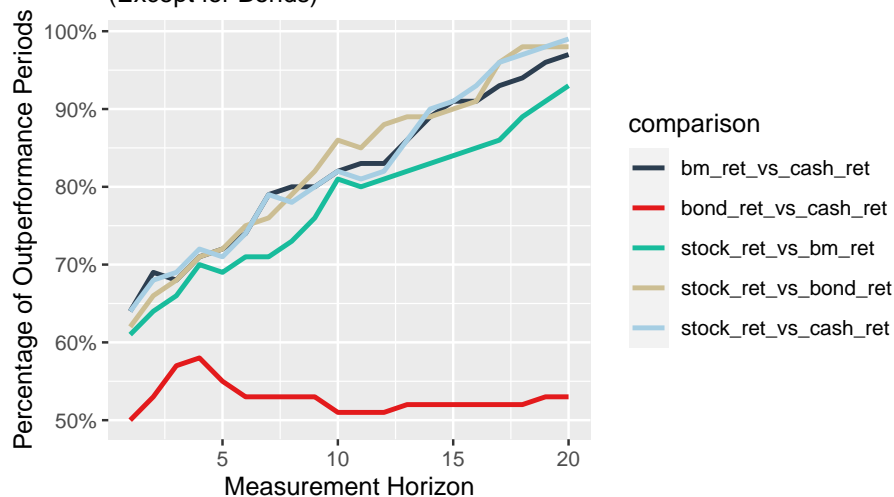
return differences among these portfolios. Over the long term, being measured on 3-year performance or 10-year shouldn't make much difference IF there are sufficiently many performance measurement snapshots. Over the career of an endowment manager there may not be enough snapshots to show this.



The results get more interesting when we examine the frequency of outperformance periods over different horizons. It is not surprising to see that longer horizons mean more certain outperformance for most asset pairs. This does illustrate how powerful the effect is. It may be surprising to see the effect is linear.

It is even more surprising to those of us whose careers spanned a multi-decade bull market in bonds, to see that a longer measurement horizon does not improve the perceived performance of bonds. Stock returns are distributed normally with an upward trend bias over time. Bonds do not share this happy circumstance, further eroding the case for fixed income as a long-term hold in an endowment portfolio.

A Longer Horizon Means More Certain Outperformance (Except for Bonds)



A Monte Carlo Approach

So far we have been making observations about the past. Now let's imagine the future. Consider an endowment which starts with \$100mm. A new endowment manager is hired with a 10-year contract (not realistic) and gets measured after 10 years with the goal of maximizing the size of the endowment. We test terminal endowment wealth under two cash-flow scenarios.

1. Level pay adjusted for inflation. The endowment starts out paying \$5mm and receiving contributions of \$1mm once a year for a net of \$4mm. This amount grows with inflation but is otherwise fixed, on the notion the college spends the same every year, irrespective of market gyrations.
2. The net cash flow starts at \$4mm but the payout is fixed at 5% of the endowment value. The payout shrinks in down markets and we also assume donor contributions are a function of the market as well. Inflation is not a factor.

To create possible future return paths we draw returns at random from the historical series. We pull 40 quarters (10 years) “with replacement” and compute the terminal value of the endowment using the two cash draw scenarios.

We then run 1000 trials and compare the terminal portfolio values.

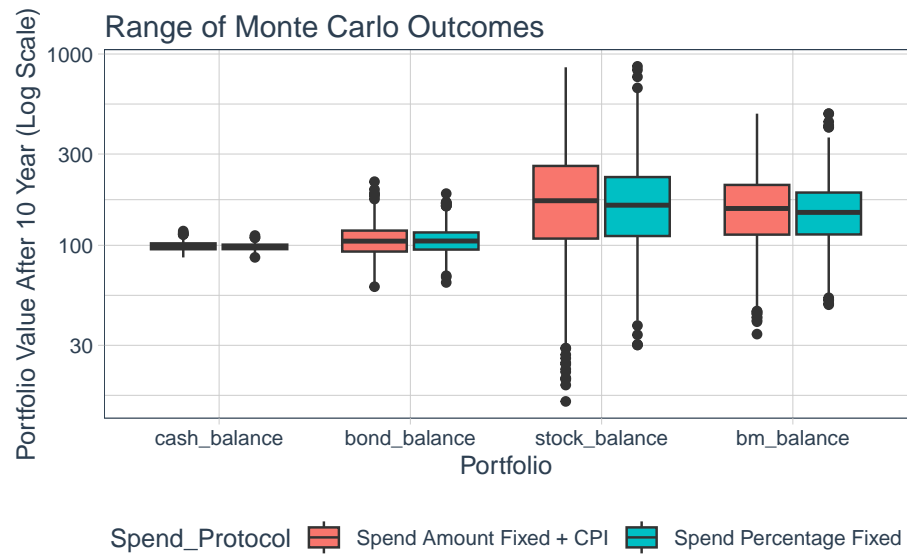
Simulated Endowment Value After 10 Years

Using Alternative Spend Protocols

Statistic	cash	bond	stock	bm
Spend Amount Fixed + CPI ¹				
Min	86	61	15	34
Bottom_Quartile	95	93	108	114
Median	99	105	171	156
Top_Quartile	102	119	260	207
Max	119	215	853	488
Spend Percentage Fixed ²				
Min	86	64	30	49
Bottom_Quartile	95	95	112	114
Median	98	105	162	149
Top_Quartile	101	117	227	189
Max	113	186	862	488

¹Assuming Fixed Payout Adjusted For Inflation

²Assuming Payout as a Fixed Percentage of Endowment Balance



"Batting Averages"

Spend_Protocol	stock_beats_bm	stock_under_par	bm_under_par
Spend Amount Fixed + CPI	68.5%	21.9%	16.5%
Spend Percentage Fixed	71.5%	19.4%	15.8%

The first thing to notice is the Monte Carlo results are worse than the historical returns. The historical returns are rolling 10-year periods so there is a lot of serial correlation and the upward bias over time is reinforced. Each 10-year window in the Monte Carlo test is independent so the median is more realistic. On the other hand, the worst case and best case tails are less realistic. In a random draw over 1000 trials some of the 10-year spans are bound to be an unbroken series of negative or positive returns. This is far less likely to be true than our model allows.

The superior result from "level-spend" means that the ability to spend a smaller fraction of the endowment when the market is up outweighs the burden of spending more when the market is down. This is the opposite of what a dollar-cost-averaging strategy would suggest.

While the median outcome for each spending protocol is different, the "batting averages" are very similar. The stock-only portfolios outperform the 70/30 benchmark 68% of the time over a 10-year horizon. The probability that the stock-only portfolio winds up below its initial value is 22%.

Conclusion

If we look at the characteristics of the return distribution we sampled from, we see that bonds have, as expected, lower volatility and essentially zero correlation to stocks. Is that enough to make up for half the annualized return? It's worth noting that over the last twenty years the negative correlation between bonds and stocks was much stronger, -0.44 vs. -0.06 for the full period. What you think this relationship will be going forward has a lot to do with what bonds might contribute to an endowment portfolio. With such low real yields today there is not a lot of cushion for bonds to move in the opposite direction of stocks.

More importantly, though, is the simple truth that, if the expected return of a portfolio is less than the required drawdown of cash from the endowment, the portfolio is unsustainable. You may observe that, in an efficient frontier context, it is possible to add a low return asset and get a better risk-reward ratio. Unfortunately, the optimal portfolio blend will have a return below the all-stock portfolio and possibly below the required return. Yes, leverage can be used to increase the return to an arbitrary level. This is Asness' point but, again, few endowments would accept so much borrowing. A "risk parity" portfolio with a 70/30 base would require about 40% leverage, assuming historical returns and

borrowing at the risk-free rate (see appendix). Of course, many hedge funds play exactly this game. Whether they succeed is a separate question.

We have shown that a passive benchmark containing “defensive” assets derives from placing too much emphasis on short-term loss potential. Taking a longer measurement horizon can allay those fears. Adding diversifying assets that put a drag on return and lower the expected future wealth of the endowment should be avoided. Especially problematic are those which have an expected return at or below the rate of cash draw, such as Treasury debt. There is certainly merit to diversifying an investment portfolio provided the alternative assets add to the expected return. If less liquid, high expected return assets are used to enhance the portfolio then two questions for endowment managers to ask are:

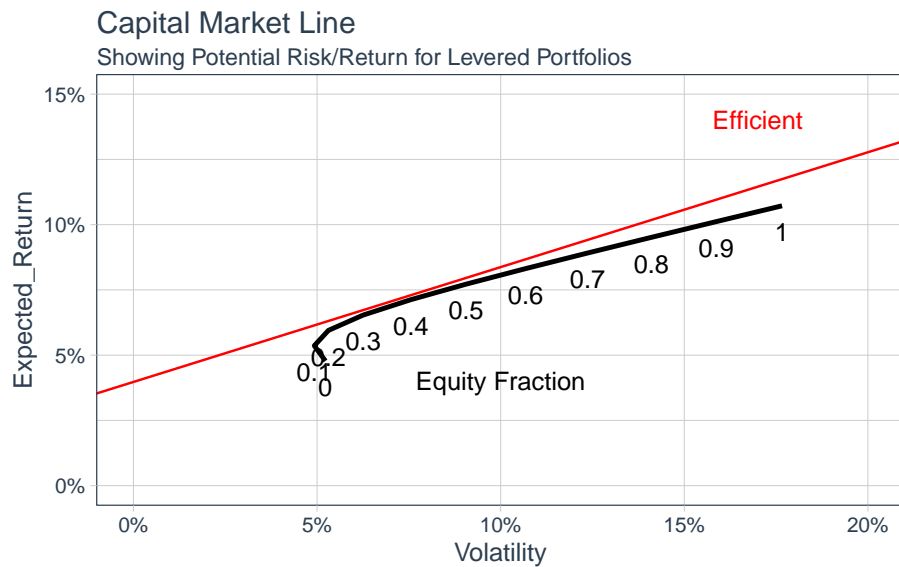
1. How much risk do return-enhancing investments add - at the portfolio level, taking correlation into account?
2. How much liquidity is needed to meet any possible immediate cash need of the sponsor, liquidity that can be provided by publicly traded stocks.

The answer to the question of how much to hold in “defensive” assets is the same as it was thirty years ago, zero.

Appendix

A Risk Parity Portfolio

Below we see the efficient frontier using the entire history for various mixes of stocks and bonds. Given cash returns the “optimal” portfolio would have been a mix of about 40% stocks, 60% bonds. A risk parity portfolio to match even the volatility of a 70/30 portfolio would require an enormous amount of leverage. This is due to the sad fact that 10-year T-Notes have scarcely returned above short term T-Bills over the full history.

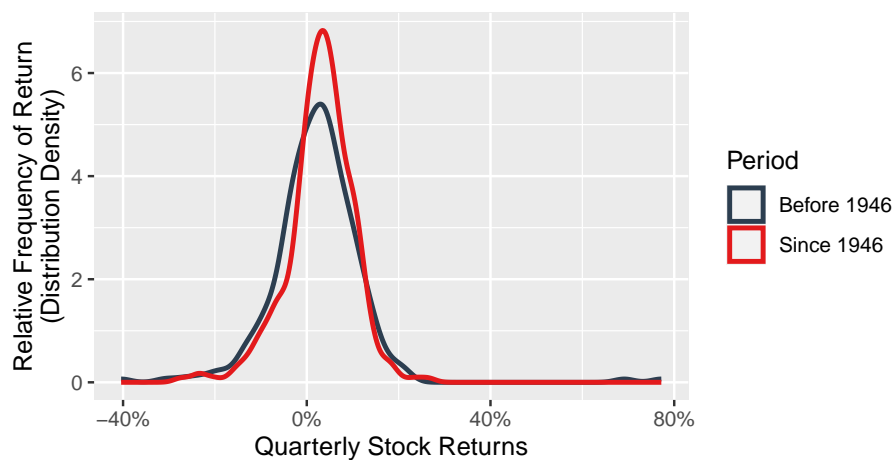


A Note About the Long History of Data

As noted earlier, we use a longer history than a typical analysis of this type. Often they confine their history to the “post-war” period or shorter. This makes our results more conservative because poor stock and bond markets were slightly more common earlier and volatility was higher. Below we can see that the distribution of outcomes for stocks shifts to the right (better) after 1946.

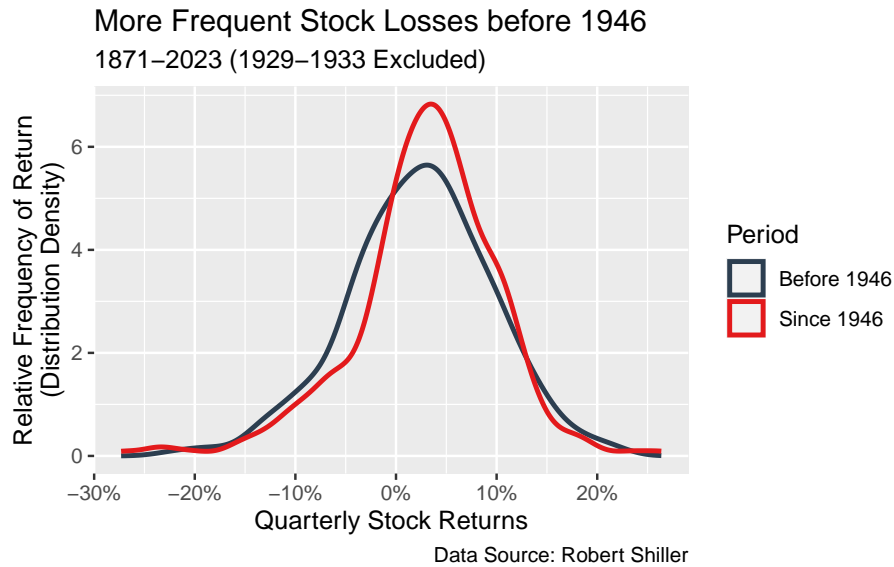
More Frequent Stock Losses before 1946

1871–2023



Data Source: Robert Shiller

If we filter out the wild gyrations around “The Crash” from 1929 to 1933, the results do not change.



Is the more recent period “normal” for the future? We make no assertions about the answer.

References

- Asness, Clifford S. 1996. “Why Not 100 Percent Equities.” *The Journal of Portfolio Management* 22 (2): 29–34. <https://doi.org/10.3905/jpm.1996.29>.
- Bordo, Michael, and Joseph Haubrich. 2012. “Deep Recessions, Fast Recoveries, and Financial Crises: Evidence from the American Record.” <https://doi.org/10.3386/w18194>.
- Hayley, Simon. 2012. “Dollar Cost Averaging - The Role of Cognitive Error.” *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1473046>.
- Thaler, Richard H., and J. Peter Williamson. 1994. “College and University Endowment Funds.” *The Journal of Portfolio Management* 21 (1): 27–37. <https://doi.org/10.3905/jpm.1994.409492>.
- Waiwood, Patricia. n.d. “Recession of 1937–38.” *Federal Reserve History*. <https://www.federalreservehistory.org/essays/recession-of-1937-38>.
- William Reichenstein, CFA, and Dovalee Dorset. 1995. “Time Diversification Revisited.” *CFA Research Foundation Books* 1995, Issue 2 (February). <https://doi.org/10.2470/rf.v1995.n2.4451>.