Volatility Management in Defined-Benefit Pension Plans: Basic Optimization

BY ROBERT T. MCCRORY, FSA
EFI ACTUARIES
1532 EAST MCGRAW ST.
SEATTLE, WA 98112

Abstract

The purpose of this paper is to explore a methodology for choosing among techniques for managing contribution volatility in defined-benefit pension plans. A frequently used—and as often criticized—method for managing contribution volatility is to base actuarial cost on smoothed assets. Our goal in this paper is to develop a methodology for measuring and evaluating the quantitative impact of competing asset smoothing policies. We proceed by investigating a series of basic pension metrics, evaluating the characteristics of each using a simulated pension plan. We then use normalized and weighted combinations of these metrics to determine an optimal smoothing policy.

1. Introduction

Optimization is in the eye of the beholder. This has been said by many people in many different contexts, but it is never said often enough. Any attempt to find an optimal solution to any problem must always involve choosing criteria to measure the degree of success achieved, and the selection of such criteria is always—100 percent of the time—an emotional decision.

So, what do we do? Decisions must be made; policies put in place, implemented, and managed. We can't give up. But what we can do is to experiment, selecting some measurements that we think may be useful and evaluating the performance of competing policies based on these measurements. That is the goal of this paper, in which the competing policies will be concerned with asset smoothing.

We will proceed as follows:

- Problem Definition

  We will define the problem and try to determine the Policy Space we are interested in; this Policy Space is the range of potential solutions to our problem that we will investigate, measure and assess.

- Methodology

  We will describe the methodology we will employ in our investigation. As in previous papers I have written, the approach will be centered on simulation experiments conducted on a Model Plan. This plan will be described.

- Exploration

  Some proposed metrics will be tested next, using a few sample policies from the Policy Space. Our goal is to find measurements of policy effectiveness that can be used to characterize and score competing policies.

- Optimization

  As noted above, all optimization is based on the selection and weighting of criteria, or performance metrics. We will define optimal in several different ways, each definition based on a set of metrics and a set of weights for these metrics. We will then score policy alternatives to determine which is optimal under the criteria selected.

In summary, we intend to demonstrate a method of discovering optimal solutions to pension plan problems. Our demonstration will be confined to limiting cost volatility through asset smoothing, but the method can be extended and applied to almost any policy decision confronting actuaries, plan administrators, plan sponsors and plan fiduciaries.

1.1 Problem Definition

Since private and public pension plans began investing in risky assets, actuaries have developed techniques to smooth the impact of investment market fluctuations on required plan contributions. A number of approaches have been developed and applied over the decades, but the one most frequently used spreads investment gains and losses over a number of years. In essence, the annual actuarial contribution is based on an actuarial, or smoothed, value of assets that may differ significantly from market value. These methods are the subject of Actuarial Standard of Practice (ASOP) 44.

There are a wide variety of approaches to asset smoothing, but for this analysis we will focus on methods that are based on market value and that recognize investment gains and losses gradually. Gradual recognition will be achieved by spreading the gain or loss over a specified number of years, or by dividing the difference between actual and expected market value by a factor. Three variables will be employed in this analysis:

- Will the smoothing algorithm be year-based or factor-based?
- How many years or how large a factor will be used?
- Will there be a corridor around market value within which the smoothed value is constrained, and how large will the corridor be?

To proceed we will identify a Policy Space with the following dimensions:

- Both year-based and factor-based algorithms will be considered (2).
- In the year-based approach, investment gains and losses will be recognized over from zero (assets are at market) to 25 years in steps of one year; in the factor approach, factors from 1.0 to 25.0 will be used, in steps of 1.0 (25).
- A corridor around market value will be enforced, where the corridor ranges from 0 percent (assets are at market) to 50 percent in steps of 2 percent (26).

This gives us a Policy Space of about 2 × 25 × 26, or 1,300 points. That's a lot of computing.

While it may be a lot of computing, it is not exhaustive. There are many, many more smoothing techniques that could be considered. However, the goal here is to demonstrate a methodology on a reasonable Policy Space and possibly to learn a little about asset smoothing as we do so.

1.2 Methodology

We will explore policy metrics and optimization using the Model Plan and Model Economy introduced in a prior paper. The Model Plan is highly simplified, with one age at hire, a single retirement age and a stable population, with new members replacing those who retire or die. The Model Economy features independent normally distributed asset returns and inflation.

In Graphs 1 and 2 below we reproduce graphs of the simulated cost and funded ratio of the Model Plan operating in the Model Economy, principally as a point of reference for the rest of this paper. We summarize below a few of the salient behavioral features of the Model Plan in Figures 1 and 2:

- Despite the fact that actuarial assumptions are met on average, we see in Graph 1 that the average cost stays roughly level for 10 years or so, and then gradually decreases. The median cost begins declining sooner, and we reach a point at about 80 years at which more than half of the trials are producing no cost at all. Note that at any point in time, the 75th percentile is about twice the average cost.

- The pattern of decreasing mean and median actuarial cost exists because of the effect of the Exclusive Benefit Rule of the Internal Revenue Code. The Exclusive Benefit Rule allows withdrawals from a pension plan only for the "exclusive benefit" of the plan's members and beneficiaries. Refunds to employers in the event the plan becomes overfunded are forbidden.

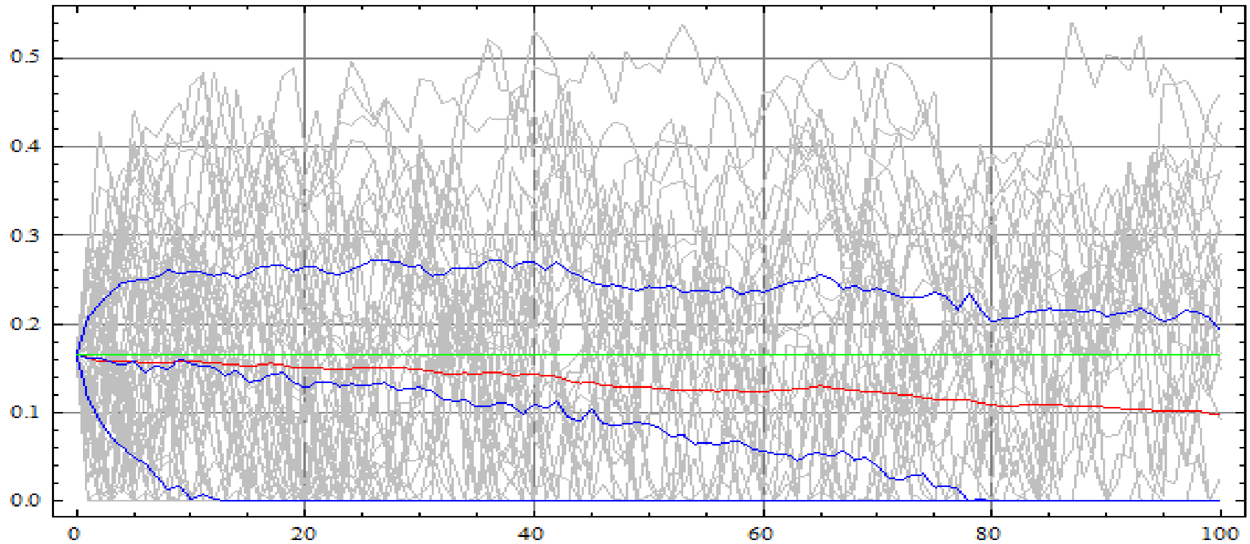
  The Exclusive Benefit Rule means that contributions to the plan must be made in times of poor returns and underfunding, but in times of good returns and overfunding money cannot be removed from the plan. The result is that the contribution pattern is asymmetrical, and overfunding is compounded rather than corrected during a series of favorable returns; on average this drives the mean and median cost down over time.

- Note the wide range of funded ratios in Figure 2: Many of the simulation trials soar well over 700 percent funding, some getting there quite quickly. As a result of such runaway overfunding, the mean funded ratio exceeds the top quartile of results within 50 years. Since funds cannot be withdrawn from the plan except in the form of benefits and expenses (the Exclusive Benefit Rule again), in good trials with high returns the plan becomes overfunded and assets continue to grow, compounding without limit. The median funded ratio also increases, but is less affected than the mean by the extreme values that arise from overfunding.

- Note in Figure 2 that there is an effective floor of about 40 to 50 percent for the funded ratio. At this low level of funding, the actuarial cost is so high that it is sufficient by itself to offset benefit payments plus any possible loss on the remaining plan assets. Under these circumstances, assets are bound to increase and fund insolvency is impossible—as long as the required actuarial contributions are made on time. A similar minimum funded ratio is observed in simulations of actual plans, so it is not an artifact of the Model Plan being used here.

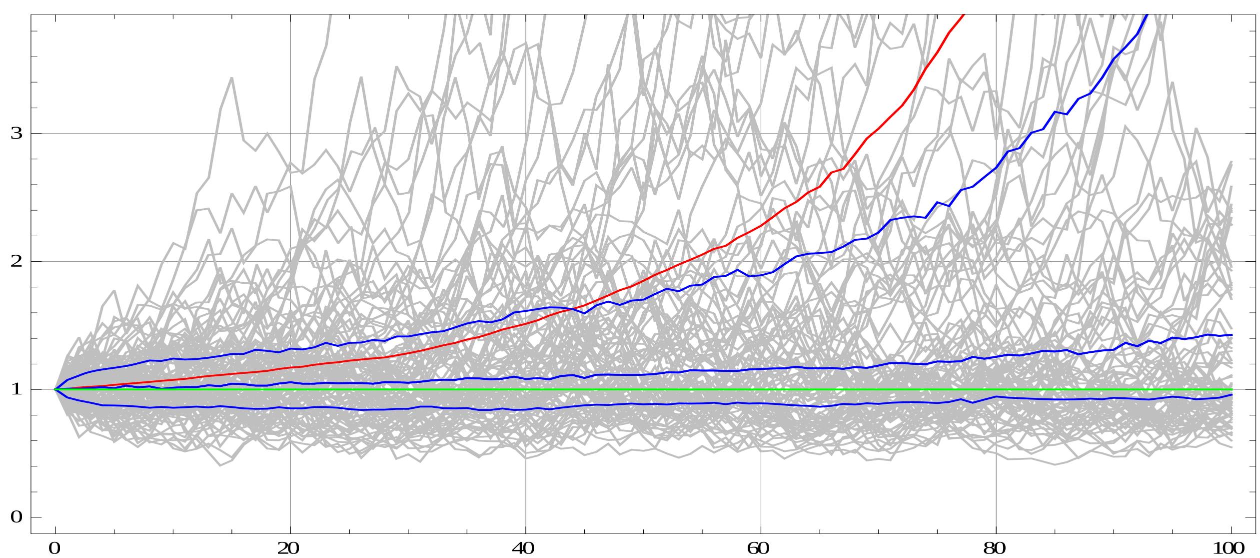
- As a consequence of runaway overfunding, the distribution of funded ratios at any time is strongly positively skewed: The funded ratios can equal or exceed 1,000 percent in some trials.

While simplified, the Model Plan and Model Economy deliver rich behavior similar to that I have observed in my clients' plans. Your mileage may differ, but this is a place to start.



Year	0	25	50	75	100
Actuarial Projection	0.165275	0.165275	0.165275	0.165275	0.165275
Mean	0.165275	0.150708	0.128161	0.114998	0.0974617
SD	0.0	0.143748	0.140221	0.139323	0.134221
75th Percentile	0.165275	0.266623	0.241916	0.235827	0.192109
Median	0.165275	0.13047	0.086934	0.015328	0.
25th Percentile	0.165275	0.	0.	0.	0.

Figure 1: Simulated Actuarial Cost of the Model Plan with Default Assumptions
(Each trial is shown as a gray line; the actuarial projection is in green; the mean is in red; and the median and quartiles are blue.)



Year	0	25	50	75	100
Actuarial Projection	1.	1.	1.	1.	1.
Mean	1.	1.22554	1.84974	3.63337	9.48764
SD	0.	0.648464	2.29238	8.29702	30.6395
75th Percentile	1.	1.36507	1.70268	2.46213	5.15995
Median	1.	1.05067	1.11554	1.22023	1.42662
25th Percentile	1.	0.846691	0.884112	0.894278	0.959494

Figure 2: Simulated Funded Ratios (Using Market Value of Assets) of the Model Plan with Default Assumptions

1.3 Exploration

Our goal is to find a set of plan measurements—metrics—that we can combine in a weighted average to score competing policies. In order to accomplish this, we will first explore the effect of a number of individual metrics on a small set of sample smoothing policies. The measurements we select will share a few characteristics:

- Semantics—the metric should mean something of importance to the stakeholders of the pension plan.

- Discrimination—the metric should vary significantly in value among the policies being tested, so that clear winners and losers in Policy Space are clear.

- Consistency—the metric should play well with other metrics, with roughly equivalent variation by policy. This will allow the calculation of a score for each policy based on a weighted average of measurements.

- Incompatibility—The set of metrics should represent trade-offs among competing goals: Risk and reward, cost level and cost stability, high funding and low risk. Defining a score by weighting the metrics should force us to make hard choices.

In order to select among the various metrics available to us, we will test how they vary with different asset smoothing policies. To do this, we will select four policies that we hope will reasonably represent the set of alternatives. These policies are:

- Cost computed using market value;

- An aggressive smoothing policy, with investment gains and losses spread over 20 years and with a 50 percent corridor around market;

- A traditional smoothing policy, with gains and losses recognized over five years with a 20 percent corridor around market value; and

- A policy adopted by the California Public Employees' Retirement System (CalPERS), using a factor of 15.0 and a 20 percent corridor around market.

Each candidate for a metric will be tested with these four policies. Metrics that have significantly different values for each of the four test policies will be selected for further consideration. Those that do not discriminate among policies will be dropped from further consideration. Again, we will be looking at a set of meaningful metrics with significant variation and with cost/benefit trade-offs that force us to make difficult choices.

1.4 Optimization

Our final step will be to scale the selected metrics so that they are consistent with one another, so that no single measurement will dominate in a weighted average score. We will then define several possible scoring systems, involving different sets of metrics and weights. For each, we will abuse the computer by having it compute a score for each of the 1,300 policies under consideration. Then we will consider what the score means and what light it sheds on the behavior of the Plan.

No great truth will emerge. Each optimal portfolio will be best only when judged with a specific set of metrics and weights. Optimization is in the eye of the beholder.

2. Candidate Metrics

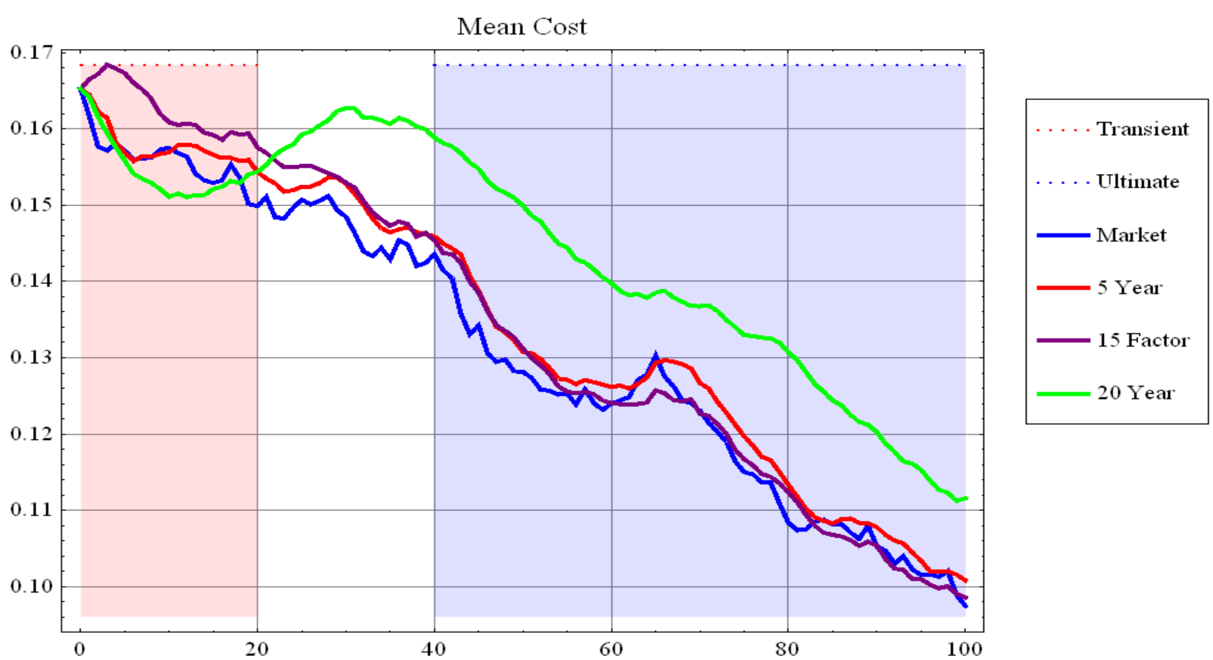
In this section we will explore some measurements that are potentially of interest in selecting an asset smoothing method. Each candidate will be tested against the four sample asset smoothing policies discussed in Section 1.3 above.

2.1 Cost Mean

The mean actuarial cost is almost always of interest to all plan stakeholders, especially the plan sponsor and taxpayers. Therefore, it certainly has semantic content. How well does it do in other respects?

Figure 3 below shows the mean cost for each year of a 100-year simulation of the Model Plan with cost computed using assets calculated four ways: At market, five-year/20 percent smoothing, factor of 15.0/20 percent smoothing (CalPERS), and 20-year/50 percent smoothing.

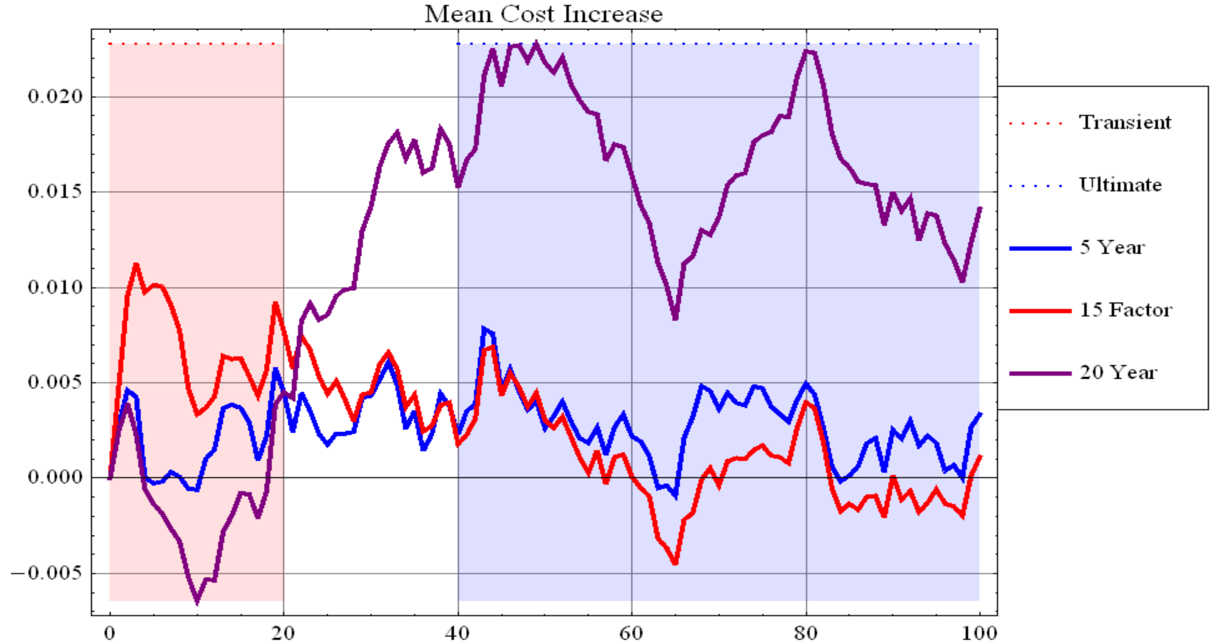
Frequently, funding policies have different effects just after adoption than they do once the policy has been in place for some time. In this paper, the initial period after adoption is labeled as Transient; it lasts for 20 years, and is shaded red in Figure 3. The years after 40 are shaded in blue and labeled Ultimate.



Year	0	25	50	75	100
Market	0.165275	0.150708	0.128161	0.114998	0.0974617
Five-Year/20%	0.165275	0.152456	0.130776	0.119707	0.100806
15 Factor/20%	0.165275	0.155127	0.131196	0.116715	0.0985721
20-Year/50%	0.165275	0.159265	0.149936	0.132973	0.111583

Figure 3: Cost Mean by Year and Asset Smoothing Policy for the Model Plan

In Figure 3, we note that the average cost decreases over time, which was noted in connection with Figure 1. During the Transient period, the smoothing methods have different effects, some producing higher and some lower costs than when assets are held at market. In the Ultimate period after 40 years, additional smoothing produces higher mean cost. This is confirmed in Figure 4 below, which shows the mean difference between smoothed cost and cost at market value by smoothing method.



Year	0	25	50	75	100
Five-Year/20%	0.	0.00174824	0.00261482	0.00470903	0.003344
15 Factor/20%	0.	0.00441923	0.00303449	0.0017163	0.00111036
20-Year/50%	0.	0.00855722	0.0217741	0.0179744	0.014121

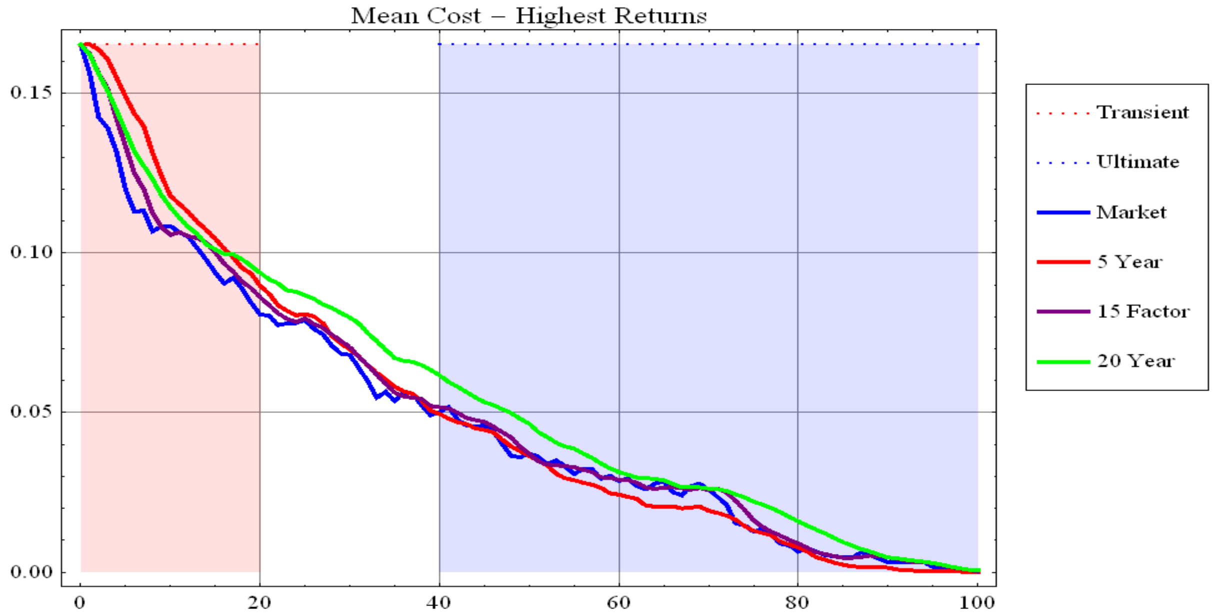
Figure 4: Difference Between Cost at Market and Cost for Each Asset Smoothing Policy for the Model Plan

We note in Figure 4 that the two smoothing methods with a 20 percent corridor generally produce higher costs than market, though the difference is fairly small (less than 0.5 percent of payroll) and tends to fade as time goes on. The 20-year/50 percent approach has much different behavior, with lower costs than market in the first 20 years but significantly higher costs than market after 40 years. This suggests that corridor width may have a material effect on cost level, with wider corridors causing an increase in mean cost after 40 years.

I want to take some time to delve more deeply into this behavior, not because it is so important or interesting in its own right, but rather because we can demonstrate an interesting and effective way to explore the effect of proposed policies.

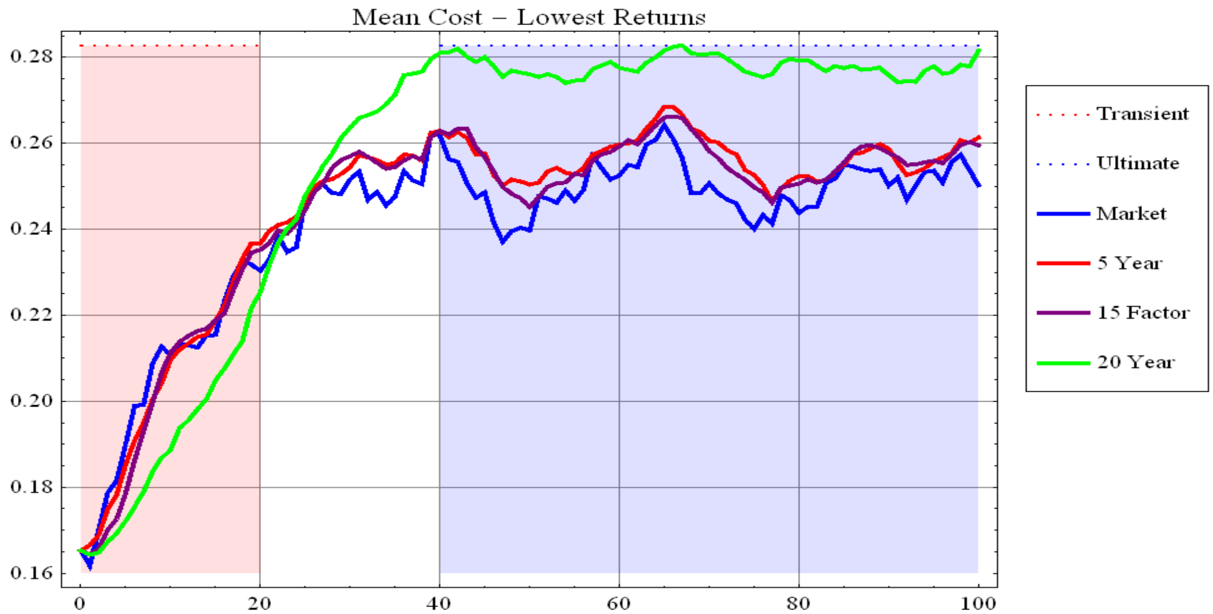
Figure 5 shows the mean cost for each asset smoothing policy over time for the top quarter of trials: The compound average return for each trial was measured, and the 250 trials (25 percent of 1,000 trials) with the highest geometric average return have their mean cost plotted in Figure 5. Note here that asset smoothing makes little difference in cost. Everything is compressed by the zero line, so all roads lead to zero. Even here, though, the 20-year smoothing produces higher mean cost.

Figure 6 shows the mean cost for the 100-year projection of the Model Plan only for those trials in the bottom 25 percent of geometric average return. Note that average cost in the ultimate period increases with more smoothing: Cost increases with poor returns are delayed by the smoothing process, so assets grow more slowly than if costs were determined at market. Consequently, the ultimate cost is elevated.



Year	0	25	50	75	100
Market	0.165275	0.0789462	0.0372002	0.0127128	0.
Five-Year/20%	0.165275	0.080847	0.0364567	0.0129918	0.
15 Factor/20%	0.165275	0.0795121	0.03708	0.0162378	0.000177414
20-Year/50%	0.165275	0.0867714	0.0465347	0.0221394	0.00056039

Figure 5: Cost Mean by Year and Asset Smoothing Policy for the Best 25% of Trials



Year	0	25	50	75	100
Market	0.165275	0.246511	0.239725	0.240075	0.25022
Five-Year/20%	0.165275	0.24594	0.250407	0.252619	0.261339
15 Factor/20%	0.165275	0.244447	0.245172	0.249791	0.259521
20-Year/50%	0.165275	0.247863	0.276078	0.276046	0.281568

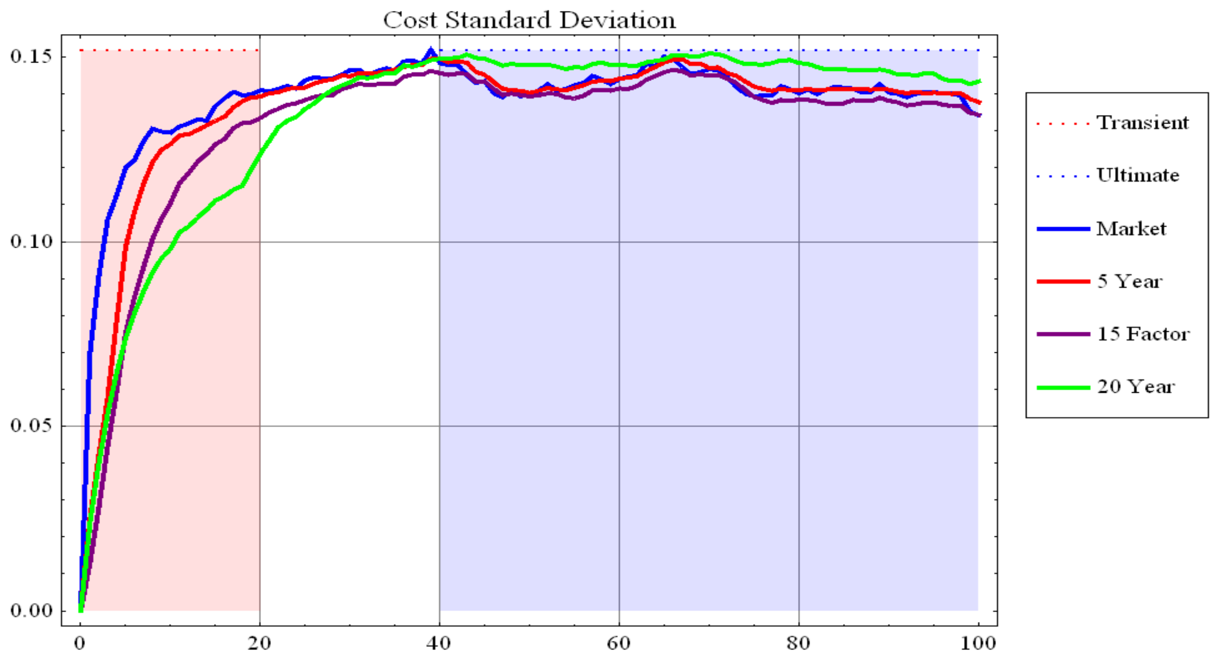
Figure 6: Cost Mean by Year and Asset Smoothing Policy for the Worst 25% of Trials

If smoothing raises the average cost in the bad trials and doesn't change it much in the good trials, it is reasonable to expect an overall increase in average cost and a decrease in average funding from smoothing. Therefore, you have to pay for asset smoothing in the form of higher costs and lower funding. Another way of looking at it is to note that contributions delayed must be repaid with interest.

2.2 Cost Standard Deviation

A traditional measure of cost risk is the standard deviation of each year's cost measured about the mean cost that year. This is a detrended standard deviation, in that the standard deviation depends only on local variation, rather than on the trend of cost from year to year. This measure should be of interest to plan sponsors, since it relates to the uncertainty associated with the plan cost; however, it is seldom discussed. For our purposes here, we would expect that the standard deviation would be a good measure of the effectiveness of an asset smoothing policy. How well does it do in this regard?

Figure 7 below shows the standard deviation of cost for each year of a 100-year simulation of the Model Plan with cost computed using assets at market, five-year/20 percent smoothing, factor of 15.0/20 percent smoothing (CalPERS), and 20-year/50 percent smoothing.

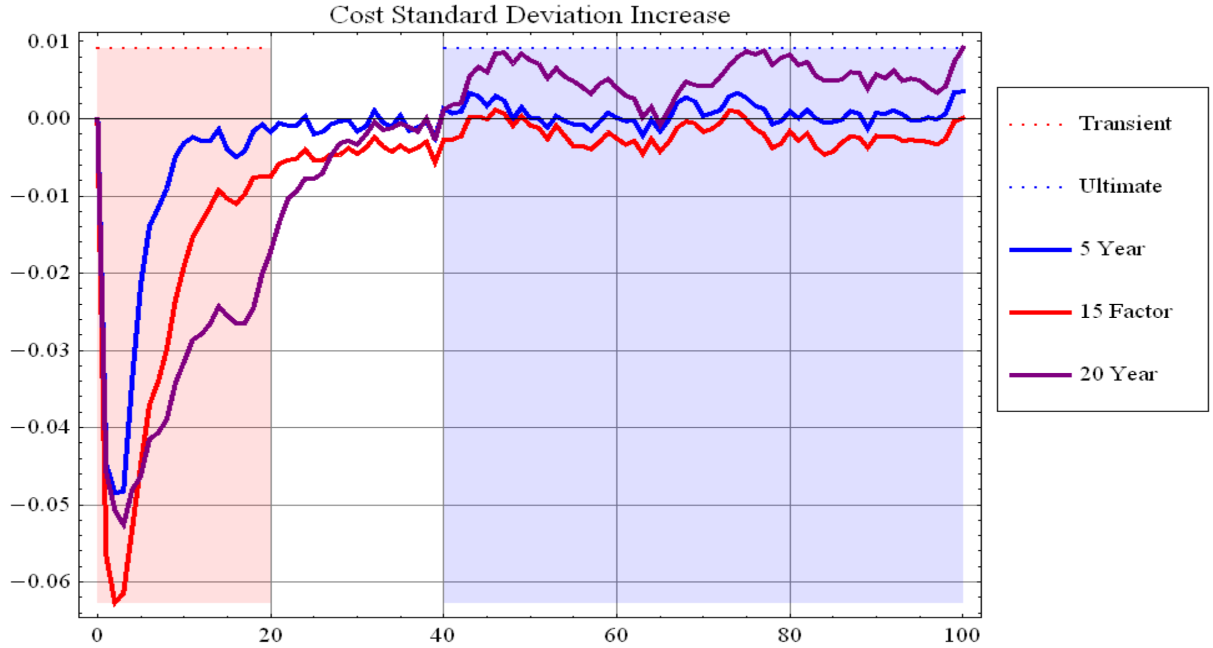


Year	0	25	50	75	100
Market	0.0	0.143748	0.140221	0.139323	0.134221
Five-Year/20%	0.0	0.141767	0.140281	0.142006	0.137755
15 Factor/20%	0.0	0.13839	0.139359	0.139205	0.134336
20-Year/50%	0.0	0.13596	0.147772	0.148047	0.143407

Figure 7: Cost Standard Deviation by Year and Asset Smoothing Policy for the Model Plan

We note that in the Transient period during the first 20 years after the smoothing policy is adopted, any of the smoothing policies result in significantly reduced standard deviation of cost. Therefore, in the years immediately following adoption, there is a material benefit.

We see this more readily in Figure 8, which shows the *change* in the standard deviation of cost as we move from assets at market value to smoothed assets in computing cost.



Year	0	25	50	75	100
Five-Year/20%	0.	-0.00198167	0.0000603389	0.00268344	0.00353414
15 Factor/20%	0.	-0.0053579	-0.000861366	-0.000117898	0.000114599
20-Year/50%	0.	-0.0077887	0.00755091	0.00872345	0.00918626

Figure 8: Difference in Standard Deviation of Cost at Market and Standard Deviation of Cost for Each Asset Smoothing Policy in the Model Plan

After 40 years, the aggressively smoothed asset policy with the wide corridor around market value has volatility equal to or slightly exceeding the less smoothed costs. Note that traditional smoothing, factor 15 smoothing, and no smoothing have about the same variation after 40 years; in both Figure 3 and Figure 7 we see that the 20 percent market value corridor keeps the costs relatively close to cost computed at market value, both in mean and standard deviation.

The sharp difference between the effect of asset smoothing in the Transient and Ultimate periods forces us to ask what our principal goals are. Specifically, should we measure the temporary phenomenon in the Transient period, or should we focus on long-term global effects in the Ultimate period?

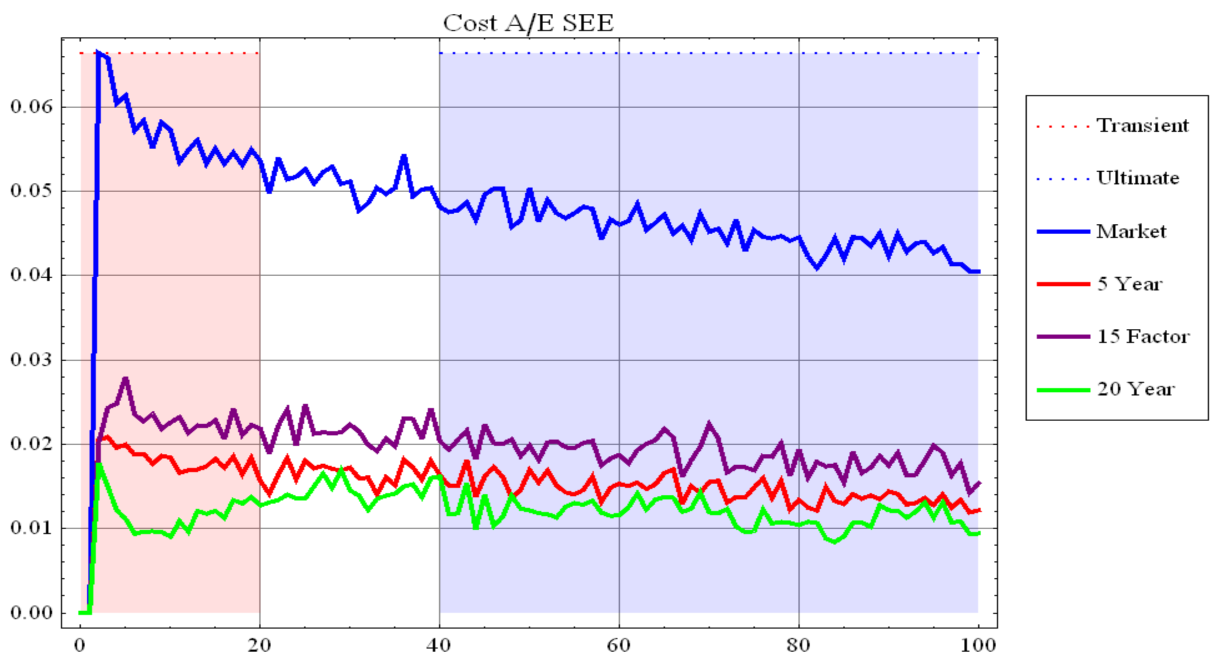
2.3 Cost Predictability

In the previous section, we discussed cost variability using the standard deviation of each year's cost measured about the mean cost that year. But variability by itself may be a poor measure of risk, particularly when part of the variability may be predictable. Accordingly, we develop a different measure of risk, based on the predictability of cost from one year to the next.

In this measurement, we compare the actual cost in a year with the cost that would have been predicted the year before using the plan's funding status and the number and magnitude of past investment gains and losses being deferred by the asset smoothing policy. We fit a regression line of actual cost to expected cost; the standard estimate of the error of this regression line (the A/E SEE) is a measure of the predictability of costs. We would hope this would improve (the standard error of the actual vs. expected cost would decrease) as we add asset smoothing.

Figure 9 below shows the A/E SEE for the 100-year projection of the Model Plan with cost computed using assets at market, five-year/20 percent smoothing, factor of 15.0/20 percent smoothing (CalPERS), and 20-year/50 percent smoothing. The first two years are shown with an A/E SEE of 0.0 because the variance of the expected cost is zero in years 0 and 1.

We note primarily in Figure 9 the fact that the A/E SEE is fairly level over time, making it a good metric. In particular, there is little need to distinguish between behavior during the Transient and Ultimate periods. We also note that asset smoothing—with any of the three alternatives considered here—materially improves the predictability of plan costs. Moreover, the expected improvement in predictability is robust over time. This contrasts with the situation for the cost standard deviation, in which smoothing reduces standard deviation in the short term, but does little to reduce this measure of variability in the long term.



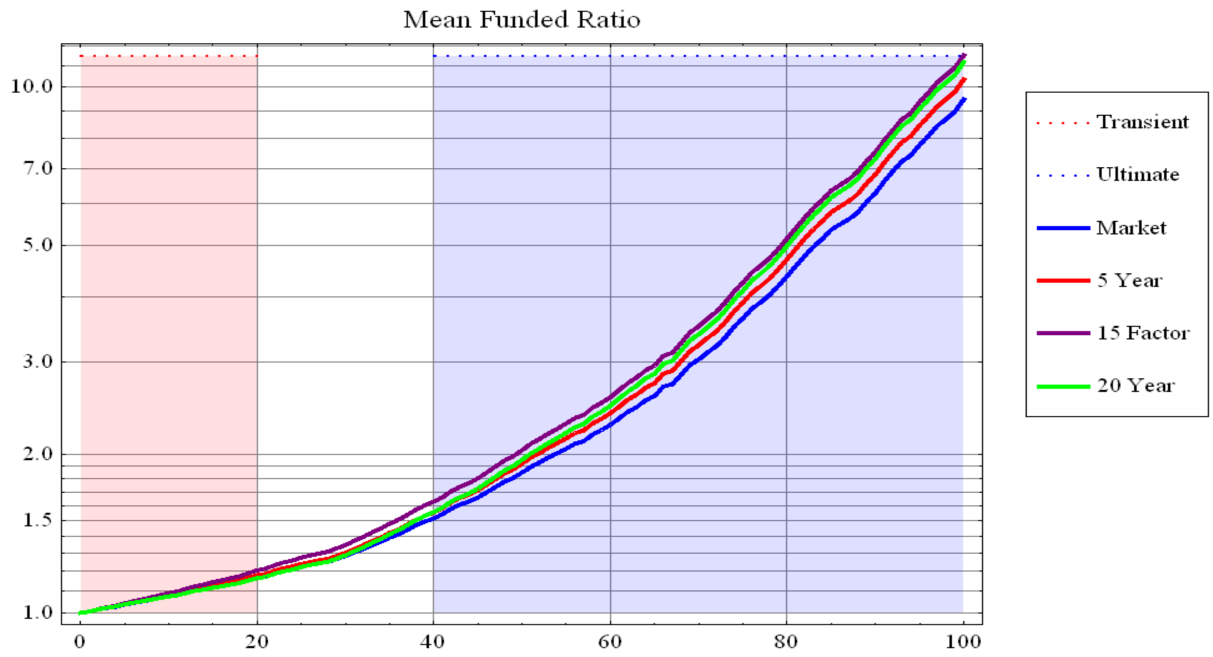
Year	0	25	50	75	100
Market	0.0	0.052655	0.050395	0.045425	0.040518
Five-Year/20%	0.0	0.01805	0.016944	0.014876	0.012169
15 Factor/20%	0.0	0.024691	0.019851	0.016915	0.015362
20-Year/50%	0.0	0.013558	0.012146	0.009689	0.009412

Figure 9: Cost Standard Error of Estimate by Year and Asset Smoothing Policy for the Model Plan

2.4 Funding Metrics

We won't spend a lot of time in this paper discussing funding metrics as they pertain to asset smoothing policies. There are a couple of reasons for this. First, in considering asset smoothing policies we are usually primarily concerned with actuarial cost levels and variability. Our concern with funding levels is generally to be sure that they are not affected too seriously. The second reason is that funding levels are difficult to measure and compare with different funding policies.

Consider Figures 10, 11 and 12 below. Here we plot the log of the funded ratio (Figure 10), the log of the funded ratio for the 25 percent of trials with the highest geometric average returns (Figure 11), and the log of the funded ratio for the 25 percent of trials with the lowest returns (Figure 12).



Year	0	25	50	75	100
Market	1.	1.22554	1.84974	3.63337	9.48764
Five-Year/20%	1.	1.23896	1.92369	3.89126	10.354
15 Factor/20%	1.	1.27392	2.04025	4.24892	11.4958
20-Year/50%	1.	1.22286	1.95568	4.10018	11.163

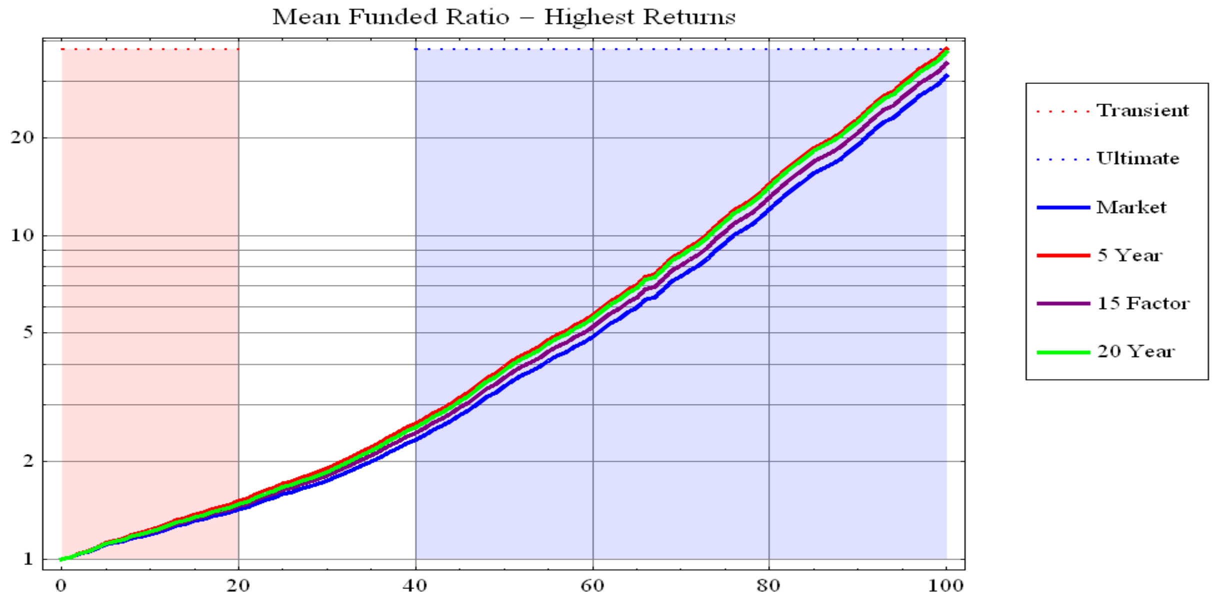
Figure 10: Log Plot of Funded Ratio by Year and Asset Smoothing Policy for the Model Plan

In Figure 10, we see that the mean funded ratios appear very near to one another, but it's difficult to tell much because of the high values. Note that the average funded ratios after 100 years are around 1,000 percent. It's interesting to note that the lowest funded ratio is for cost computed using market value of assets. The reason for this result can be inferred from Figures 11 and 12.

In Figure 11 we see the funded ratio for the 25 percent of trials in which the geometric return on assets is highest. We see here that the funded—or overfunded—ratio reaches staggering heights over 3,000 percent, and the funded ratio is lowest when assets are held at market value. On the other hand, we note in Figure 12—which consists only of those trials with returns in the lowest 25 percent—funded ratios are between 80 and 90 percent, and the funded ratio is highest when assets are at market value. Note in particular that the most aggressive smoothing—20 years with a 50 percent corridor—has the most profound impact on funding.

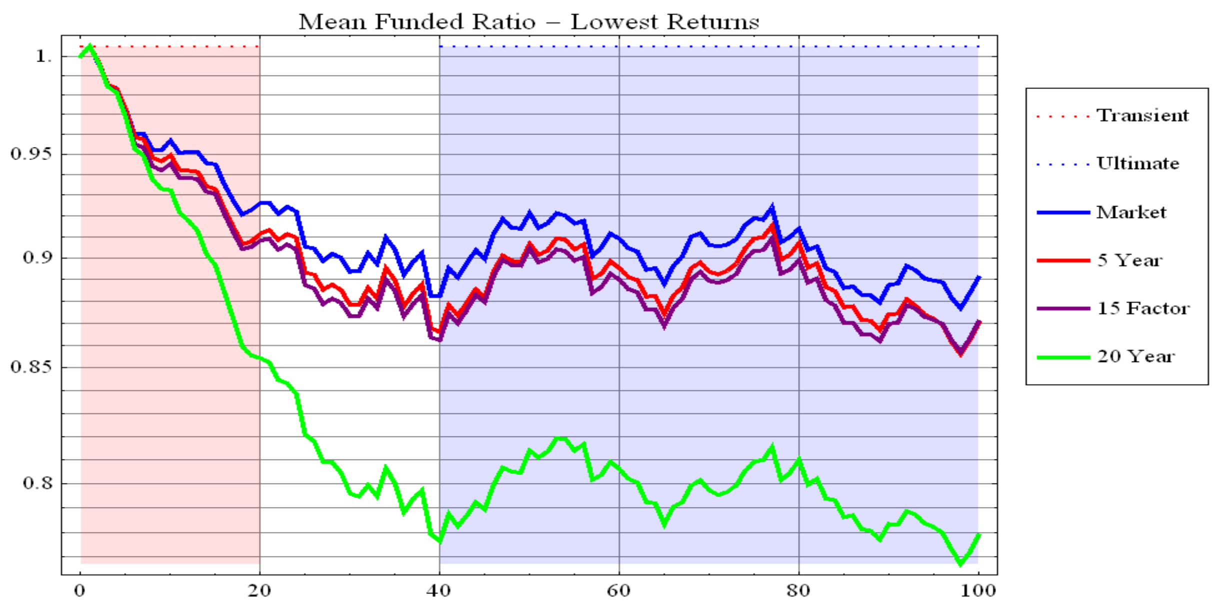
This contrast helps us explain the relationship of funding to asset smoothing. When returns are good (Figure 11), smoothing slows down the recognition of gains and keeps the actuarial cost above the cost using market value, thus pushing the funded ratio up. The reverse occurs with bad returns (Figure 12): Asset smoothing delays recognition of losses, keeping actuarial costs low, and producing lower funded ratios than if cost were computed at market.

Insofar as the average in Figure 10 is concerned, the extraordinarily high funded ratios in times of good returns in Figure 11 swamp the funding shown with bad returns in Figure 12, creating the mistaken impression that asset smoothing improves funding. In fact, the effect of asset smoothing on funding depends entirely on the return on assets. Reviewing the impact of asset smoothing on median funded ratio produces the same conclusions.



Year	0	25	50	75	100
Market	1.	1.59092	3.41573	9.49498	31.141
Five-Year/20%	1.	1.71483	3.92542	11.3091	37.7501
15 Factor/20%	1.	1.63934	3.63147	10.2942	34.0418
20-Year/50%	1.	1.67441	3.82959	11.0503	36.8761

Figure 11: Log Plot of Funded Ratio for Highest Return Trials by Year and Asset Smoothing Policy



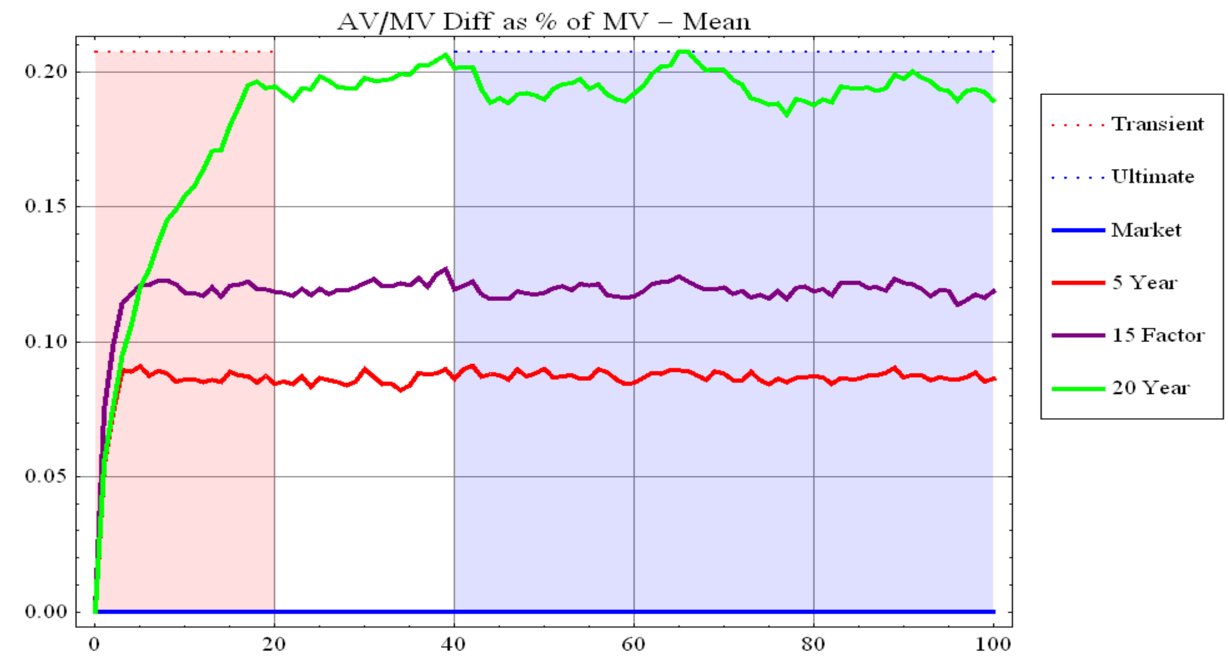
Year	0	25	50	75	100
Market	1.	0.905323	0.921375	0.919047	0.890875
Five-Year/20%	1.	0.893281	0.906692	0.909194	0.869757
15 Factor/20%	1.	0.887367	0.904516	0.903481	0.870576
20-Year/50%	1.	0.820681	0.81414	0.809281	0.778807

Figure 12: Log Plot of Funded Ratio for Lowest Return Trials by Year and Asset Smoothing Policy

Measures of funding variation—the standard deviation and the standard error of regression of actual versus expected funded ratios—are affected only indirectly by asset smoothing. Contributions to the Model Plan are mostly in the range of 0 percent to 5 percent of assets, so smoothing actuarial contributions will have little effect on the variability of plan assets; plan asset variation is caused predominantly by variation in investment returns rather than contribution levels. Instead, as noted above, cost smoothing has an impact on funding levels, which in turn affects measures of funding variation: Higher funding levels experience higher measures of variation in absolute terms.

2.5 Special Metrics

In any policy discussion, there are likely to be metrics that are unique to the policy in question. In the case of asset smoothing, one such metric could be the average absolute value of the distance between actuarial value of assets and market value, expressed as a percentage of market value. We see a graph of this metric below in Figure 13.



Year	0	25	50	75	100
Market	0.	0.	0.	0.	0.
Five-Year/20%	0.	0.086612	0.089974	0.08423	0.086335
15 Factor/20%	0.	0.119661	0.118173	0.11597	0.118721
20-Year/50%	0.	0.198355	0.189761	0.187835	0.189332

Figure 13: Absolute Difference Between Smoothed Asset Value and Market as a Percentage of Market

We note that, in each case, the level of the difference between smoothed and market values of assets (the AV/MV distance) rises to an ultimate value and stabilizes. This process takes longer when the smoothing period is longer, and the ultimate level of the average AV/MV distance is strongly influenced by the required corridor around market value, which is no surprise. Note that this measure is stable after 20 years, which offers the hope of a reliable long-term metric.

2.6 Summary of Metrics

We have touched on the impact of a number of metrics above. In some cases we have gone into detail regarding the measurement and the reasons for its behavior. There are many metrics we have not discussed, either in the interest of space or because the metric duplicates the probative value of another metric. The table below summarizes the behavior of the metrics we have discussed.

Metric	**Cost**	**Funding**
Mean Value	In the first 20 years, no pattern emerges, with some smoothing methods increasing mean cost and some decreasing mean cost. After 40 years, asset smoothing increases mean cost: In trials with high returns, smoothing makes little difference, but in trials with low returns smoothing delays higher contributions and increases ultimate cost levels as a result.	The impact of asset smoothing on funding is to delay the recognition of returns in the form of changes in actuarial cost. In high-return scenarios, this improves funding; while in low-return scenarios funding is lower than when assets are held at market.
Median Value	Median cost behavior is identical to the behavior of the mean cost.	Median funding behavior is identical to the behavior of the mean funded ratio.
Standard Deviation	In the first 20 years after the smoothing policy is adopted, the smoothing policy results in significantly reduced standard deviation of cost. Note that after 40 years, smoothing methods have about the same standard deviation as cost at market value, with the most aggressively smoothed cost having the highest volatility.	Smoothing of contributions has little effect on variability of funding, since contributions are small relative to assets. Asset smoothing affects the variability of the funded ratio through its effect on the level of funding. With good returns, funding increases, and with it variability of funding. The opposite happens when returns are unfavorable.
Predictability (A/E SEE)	In all years, plan cost is much more predictable with asset smoothing than without.	Asset smoothing operates on funding predictability as it does on funding standard deviation.
Absolute Difference Between Smoothed and Market Assets as a Percentage of Market (AV/MV Distance)	The mean AV/MV distance is level after an initial phase-in and increases with additional smoothing.	

3. Policy Optimization

Recall the problem definition in Section 1.1: We want to explore a large number of possible asset smoothing policies. Of the infinite number of potential policies available, we have settled on 1,300 candidates, which can be described as follows:

- We consider both year-based and factor-based smoothing. In year-based smoothing, investment gains and losses are recognized over a fixed number of years, typically five. In factor-based smoothing the difference between market value and expected actuarial value is divided by a factor and added to expected actuarial value. This has the impact of going 1/factor of the way between expected value and market value each year.

- In the year-based approach, investment gains and losses will be recognized over from zero (assets are at market) to 25 years in steps of one year; in the factor approach, factors from 1.0 to 25.0 will be used, in steps of 1.0.

- Smoothed assets will be constrained to remain within a corridor around market value from 0 percent (market value) to 50 percent in steps of 2 percent.

3.1 Basic Measurements

We proceed by measuring performance of the Model Plan under each of the 1,300 smoothing policies in our Policy Space. A compact presentation of the results of this process is challenging; consider Figure 14 below.

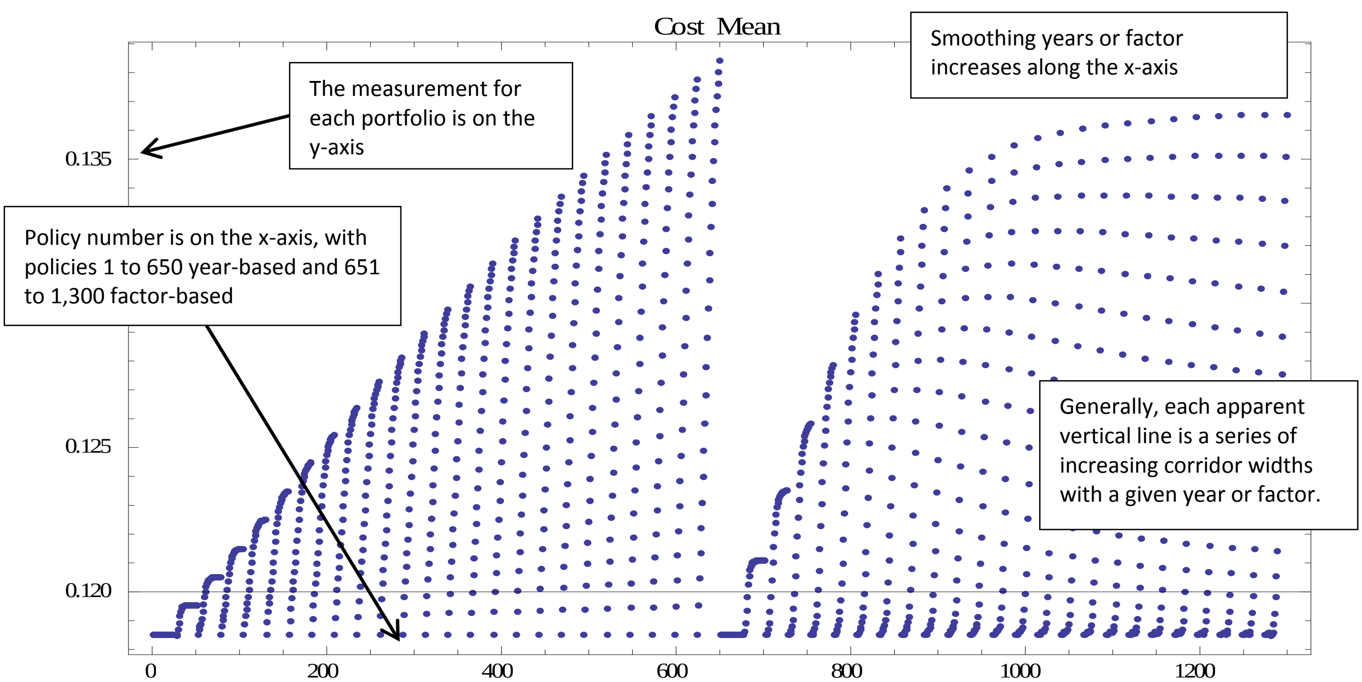


Figure 14: Mean Actuarial Cost for 1,300 Candidate Asset Smoothing Policies for Years 40 Through 100
(The policy with the highest mean cost is year-based, with a 25-year smoothing period and a 50% corridor. The policy with the lowest mean cost is year-based with two-year smoothing and a 2% corridor.)

In Figure 14 we attempt to present an overview of the results of a given measurement of the performance of the 1,300 portfolios. The graph is split in half: Year-based policies in the left half; factor-based policies in the right. Along the horizontal axis, within each half, the number of years of smoothing or the size of the smoothing factor increases from one year or 1.0, respectively, to 25 years or a factor of 25.0. Within each year or factor value, the width of the market value corridor increases.

For example, policy numbers 0 through 25 use one-year smoothing with corridors of 0 percent to 50 percent in 2 percent increments. Then policies 26 through 51 use two-year smoothing, with corridors from 0 percent to 50 percent in 2 percent increments, making policy 27 two-year smoothing with a 2 percent corridor.

The value of the measurement is shown in the vertical axis. The measurements tend to fall into a series of nearly vertical lines, each line representing a given smoothing period or factor with a series of corridor widths. We notice that each line starts along a horizontal axis, representing a zero corridor width, so that assets are at market value. As the corridor width increases, the mean cost increases with it, which was noted earlier.

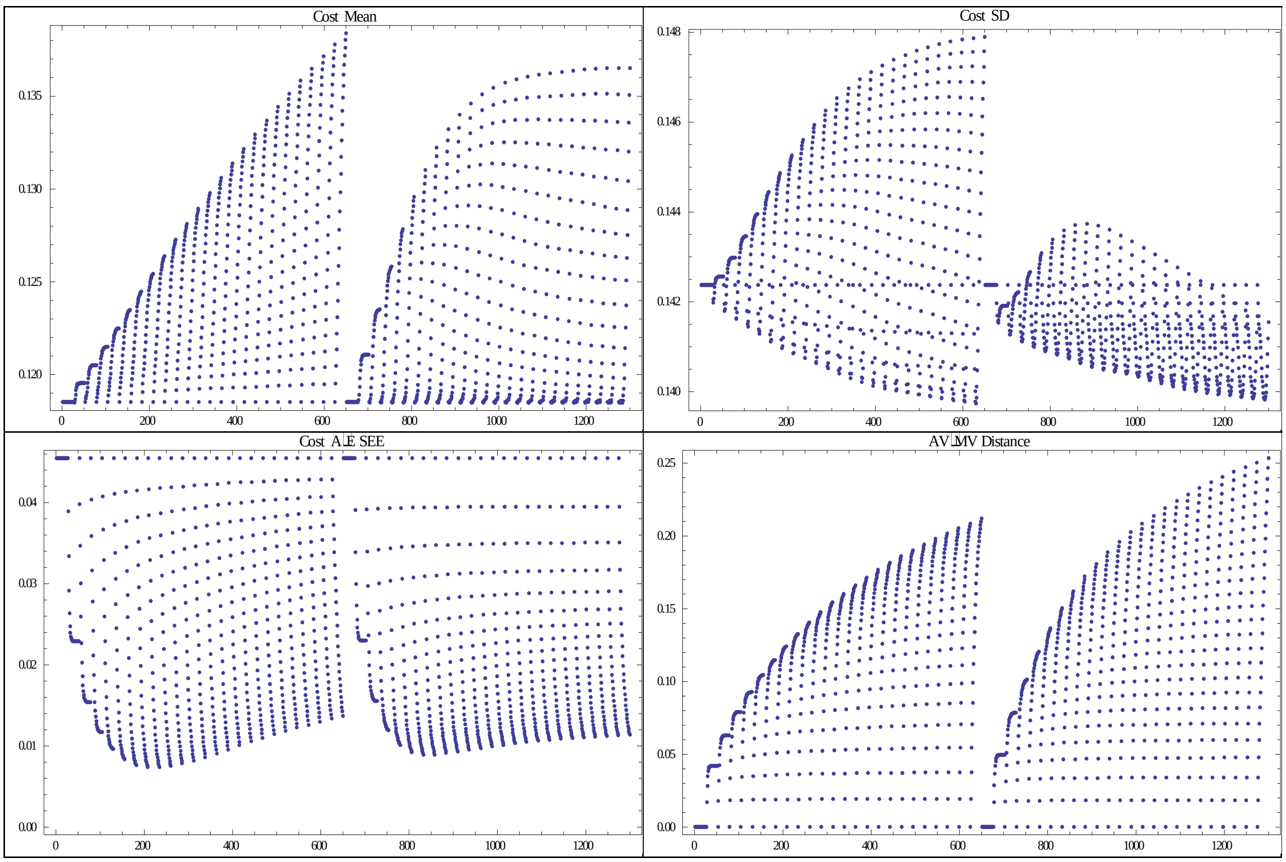


Figure 15: Four Metrics for 1,300 Candidate Asset Smoothing Policies
(Shown are mean actuarial cost (upper left), standard deviation of cost (upper right), standard error of estimate of actual vs. expected cost (lower left), and mean absolute difference of actuarial value from market value (lower right).)

In Figure 15 above we have displayed an array of four policy graphs. At the top left is the mean actuarial cost for simulation years 40 through 100, with assets computed using the 1,300 candidate smoothing policies under consideration. The top right graph shows the annual standard deviation of the actuarial cost; the bottom left shows the standard error of estimate of the regression of actual versus expected cost;

and the bottom right graph shows the mean absolute difference between the actuarial and market values of assets as a percentage of the market value of assets.

Even a quick glance shows how differently the candidate policies behave under each of the four metrics. The mean cost plot is a repetition of Figure 14. It shows that year-based and factor-based smoothing have comparable behavior: In general, more smoothing and a wider corridor produces an increase in the mean cost.

However, the policy plot of the cost standard deviation shows much different behavior. Factor-based smoothing methods have lower standard deviations of cost, and the changes in the standard deviation with increased smoothing years and wider corridors are fairly complicated. The measure of cost predictability (the A/E SEE) shows decreasing standard errors of estimate with increased smoothing, but only to a point. Lastly, the AV/MV distance graph is similar, but not identical, to the graph of mean cost.

The implications of the measurements vary as well. High mean cost, standard deviation and A/E SEE are to be avoided, but a high value of the funded ratio is to be desired. Any realistic policy analysis will involve trade-offs: We may want predictable cost, but we may want to avoid policies that result in an excessive difference between smoothed asset value and market value.

How can we combine these and other metrics to evaluate policy choices?

3.2 Normalization

When we look at the metrics in Figure 15, we note that they vary on different scales. For example, the A/E SEE varies from a bit under 1 percent to a little below 5 percent, while the AV/MV distance ranges from 0 percent to 25 percent. The cost standard deviation has a narrow range of just 0.8 percent, from 14 percent to 14.8 percent. If we were measuring the mean funded ratio, that metric could range from around 40 percent to well over 100 percent. Combining these measurements into a composite score requires that they be scaled, or normalized, so that one measurement does not affect the outcome in a way that is out of proportion to the importance assigned to the metric by the user.

In Figure 16 below, we put all the metrics on the same scale; when we do so, another issue emerges. The AV/MV distance varies much more widely among the various candidate policies than the other metrics, so differences in the AV/MV distance metric will overshadow those arising from cost mean and standard deviation and will largely determine the outcome of any reasonable weighting of the different metrics.

Therefore, there are two issues to be resolved:

- The metrics to be combined must be put on the same scale so that the magnitudes are comparable. For example, funded ratios cannot be compared with actuarial costs (usually).

- The amount of variation of each metric among the candidate policies must be reasonable. Otherwise, the metrics with the greatest variation will determine which of the policies will be selected. On the other hand, if the changes in policy do not result in a significant change in a metric, its variation should not be exaggerated. For example, the cost standard deviation is really not affected much by the asset smoothing policy, so its variation should not be artificially increased.

There are a number of algorithms for normalizing metrics. For example, we could compute the mean and standard deviation of the metric and compute a scaled metric by subtracting the mean from each data point and dividing the result by the standard deviation.

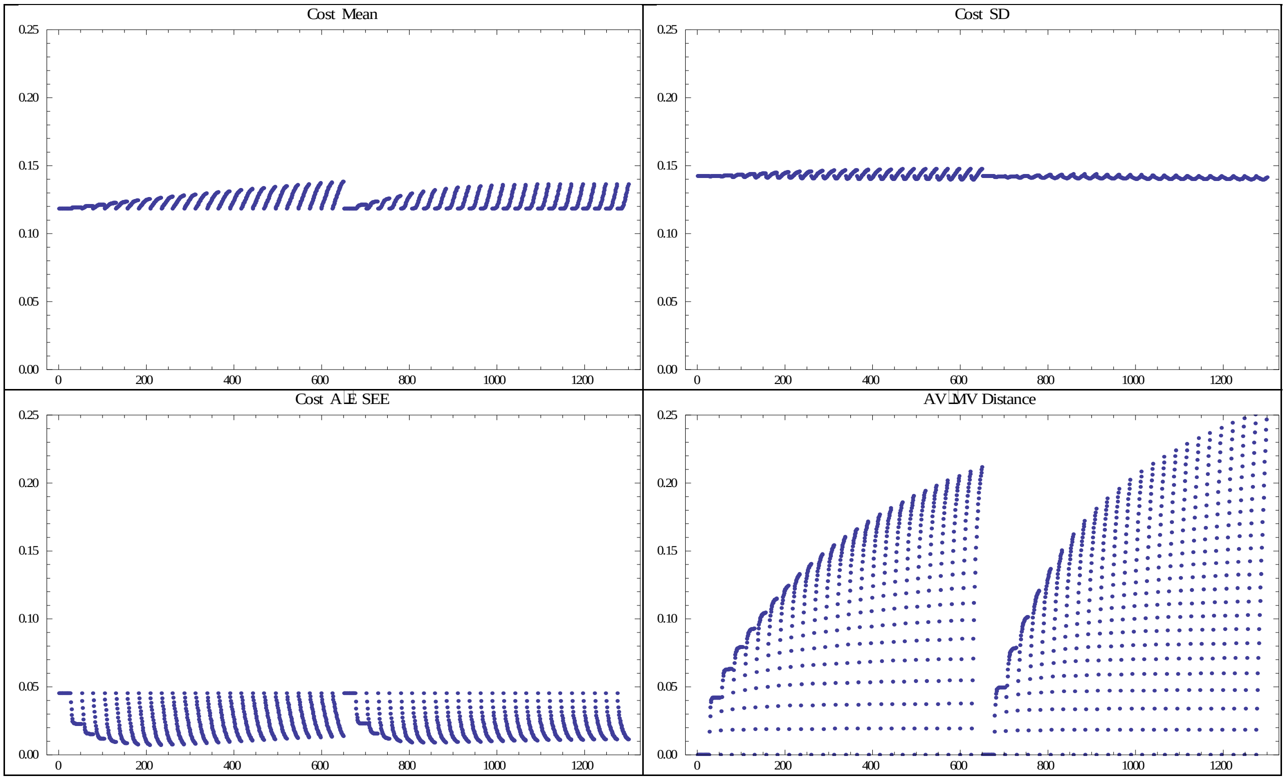


Figure 16: Four Metrics for 1,300 Candidate Asset Smoothing Policies Graphed on a Common Scale
(In each case the vertical axis varies from 0 percent to 25 percent.)

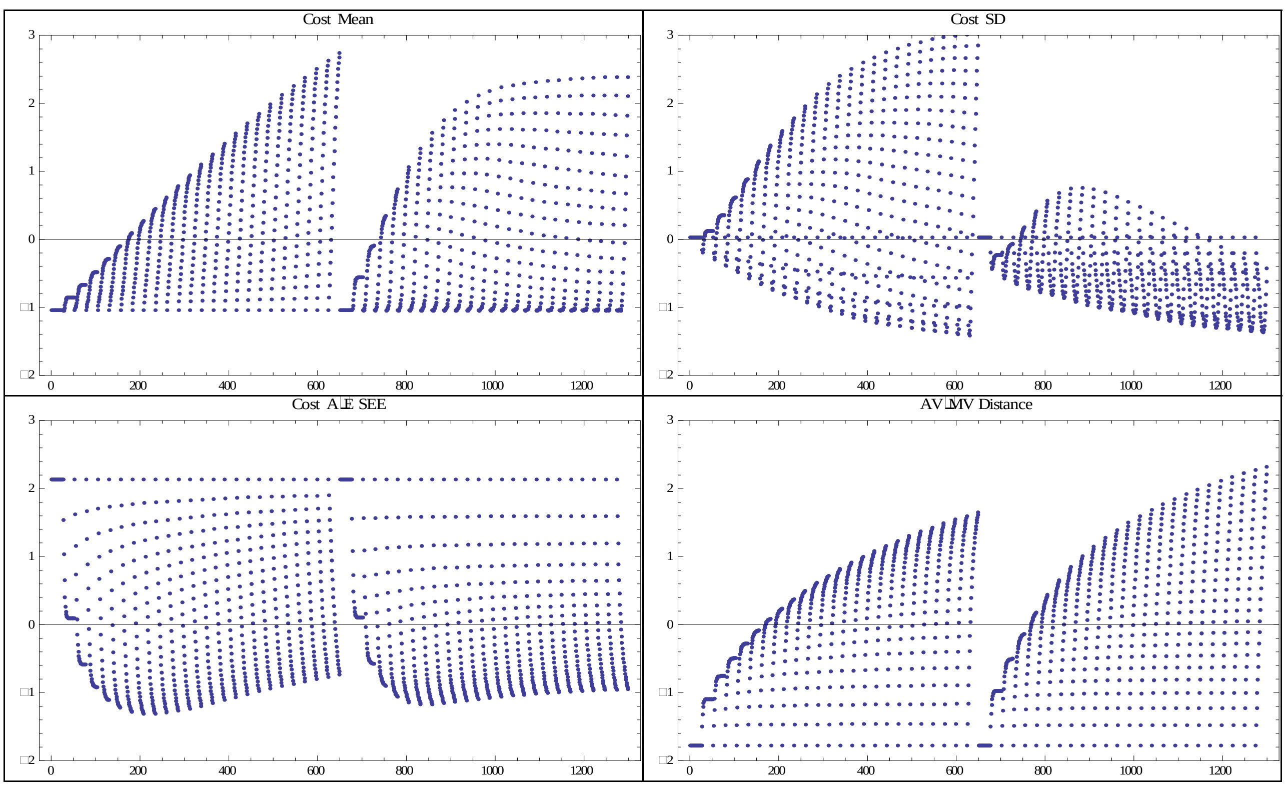


Figure 17: Four Metrics for 1,300 Candidate Asset Smoothing Policies Graphed on a Common Scale
(In each case the metric was normalized by subtracting the mean and dividing by the standard deviation.)

Figure 17 shows the result of this approach. All of the subcomponent graphs have a range of from -2 to +3 standard deviations around the mean.

We have not adopted that approach here for two reasons: First, our measurements are not normally distributed, which is clearly seen in the graph. We note that variations are both excessively wide and asymmetrical for a normally distributed value.

Second, the traditional method of normalizing tends to exaggerate the importance of measurements with a small standard deviation. We note that the normalized cost standard deviation now shows wide variation by smoothing policy, and as such it would play a major role in selecting a policy. This is unwarranted since cost standard deviation is relatively unaffected by smoothing policy.

Another approach would be to scale all measures linearly from 0 to 1, with the minimum value being assigned 0 and the maximum the value 1. This algorithm succeeds in putting all the competing metrics on the same scale, but it does so by stretching those metrics with small variation and overstating their importance, just as traditional normalization does. The graph for this approach (not shown) looks very similar to Figure 17, but with a scale of 0 to 1.

We use a different approach here. In normalizing each metric we will divide it by its maximum absolute value. Therefore, all metrics will lie between 0 and 1, but those metrics with limited range will have their normalized values clustered just below 1.0, so they will not have a disproportionate impact on the combined metric when comparing policies. The result of normalizing with this approach is shown in Figure 18 below.

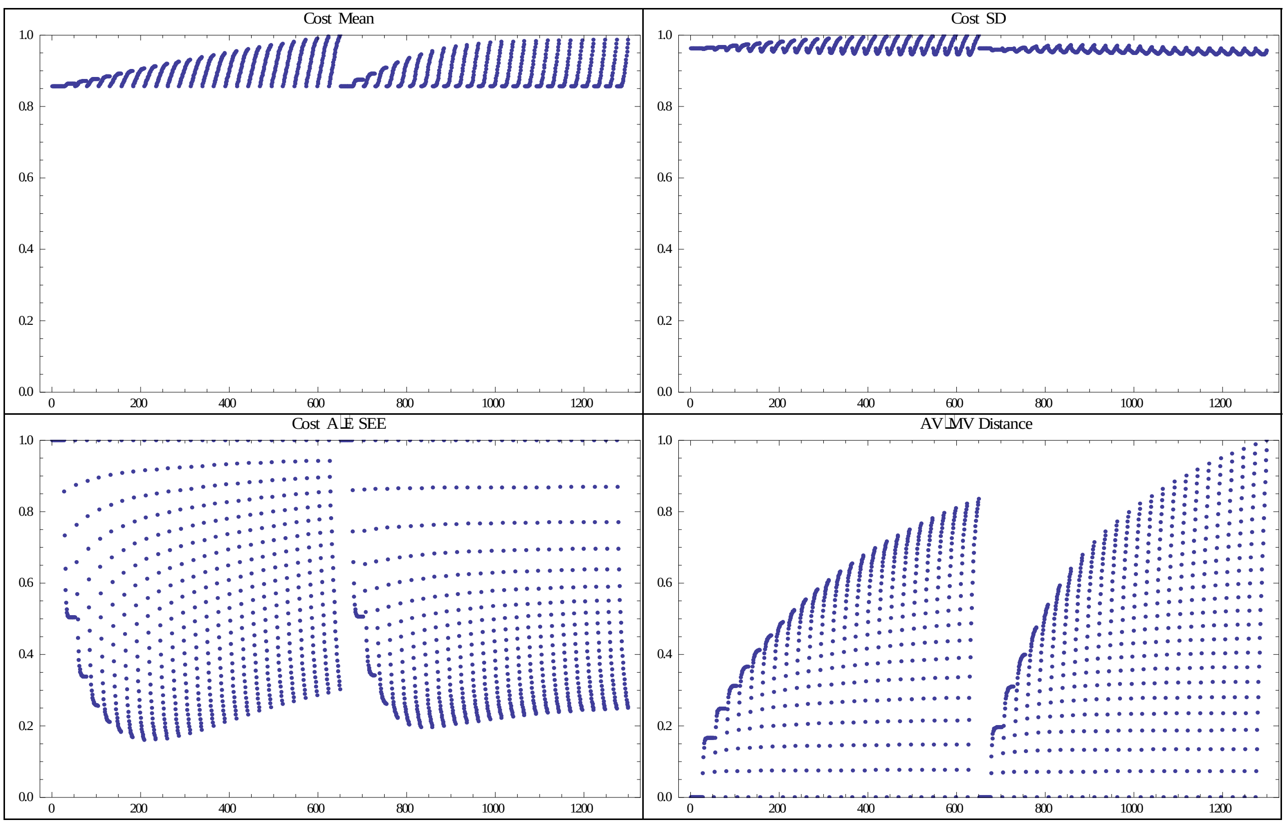


Figure 18: Four Metrics for 1,300 Candidate Asset Smoothing Policies Graphed on a Common Scale
(In each case the metric was normalized by dividing by the maximum value of the metric.)

We note in Figure 18 that we have achieved our two goals: The metrics are all on a common scale, and those metrics with minimal variation among policy choices continue to have small variation when normalized.

So far we have just concerned ourselves with deriving a common scale for our metrics. There is a second issue of combining metrics that have desirable and undesirable consequences. The normalized values of metrics should be on the same scale, and the value assigned to each metric should reflect its value to the plan. Accordingly, for metrics representing goals to be sought, we will divide the value of the metric by its maximum value. For metrics representing events to be avoided, we will subtract the metric divided by its maximum from 1.

Therefore, desirable metrics will be toward the high end of the range from 0 to 1, while undesirable metrics will cluster at the low end of the range. As a result, all combined metrics will be in the range from 0 to 1, and a high value of the combined metric is a result to be sought, indicating the best policy choice.

In Figures 15 through 18, the four metrics illustrated all define measurements that we would like to minimize—mean and variation of actuarial cost, the unpredictability of cost (the A/E SEE) and the difference between smoothed assets and market value of assets. Accordingly, all would be subtracted from 1 to compute the final metric. For the sake of completeness, the results are shown in Figure 19.

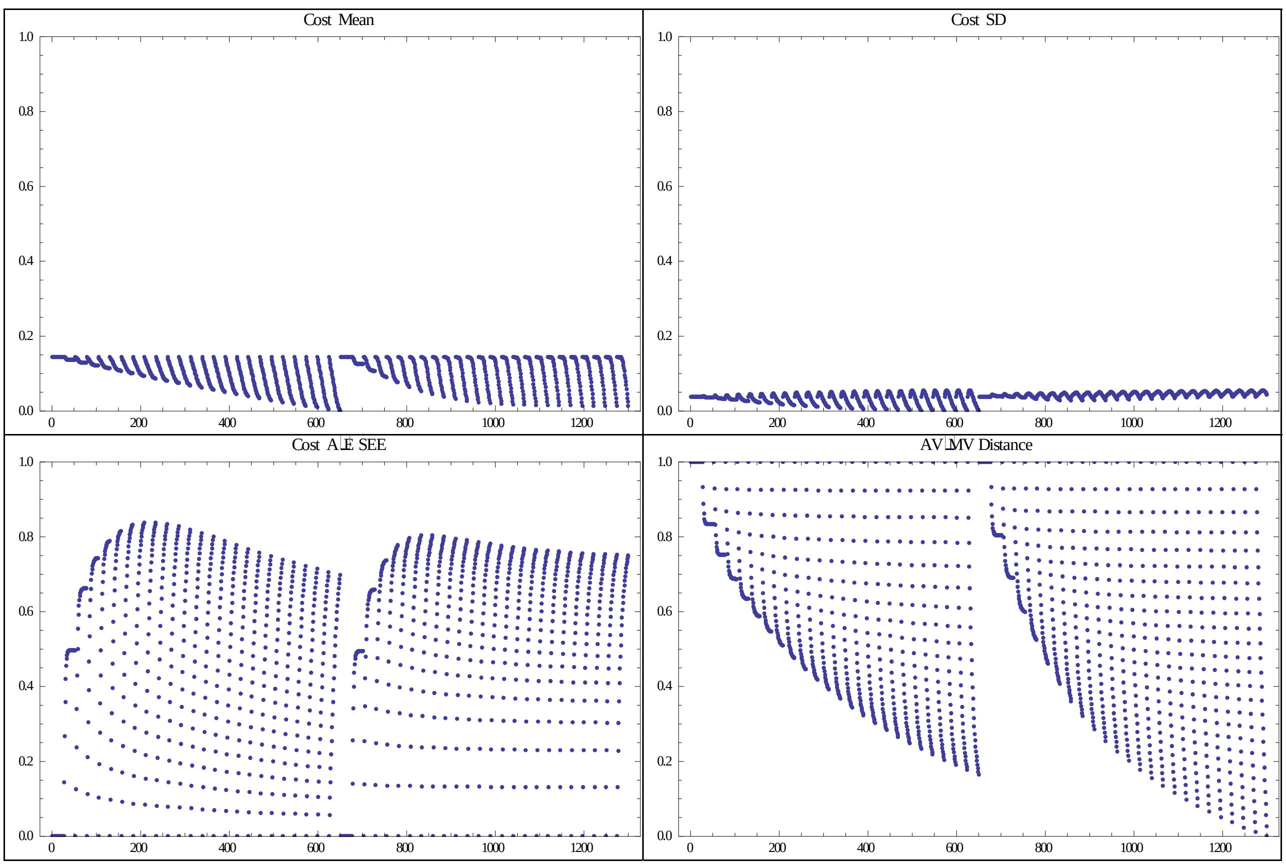


Figure 19: Four Metrics for 1,300 Candidate Asset Smoothing Policies Graphed on a Common Scale
(In each case the metric was normalized by dividing by the maximum value of the metric, then subtracted from 1 to code the undesirability of high values.)

3.3 Policy Decisions

Now we can combine our metrics into a single score, using this score to find the best policy given our weightings of the metrics that matter to us. We will show the results of three weightings:

- A combined score composed of the average of the normalized values of the mean and standard deviation of the actuarial cost;

- A combined score equal to the average of the normalized predictability of actuarial cost (the A/E SEE) and the normalized difference between smoothed and market asset values; and

- A combined score formed by weighting the normalized mean cost, standard deviation of cost, predictability of cost and AV/MV distance by 25 percent each.

We will see in each case that changes in the weights assigned to the various metrics will change the policy that scores highest and would be selected. In other words, the policy you choose depends on your priorities, a commonsense result.

3.3.1 Cost Mean and Standard Deviation

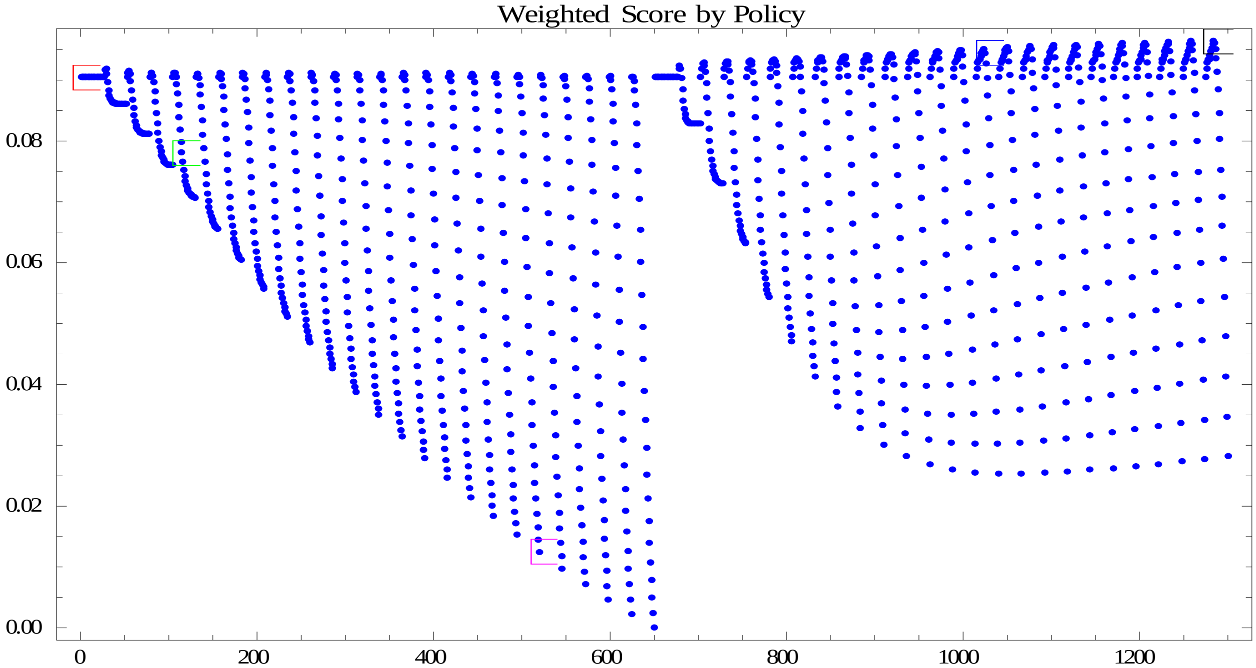
As an initial point of departure, suppose we focus on two traditional measures of plan cost. From the discussion above, we know that adding asset smoothing can reduce cost standard deviation, but the more smoothing we introduce, the more the average plan cost tends to increase. You don't get reduced risk without paying for it. Now suppose we want to improve the stability of plan cost by reducing the standard deviation of cost, but we don't want to have the mean cost increase too much.

Assuming that we want both of these goals equally, we will weight both cost mean and cost standard deviation at 50 percent in creating a combined score equal to 50 percent of the normalized mean and 50 percent of the normalized standard deviation. Recall the normalized value of each of these is obtained by dividing the metric by its maximum value and subtracting it from 1.0, since both higher mean and higher standard deviation are to be avoided.

Figures 20 and 21 below show the result. In Figure 20, each small blue dot is a policy that is plotted with its policy number on the horizontal axis and its weighted score on the vertical axis. The large black dot is the policy with the highest weighted score. The red square is the policy in which actuarial cost is computed using the market value of assets, and the green square is traditional five-year, 20 percent corridor smoothing. The blue square is factor-based smoothing, with a factor of 15.0 and a 20 percent corridor, and the magenta square is 20-year smoothing with a 50 percent corridor.

The smoothing policy with the highest score recognizes asset gains and losses over a 25-year period, but requires an 18 percent corridor around market value. There is nothing particularly intuitive about this, except to note that this is the best compromise between reducing the standard deviation of cost without an excessive increase in average cost. It is also worthy of note that this policy would be unlikely to be considered other than as a result of the technical process undertaken here.

The factor of 15/20 percent corridor policy adopted by CalPERS is fairly close in results to the highest scoring policy, but it does so without the excessively large smoothing factor. This suggests that implicit in the choice of smoothing policy for CalPERS was a balance of reducing cost variation but not incurring a large increase in mean cost by doing so. The most aggressive smoothing method—20-year smoothing with a 50 percent corridor (the magenta point)—scores worse than the other principal policy choices, and worse than the majority of all policy options tested.



	Policy	Type	Factor	Corridor	Cost Mean	Cost SD	Score
● (black)	1284	Factor	25.	0.18	0.118597	0.140523	0.0964594
■ (red)	1	Years	1.	0.	0.118522	0.142381	0.0904488
■ (green)	115	Years	5.	0.2	0.121306	0.14306	0.0780954
■ (blue)	1025	Factor	15.	0.2	0.119277	0.140387	0.0944625
■ (magenta)	520	Years	20.	0.5	0.135146	0.147671	0.0125062

Figure 20: Score for 1,300 Asset Smoothing Policies: Equally Weighted Cost Mean and Cost Standard Deviation (Black—Highest score; Red—Market; Green—Five-year, 20%; Blue—Factor of 15, 20%; Magenta—20-year, 50%)

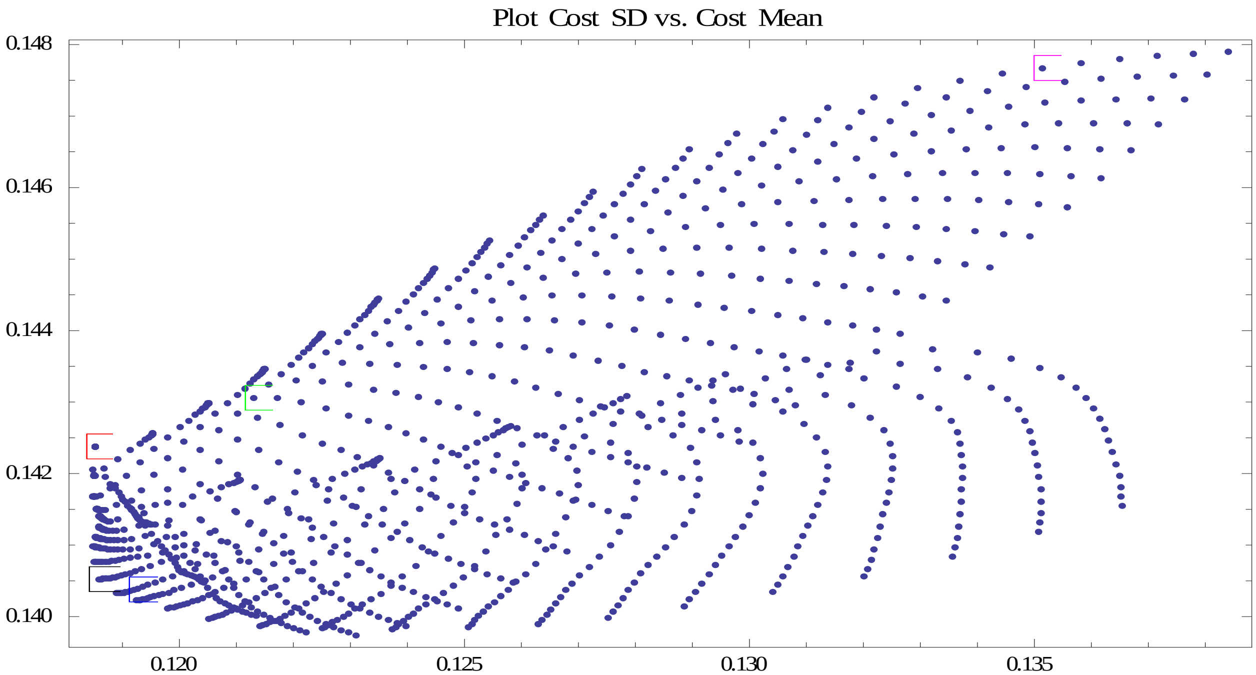
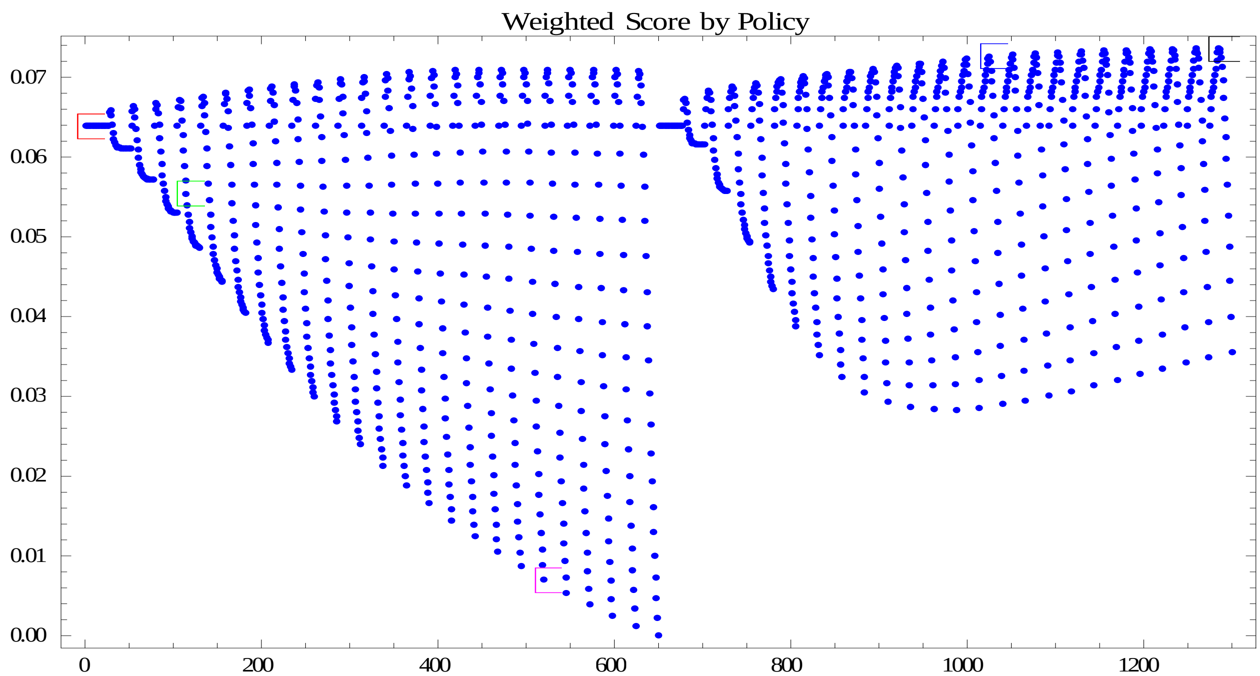


Figure 21: Mean Cost Standard Deviation vs. Mean Actuarial Cost for Candidate Smoothing Policies (Black—Highest score; Red—Market; Green—Five-year, 20%; Blue—Factor of 15, 20%; Magenta—20-year, 50%)

Figure 21 is interesting in that it plots the standard deviation of actuarial cost against the mean actuarial cost for the candidate portfolios. We note that the portfolio with the highest score (the black dot) does the best job of simultaneously minimizing mean cost and standard deviation. The 20-year/50 percent corridor policy is spectacularly unsuccessful in this regard.

Again we should note an advantage of the experimental approach we are taking: The rich interaction of the cost mean and standard deviation revealed in Figure 21 would not have been imagined without it.



	Policy	Type	Factor	Corridor	Cost Mean	Cost SD	Score
● (black)	1285	Factor	25.	0.2	0.118887	0.140324	0.0736349
■ (red)	1	Years	1.	0.	0.118522	0.142381	0.0638628
■ (green)	115	Years	5.	0.2	0.121306	0.14306	0.0553905
■ (blue)	1025	Factor	15.	0.2	0.119277	0.140387	0.0726109
■ (magenta)	520	Years	20.	0.5	0.135146	0.147671	0.007007

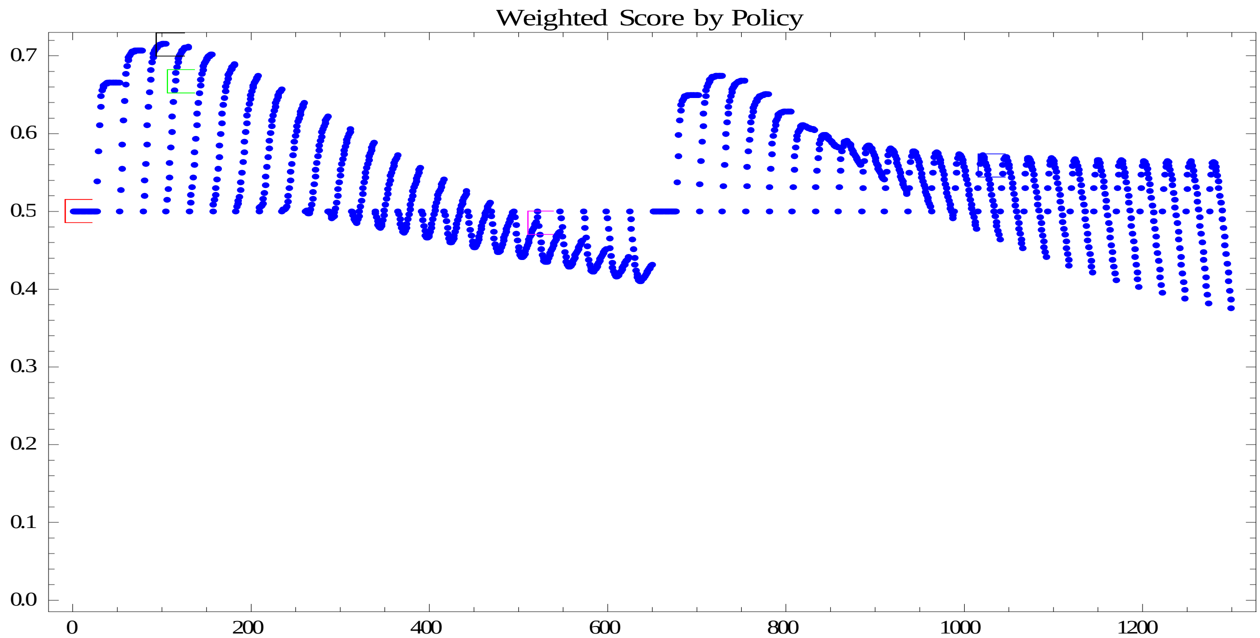
Figure 22: Mean Actuarial Cost for 1,300 Candidate Asset Smoothing Policies with 75% Weight on Cost Standard Deviation

Figure 22 shows the result if we change our judging criteria from a 50 percent/50 percent weight on mean/standard deviation of cost to a 25 percent/75 percent weight. The new weighting puts more emphasis on reducing cost variation, with the acknowledged effect of tolerating a larger mean cost. We note that there has indeed been a slight change in the selected policy, which now has a 20 percent corridor, and a corresponding change in the pattern formed by all 1,300 candidate policies.

No change in the weights used to evaluate the policies will produce a dramatic shift, because the two metrics used to evaluate smoothing policies show little variation by policy (see Figure 19). Perhaps using different metrics would be in order.

3.3.2 Cost Predictability and Asset Difference From Market Value

Now suppose we want to improve the predictability of plan cost by reducing the cost A/E SEE, but we don't want to have our smoothed asset values too far from market value. Assume we want both of these goals equally, so we weight each 50 percent. The result is Figure 23, and the highest scoring smoothing method is year-based, with gains and losses recognized over four years and with a 50 percent corridor around market value.



	Policy	Type	Factor	Corridor	Cost Mean	Cost SD	Score
●	104	Years	4.	0.5	0.01169	0.079214	0.715305
■ (red)	1	Years	1.	0.	0.045496	0.	0.5
■ (green)	115	Years	5.	0.2	0.014556	0.087258	0.667944
■ (blue)	1025	Factor	15.	0.2	0.018779	0.119115	0.558706
■ (magenta)	520	Years	20.	0.5	0.011875	0.194433	0.486042

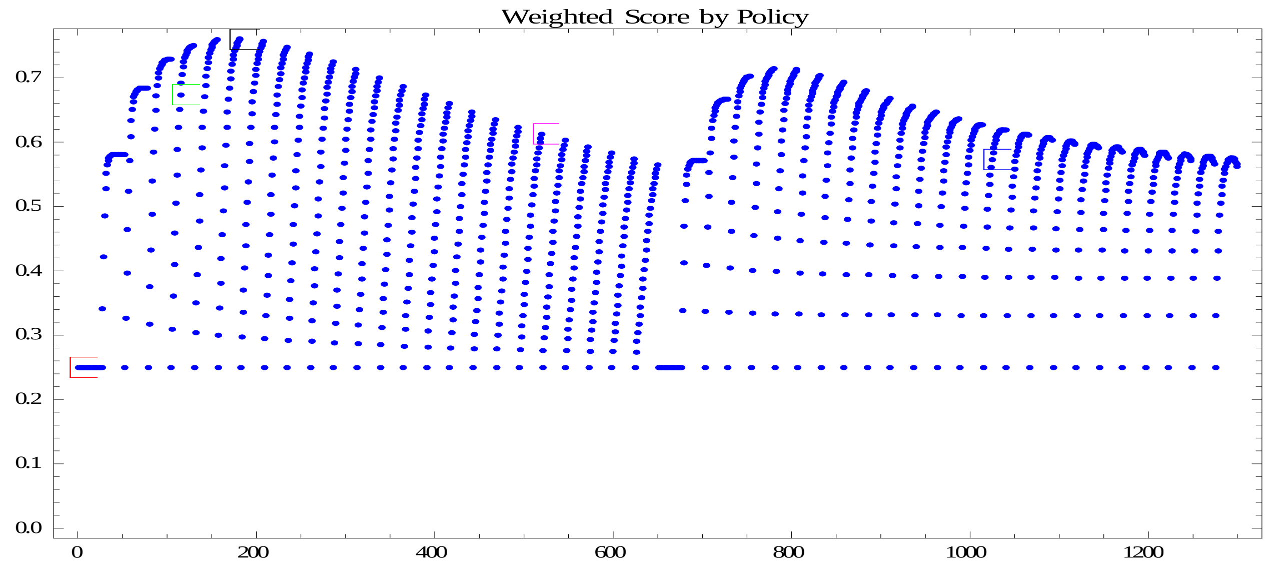
Figure 23: Score for 1,300 Asset Smoothing Policies: Equally Weighted Cost A/E SEE and AV/MV Distance (Black—Highest score; Red—Market; Green—Five-year, 20%; Blue—Factor of 15, 20%; Magenta—20-year, 50%)

Traditional five-year/20 percent corridor smoothing has the second best score; it is most like the four-year method with the top score, so it has similar dynamics. The factor of 15/20 percent corridor policy adopted by CalPERS finishes in third place. The most aggressive smoothing method—20-year smoothing with a 50% percent corridor (the magenta point)—again scores worse than the other principal policy choices, and worse than most of the policy options tested. This occurs because of the large values of the AV/MV distance allowed by the 50 percent corridor.

Figure 24 shows the result if we change our judging criteria from a 50 percent/50 percent weight on cost predictability (A/E SEE) and the difference between smoothed and market value of assets (AV/MV distance) to a 75 percent/25 percent weight. The new weighting puts more emphasis on increasing the predictability of actuarial cost from year to year, with the acknowledged effect of tolerating a larger deviation of smoothed asset value from market value. This change in scoring results in a significant change in the selected policy, with the period over which gains and losses are spread increasing from four years to seven.

3.3.3 Revisiting the Corners, and Some Conclusions

Whenever we are optimizing in a Policy Space, it is prudent to check the corners of the space. We do this by optimizing the smoothing policy with a 100 percent weight on each of the metrics successively. This is accomplished in Figure 25, which shows the best scoring smoothing policy with 100 percent weight on each of cost mean, cost standard deviation, cost predictability (A/E SEE) and AV/MV distance.



	Policy	Type	Factor	Corridor	Cost Mean	Cost SD	Score
● (black)	182	Years	7.	0.5	0.007648	0.114976	0.760548
■ (red)	1	Years	1.	0.	0.045496	0.	0.25
■ (green)	115	Years	5.	0.2	0.014556	0.087258	0.674002
■ (blue)	1025	Factor	15.	0.2	0.018779	0.119115	0.572972
■ (magenta)	520	Years	20.	0.5	0.011875	0.194433	0.612515

Figure 24: Score for 1,300 Asset Smoothing Policies: 75%/25% Weighted Cost A/E SEE and AV/MV Distance (Black—Highest score; Red—Market; Green—Five-year, 20%; Blue—Factor of 15, 20%; Magenta—20-year, 50%)

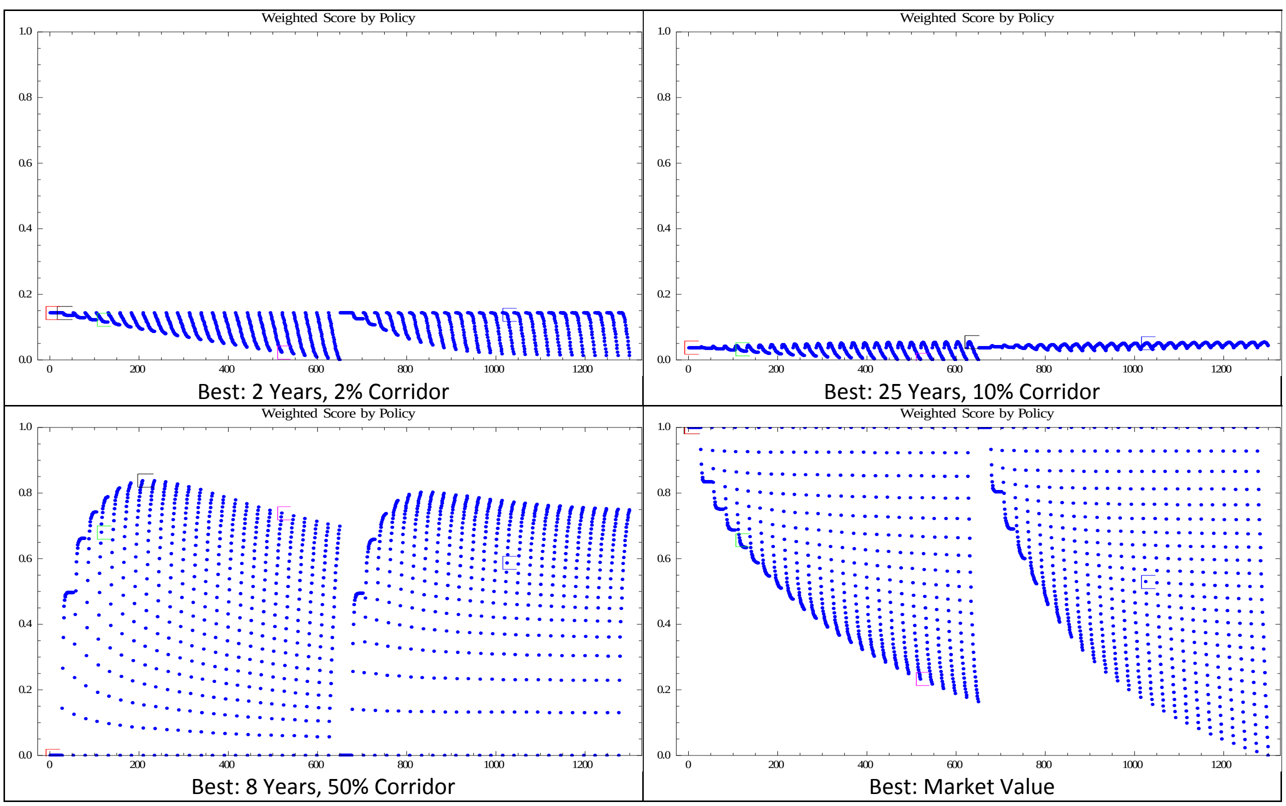


Figure 25: 100% Weight on Each of Four Metrics for 1,300 Candidate Asset Smoothing Policies

We can draw some tentative conclusions from our optimization exercise.

- Optimization using most combinations of the four metrics above seems to result in year-based, rather than factor-based, policies.

  Overall, at least for the metric weightings we have investigated, every factor-based algorithm appears to have a year-based algorithm with a better score. However, this is no reason to abandon factor-based smoothing schemes. In this paper we have focused on long-term results, 40 years into the future and beyond, well after the policy has been implemented. Optimizing over a shorter period would produce different results.

  In addition, there may be nonquantitative reasons for favoring either year-based or factor-based approaches, such as the history of the plan to date, the preferences of the responsible actuary, and the stated preferences of the plan sponsor or board. Any of these factors could indicate that we should limit the Policy Space accordingly.

- Policies with the best scores tend to have short recognition periods, with wide corridors.

  In my own practice I have favored a 20 to 30 percent corridor around market value. However, the optimizations conducted here have tended to favor wide corridors, up to the 50 percent maximum allowed. In fact, if we optimize based on a one-third weight on cost predictability and a two-thirds weight on the AV/MV distance, we get a smoothing policy with just a two-year recognition of gains and losses, but with a 38 percent corridor. This result is very much in line with the opinion of most actuaries that the corridor can be arbitrarily wide or nonexistent as long as the smoothing period is sufficiently short.

  Being stubborn, I still prefer a 20 to 30 percent corridor, but for political and public relations reasons. I cannot claim that such a corridor is technically superior to a wider one.

- Optimization with other metrics will produce different results.

  Only a small subset of the most obvious metrics has been used in the above analysis. Other metrics could be tried, in particular metrics dealing with excessive cost or unsatisfactory funding levels.

- Your mileage may differ.

  Obviously. Real pension plans differ radically from one another, with asymmetries, nonlinearities, discontinuities and legislative restrictions, each of which can have a significant impact on the scoring of any asset smoothing policy. Add to this the political realities of the plan, and the policy selected may be a long way from those considered here.

  I continue to feel, however, that the technical effort is worthwhile. When conducted on a Model Plan, as here, we can get a feeling for what is possible and what to look for. When conducted on a specific plan, we can at least determine a reasonable range of policy options to present to the decision makers, and we can discover if there are exceedingly bad options that should be excluded.

4. Summary and Conclusion

In this paper we have presented a quantitative, multifactor, simulation-based approach to selecting policies governing the operation of defined-benefit pension plans. Any policy analysis worthy of the name involves multiple conflicting goals. The approach we have adopted here is to identify the objectives we seek to achieve, find ways to measure the extent to which we have met our goals, devise ways of prioritizing and weighting our goals, and then combine weighted metrics in an overall score that will be useful in selecting among possible policies.

Note that we have said that the scores achieved by our candidate policies will be *useful* in our selection, not *determinative*. We may with good reason choose to modify, override, or even ignore the results of our exercise; it only represents a benchmark for us to refer to. Nonquantifiable factors, such as politics, public image or administrative capabilities may require that we narrow our Policy Space to an acceptable subset of all possible policies.

However, I believe strongly that the quantitative exercise is useful. It allows us to discuss options, constraints and metrics with decision makers in the broadest possible context.

We see every day the role that engineering plays in our world. From the mines in West Virginia to the space shuttle, engineers are responsible for the design, installation, operation and management of a wide range of critical and expensive systems. In my view, the actuarial profession is responsible for the design, installation, operation and management of financial security systems. Millions of people depend on the systems we design, some for their very lives. Billions of dollars are at stake.

Don’t our systems deserve the same engineering effort as all the other systems in our lives?