11-3-2015

Assessing the Impact of Investment Shortfalls on Unfunded Pension Liabilities: The Allure of Neat, but Faulty Counterfactuals

Robert M. Costrell
University of Arkansas, Fayetteville

Assessing the Impact of Investment Shortfalls on Unfunded Pension Liabilities: The Allure of Neat, but Faulty Counterfactuals

Robert M. Costrell, University of Arkansas, Department of Education Reform

November 3, 2015*

Keywords: public pensions, unfunded liabilities

Abstract

In this paper I provide a methodological critique of the conventional method for assessing the impact of investment shortfalls and other contributors to unfunded pension liabilities, and offer a methodologically sound replacement with substantive policy implications. The conventional method – simply summing the annual actuarial gain/loss figures over time – provides a neat, additive decomposition of the sources of the rise in the Unfunded Accrued Liability (UAL). In doing so, however, it implicitly assumes that in the counterfactual exercise, amortization would adjust dollar-for-dollar with the interest on additional UAL. That is, even if the total (and average) shortfall from covering interest is substantial, the marginal shortfall is assumed to be zero. This is not how contribution shortfalls arise under funding formulas typically used by public plans in the United States. Using the actual funding formula in the counterfactual – with contribution shortfalls on the margin -- leads to much higher estimates of the UAL impact of investment shortfalls than the conventional method. The reason is that there are large interactions over time between investment shortfalls and marginal contribution shortfalls. The conventional counterfactual implicitly assumes away these interactions. The resulting additivity is alluring, but illusory.

The conventional method also leads to untenable results on other UAL-drivers. Most striking is the implication that the cumulative UAL impact of pension obligation bonds (POB's) is no different from the initial impact of receiving the proceeds, independent of the return (actual or assumed) on those proceeds.

The underlying problem with the conventional framework is that it has emerged without careful attention to the counterfactual scenarios it is meant to address. This paper provides explicit and internally consistent counterfactuals to better understand the conventional method and its flaws, as well as the reasons for using instead the actual amortization formula in the counterfactual. Mathematical methods are used to illuminate the theoretical issues that lie behind any simulations.

The analytical results are illustrated empirically with an adapted version of the actuarial history of the Connecticut State Teachers' Retirement System (CSTRS), FY00-FY14. The example is instructive because it is a highly underfunded system, notable for its high (and unreduced) assumed rate of return (8.5 percent), as well as its use of $2 billion in POB proceeds to reduce the UAL in FY08, just before the market crash.

* This paper will be presented at the Fall Research Conference of the Association for Public Policy and Management, November 12, 2015, Miami, FL. A precursor of this paper was presented at the Association for Education Finance and Policy, 40th Annual Conference, February 26, 2015, Washington, DC. I would like to acknowledge the early support of StudentsFirst and EdBuild for my work on Connecticut, and the particularly helpful comments on this paper by Andrew Biggs and Josh McGee.

Assessing the Impact of Investment Shortfalls on Unfunded Pension Liabilities: The Allure of Neat, but Faulty Counterfactuals

Robert M. Costrell, University of Arkansas, Department of Education Reform
November 3, 2015.

Introduction and Summary

What has been the impact of investment shortfalls from the assumed return on public pension funds, over extended periods? How has this impact compared with that of shortfalls from liability assumptions and contribution shortfalls? These questions have been framed as the relative contributions of the investment, liability, and contribution shortfalls to the rise of a plan's Unfunded Accrued Liability (UAL). That framework, in turn, has been based on the actuarial gain/loss statement in a plan's annual valuation report, which parcels out the year's rise in UAL among the UAL-drivers. In the conventional multi-period UAL analysis (Munnell, et. al., 2015), these one-year gain/loss attributions are simply added up over time. In this paper, I closely scrutinize the methodological basis for this procedure and find it wanting. I propose a methodologically sound replacement with substantive policy implications.

The critical step in the conventional methodology is the last step -- adding up the annual gain/loss components of the rise in UAL over a period of years to arrive at a cumulative breakdown of the multi-year rise in UAL. While this procedure provides a neat, additive decomposition of the rise in UAL, it does so by effectively assuming away important interactions between investment shortfalls and contribution shortfalls, through its implicit assumption regarding the counterfactual behavior of amortization payments.

Specifically, the method implicitly assumes that the one-year impact of any UAL-driver elicits in the following year a dollar-for-dollar increase in amortization payments for the interest on the additional UAL. That is, even if the total (and average) shortfall from covering interest is substantial, the marginal shortfall is assumed to be zero. This is not how contribution shortfalls arise under funding formulas typically used by public plans in the United States. In these formulas amortization is proportional to the UAL (and its interest), so the average and marginal shortfall (if there is one) is the same.

Using the actual funding formula in the counterfactual can lead to much higher estimates of the UAL impact of investment shortfalls than the conventional method. The reason is that there are large interactions over time between investment shortfalls and marginal contribution shortfalls. The conventional counterfactual implicitly assumes away these interactions, with a resulting additivity that is alluring, but illusory.

It is important to get the impact of investment shortfalls right, not only for academic reasons, but also for practical policy purposes. In the aftermath of the 2007-09 market crash, many funds reduced their assumed returns, but not by much (0.27 percent, on average, Biggs, 2015). Several years later, despite generally strong market returns, the question returns, as in many cases the cuts have proven insufficient to prevent a further rise in UAL's. Thus, pension funds – and the general public – need to understand the historical impact of investment shortfalls to better inform the choice of rate reduction going forward. By underestimating the past impact, the conventional method implies a low-ball answer to the question of how much to cut the assumed return.

The conventional method also leads to untenable results on other UAL-drivers. Most striking is the implication that the cumulative UAL impact of pension obligation bonds (POB's) is no different from the initial impact of receiving the proceeds. If a state issues POB's to "pay down," say, $2 billion in a fund's UAL (as did Connecticut for its teachers' retirement system in 2008) the plan's annual gain/loss statement will record a $2 billion reduction in UAL due to the POB. In the conventional multi-year UAL analysis, the cumulative impact of that POB remains at $2 billion for every year thereafter. In this framework, there is no impact attributed to the POB of the assumed or actual return on the proceeds invested in the fund (let alone the debt incurred and debt service paid by the state or other issuer of the POB's). This treatment of POB's in the conventional multi-year UAL analysis is, perhaps, the canary in the coal mine, raising questions of the attribution to the other UAL-drivers as well.

The underlying problem with the conventional framework is that it has emerged without careful attention to the counterfactual scenarios it is meant to address, such as, "What if investment returns had met assumptions, but other factors had remained unchanged?" In general, such scenarios take as the baseline the time series of all UAL-drivers at their actual values (generating the actual rise in UAL), and assesses the counterfactual impact of setting the series of each driver, one at a time, at the value that generates no rise in UAL. This is what policy-makers have in mind for the "what-if" scenarios.

Concretely, it is the specification of the UAL-drivers that defines the exercise. Formally, that specification defines what is exogenous. In the conventional decomposition (based on the annual gain/loss statement), a key driver is the total dollar amount of the contribution shortfall – the difference between interest on the prior year's UAL and the current year's amortization. Taking this as exogenous for a one-year change is innocuous, but not for multiple years. For if so, then current amortization varies dollar-for-dollar with the interest on the prior year's UAL. That is, the implicit assumption for is that in each year, amortization equals interest on the counterfactual UAL for the prior year, minus the current year's actual contribution shortfall, impounded in an exogenous constant term. In this way, the total contribution shortfall reproduces the actual, but the marginal shortfall is set to zero. In actual practice, the marginal and average shortfall (if there is one) are typically the same. This is the crux of the problem.

A second issue with the conventional framework is that the counterfactual time series of asset values is endogenous, but is implicitly treated as exogenous. One can readily rectify this, treating asset values as endogenous. Otherwise, the counterfactuals are not internally consistent. Although this has no effect on the measured impact of investment shortfalls, it can significantly improve the measured impact of POB's.

The structure of the paper is as follows. First, I review the analytics behind the conventional method. I then briefly discuss the implications for this procedure of smoothed vs. market valuation. Next, I dissect the key issue of the implicit designation of what is exogenous under the conventional method, and how its counterfactual model of amortization is at odds with actual formulas. I then informally draw the implications for the conventional method's under/overestimation of the impact of investment shortfalls and other drivers. Before continuing, I correct the exogenous treatment of asset values so that the analysis proceeds with internally consistent counterfactuals.

Next, to fully illuminate the implications of what is taken as exogenous, I mathematically contrast the two extremes: the conventional framework's exogenous treatment of the total contribution shortfall vs. exogenous amortization payments. I then illustrate my proposed method of using actual amortization formulas, which, in the case of underfunding, lies between the two extremes previously analyzed. To complement our understanding of the varying impact of investment shortfalls on the UAL under the different amortization assumptions, I show that the difference is mirrored in the impact on cumulative contributions, with the total impact – paid (amortization) and not-yet-paid (UAL) -- essentially invariant. Finally, for completeness sake, I fill out the accounting for POB's, including the debt and debt service of the bonds' issuer.

It is important to note that when using the amortization formulas, the UAL impact of individual drivers will exceed the actual rise in the UAL, in the case of underfunding. That is, there are substantial interactions that are effectively assumed away under the conventional framework. The reason is that if investment shortfalls generate additional UAL, the interest on that additional UAL will, in underfunded systems, also be underfunded, adding further to the UAL. These dynamic interactions between investment shortfalls and contribution shortfalls will disrupt additivity. I argue that the adding-up problem is more aesthetic than policy-relevant. It is more important to get the individual impacts right, taken one at a time, since that is the usual policy interest. If the policy interest is in the simultaneous impact of multiple drivers, that can be readily modeled, reflecting the interactions, rather than effectively assuming away the interactions so that the individual impacts artificially add up.

Throughout the paper, I supplement the analytical results with illustrations adapted from the Connecticut State Teachers' Retirement System (CSTRS), FY00-FY14. The CSTRS case is of interest in its own right for several reasons. The system's pension funding difficulties rival those of more well-known cases such as California, New Jersey, Illinois and Pennsylvania. Unlike other systems, it has not reduced its assumed rate of return from 8.5 percent. Finally, it has made use of POB's to reduce its reported UAL.

Thus, it is an interesting case to illustrate the effect of using a more sound method of assessing the impact of investment shortfalls on the rise in UAL, as well as the more complete impact of POB's and other UAL-drivers.

The Sources of the Rise in UAL: The Conventional Method

The starting point is the decomposition of the annual rise in UAL presented in a plan's valuation report. The sources of the rise can be categorized into (i) investment shortfalls of market return from assumed return; (ii) shortfalls on the liability side due to changes in and deviations from actuarial assumptions, as well as benefit changes; (iii) contribution shortfalls relative to interest on the UAL;[1] and (iv) other events, such as the issuance of POB's (pertinent for CSTRS, as well as other plans).

Let us consider this decomposition formally, with some slightly simplified math:[2]

(1) $UAL_t \equiv L_t - A_t$

(2) $L_t = (L_t - L^e_{t|t-1}) + L^e_{t|t-1} = (L_t - L^e_{t|t-1}) + (1+r^*)L_{t-1} + NC_t - B_t$

(3) $A_t = (1+r_t)A_{t-1} + AMT_t + NC_t - B_t + POB_t$,

where

A_t = assets[3]
L_t = accrued liabilities
$L^e_{t|t-1}$ = expected accrued liabilities in period t, as of t-1, under actuarial assumptions
NC_t = normal cost
B_t = benefit payments
AMT_t = employer contributions in excess of NC, credited toward amortization
r^*, r_t = assumed and actual return on investment[4]
POB_t = proceeds from pension obligation bonds issued in period t.

Substituting and simplifying, we have the standard actuarial gain/loss result, decomposing the change in the UAL from t-1 to t:

---

[1] Munnell, et. al., 2015 rightly point out that this is the UAL-relevant contribution shortfall, rather than the commonly cited shortfall between actual contributions and the ARC, as the ARC may not cover interest.
[2] I assume cash flows and annual liability accruals are made at the end of the year, rather than the usual actuarial assumption of mid-year, thus excluding the associated half-year interest on each.
[3] These are market asset values, as are the return on assets, r_t. The implications of using smoothed asset values, while tangential to the subject of this paper, are briefly discussed below.
[4] We take the assumed return r* as constant, which is accurate for CSTRS. For cuts in the assumed return, the modifications are straight-forward.

$$(4)\ \Delta_{t-1,t}UAL = (L_t - L^e_{t|t-1}) + (r^* - r_t)A_{t-1} + (r^*UAL_{t-1} - AMT_t) - POB_t.$$

The first term on the RHS is the loss from liabilities exceeding expectations (i.e. adverse deviations from liability assumptions, changes in those assumptions, or changes in benefits); the second term is the loss from investment returns falling short of expectations; the third term is the loss from amortization payments' failure to cover accrued interest on the UAL; and the fourth term is the reduction in UAL from POB's.[5]

For the one-year change in UAL, this is an attractive formulation: it unambiguously allocates the change among liability shortfalls, investment shortfalls, contribution shortfalls, and POB's in a way that adds up to the actual change in the UAL. Specifically, note that there are no interactions among these drivers of the rise in the UAL that would disrupt additivity; some terms depend on prior values of UAL and A, but these can be taken as pre-determined. Thus, each source represents a well-defined counterfactual that is independent of one another: "How would the UAL in period t (and its rise from period t-1) differ from the actual if in period t investment assumptions had been met [or liability assumptions had been met, or amortization had covered interest, or no POB's had been issued]?" If all four conditions held simultaneously, the rise in UAL would be zero. Since there are no interactions, the separate counterfactual impacts sum to the actual rise in the UAL.

How, then, to allocate the rise in UAL over multiple periods, from period 0 to T? The conventional approach (put forth by Munnell[6] and her colleagues at the Center for Retirement Research at Boston College) is to simply add up each component of (4), year-by-year, to allocate the total change in the UAL among its various sources:

$$(5)\ \Delta_{0,T}UAL = \Sigma_{t=1}^{T}(L_t - L^e_{t|t-1}) + \Sigma_{t=1}^{T}(r^* - r_t)A_{t-1} + \Sigma_{t=1}^{T}(r^*UAL_{t-1} - AMT_t) - \Sigma_{t=1}^{T}POB_t$$

$$= \Delta_{0,T}UAL^{L:\Sigma} + \Delta_{0,T}UAL^{r:\Sigma} + \Delta_{0,T}UAL^{C:\Sigma} + \Delta_{0,T}UAL^{B:\Sigma}.$$

I label each of the four terms as the contribution to the rise in UAL from each of the four UAL-drivers, as attributed by the conventional method: $\Delta_{0,T}UAL^{L:\Sigma}$, etc., where Σ denotes the simple summation from the annual gain/loss statements. We use C in the third term to denote contribution shortfalls, $(r^*UAL_{t-1} - AMT_t)$. Using this framework, the Munnell team estimates that 60.4 percent of the rise in UAL from 2001 to 2013 for the 150 plans in Boston College's Public Plans Database was due to annual investment shortfalls, the second term of (5), $\Delta_{0,T}UAL^{r:\Sigma}$.

[5] Standard UAL accounting omits the POB debt itself, carried by another party (e.g. the state). I will provide a full accounting in a later section, below.

[6] Munnell, et. al, 2015, "How Did State/Local Plans Become Underfunded?"

I will illustrate the conventional method and alternatives to be considered with the case of the Connecticut State Teachers Retirement System, CSTRS.[7] I use an adaptation of the time series for FY00-FY14. The main difference between this adaptation and the actual series of CSTRS is the use of market asset values rather than smoothed asset values – a difference with consequences discussed below. I also exclude the midyear return for cash flows and accruals from the growth in assets and liabilities, respectively, a relatively minor difference, to keep the math simple and focused on the issues at hand.[8] The net result of these adaptations is a rise in the UAL of $8.4 billion from FY00 to FY14, rather close to the actual rise of $8.6 billion.

Figure 1 depicts the decomposition of the $8.4 billion rise under the conventional method. I find that the conventional method would attribute $5.0 billion (about 60 percent) to investment shortfalls (red line) from the assumed return of 8.5 percent. Contribution shortfalls (purple line) account for another $3.3 billion (around 40 percent). I have broken out liability losses into those due to deviations from or changes in assumptions and the FY08 benefit hike for COLA's.[9] These were both offset by the $2 billion POB, also issued in FY08. (The dotted lines will be discussed later.)

[FIGURE 1 APPROXIMATELY HERE]

In Figure 1, note that the cumulative impact of the POB is unchanged after its FY08 impact: the solid orange line is flat at -$2 billion. Similarly, the cumulative impact of the FY08 benefit enhancement is unchanged thereafter: the brown line is flat. These are the most visible manifestations of the fact that under the conventional method, the one-year impact of *any* contributor to the rise in UAL, as specified in (4), is also the cumulative impact thereafter, in (5). There is no further impact from the interest on the initial UAL impact or shortfalls from the assumed return. The implicit assumptions of the conventional method that underlie this result will be explained below.

Measured Impact of UAL-Drivers Under Smoothed vs. Market Asset Values

In many respects, the use of smoothed vs. market asset values makes little difference over long periods; the change in asset values and UAL is not much different for CSTRS over the period in question. However, the attribution of that rise can differ materially between the investment and contribution shortfalls. In the CSTRS case, I find that

[7] This study uses the series of biennial valuation reports from FY02 – FY14.

[8] One other difference is the exclusion of an actuarial adjustment made in FY08 for the interest on the integration of funds previously segregated for the COLA.

[9] Specifically, the FY08 legislative package (PA 07-186) which authorized the POB's also changed the terms of the COLA and converted it from a fund-contingent benefit to a guaranteed one. This resulted in a $1.151 billion hike to the accrued liability, effectively spending over half the proceeds of the POB's, instead of using it all to "pay down" the UAL.

smoothed asset values swings about 7 additional percentage points of the rise in UAL away from the impact of contribution shortfalls to the impact of investment shortfalls.

The reason is not hard to explain. Over the period in question, smoothed asset values have, on average, exceeded market asset values for CSTRS, and, consequently, the smoothed UAL's have, on average, been lower than the market UAL's. This means the assumed interest on the UAL was also lower under smoothing, and, thus, the contribution shortfall from assumed interest was lower, too. Conversely, since smoothed asset values were higher, so was the assumed investment income, and, therefore, the investment shortfall was greater.
The point here is that the allocation of the rise in UAL is sensitive to smoothing. The difference in average asset values is about 4 percent over this period, but this corresponds to an average difference in UAL's of about 7 percent and that results in the 7 point swing in the attribution between investment and contribution shortfalls. Although the asset-smoothing mechanism may well serve its designated purpose of smoothing out contributions, it can significantly distort our understanding of the rise in the UAL. Consequently, I focus on market values in the analysis of this paper.

What is Exogenous Under the Conventional Method?

To understand the underpinnings of the conventional method, one must consider closely the question of what is exogenous. The unstated premise of the decomposition of *any* variable's growth among the drivers of that growth is that the series of drivers are to be considered exogenous. That is, if $\Delta_{0,T}Y$ is to be attributed to the time series of drivers X^a_t, X^b_t …, t = 1,…,T, then any counterfactual on those drivers must be considered independent of one another.[10] Specifically, the portion of $\Delta_{0,T}Y$ attributed to the driver series X^a_t is answered by the question, "How much different would $\Delta_{0,T}Y$ have been, had X^a_t taken the counterfactual values $X^a_t{}^*$ that would have contributed nothing to $\Delta_{0,T}Y$, instead of its actual values X^a_t, holding the other drivers constant at their actual values?" To pose this "what if" question, one must assume that each series of drivers can be set independently of the others in the counterfactual, i.e. that each driver can be considered exogenous.

In the present case, this means that the conventional decomposition of UAL growth takes as exogenous the series of drivers $(L_t - L^e_{t|t-1})$, POB_t, $(r^* - r_t)A_{t-1}$, and $C_t \equiv (r^*UAL_{t-1} - AMT_t)$. As stated above, the variable C_t denotes the contribution shortfall. The key assumption implicit in the conventional allocation (5) is that the series of contribution shortfalls C_t is exogenous. That is what allows one to interpret each term of (5) pertaining to the other drivers as the answer to the question, "How much different would

[10] This is *not* to say that their impact on Y must be independent of one another, only that the values of the X's can be considered independent in the counterfactuals. The question of interactions in their impact on Y (and, hence, of additivity) is separate.

$\Delta_{0,T}$UAL have been, had (say) (L_t - $L^e_{t|t-1}$) taken the counterfactual values of zero that would have contributed nothing to $\Delta_{0,T}$UAL (i.e. liability assumptions had been met), instead of its actual values (L_t - $L^e_{t|t-1}$), holding the other drivers constant at their actual values, ***including C_t***?"

That is, these counterfactuals -- implicit in the conventional method -- take the series of contribution shortfalls C_t as exogenous at the actual observed levels, independent of the counterfactual values assigned to the other drivers, and (importantly) their corresponding impact on UAL. That means in the counterfactuals for the other drivers' impact, it is implicitly assumed that

(6) $AMT_t = r^*UAL_{t-1} - C_t$,

such that C_t is held constant at the actual level and AMT_t varies dollar for dollar with the interest on the UAL, as it reflects the counterfactual impact of the other drivers.

The form of (6) – implicit in the conventional method's counterfactuals for the other drivers – is important to understand. Amortization in period t is linear in the interest on UAL, with coefficient one (i.e. varies dollar-for-dollar) and a constant term equal to the negative of that year's actual total dollar contribution shortfall. Figure 2 graphs equation (6), with the counterfactual values of AMT_t on the vertical axis and the counterfactual values of r^*UAL_{t-1} on the horizontal. Equation (6) is parallel to the 45° line ($AMT_t = r^*UAL_{t-1}$), shifted down by a constant equal to the actual contribution shortfall, C_t.

[FIGURE 2 APPROXIMATELY HERE]

This differs significantly from the typical amortization formula, under which amortization is proportional to the UAL and its interest. That is, unlike (6), the typical formula has zero constant and generates contribution shortfalls (surpluses) through a coefficient on the interest that is less than (greater than) one. Specifically, the formula commonly employed is "level percent [of payroll]." With assumed growth rate g, the basic form is:[11]

(7) $AMT_t = UAL_{t-1}(r^*-g)/\{1 – [(1+g)/(1+r^*)]^N\}$,

where N is the remaining amortization period (closed or open). We can then write

(8) $AMT_t = \alpha_t r^*UAL_{t-1}$, where $\alpha_t = [(r^*-g)/r^*]/\{1 – [(1+g)/(1+r^*)]^N\}$.

[11] As before, we simplify by ignoring the actuarial conventions regarding mid-year cash flows. Another departure from typical actuarial practice is that AMT will often be based on 2 (or more) years lagged UAL, due to timing of actuarial reports.

As stated above, this differs from (6) by virtue of a zero constant and a coefficient on interest α_t that can be less than (greater than) one.[12] For example, with r* = 0.085 and g = 0.0375, $\alpha_t < 1$ for N ≥ 19. That is how contribution shortfalls are typically generated. Figure 2 graphs equation (8). It is a clockwise rotation of equation (6), through their common point, which is the actual value of (r^*UAL_{t-1}, AMT_t).

Munnell's team finds that contribution shortfalls are more common than not,[13] but the mechanism by which those shortfalls are generated is misrepresented by (6), as embedded in the conventional method. Both (6) and (8) model the same total shortfall, ($r^*UAL_{t-1} - AMT_t$), and the same average shortfall, ($r^*UAL_{t-1} - AMT_t)/r^*UAL_{t-1}$ for the actual values AMT_t and r^*UAL_{t-1}. However, they differ importantly in the marginal shortfall for the counterfactual values, $\Delta(r^*UAL_{t-1} - AMT_t)/\Delta(r^*UAL_{t-1})$ – one minus the slope of the counterfactual line. Under (6) it is zero, and under (8) it is ($1 - \alpha_t$). <u>This is the crux of the problem with the conventional method – the main source of its misleading assessments of driver impacts -- as I will explain.</u>

In some cases, budgeters do not actually fund the amortization formula in place, but instead contribute "what they can afford." This is an elusive concept, to be sure, governed by tax revenues and non-pension expenditures that are considered non-discretionary. But it seems clear that in such cases the amortization payments are even less responsive to the UAL than the amortization formula. In effect, AMT_t is exogenous, corresponding to the horizontal line in Figure 2.

<u>What are the implications of Taking C_t as Exogenous?</u>

Before completing my formal analysis, I can sketch the implications of the conventional method's treatment of C_t as exogenous, i.e. modeling AMT_t by (6) instead of (8). Consider the measured impact of investment shortfalls, liability shortfalls and POB's.[14]

The main issue here is investment shortfalls. In short, if $\alpha_t < 1$, the conventional method will underestimate the impact of investment shortfalls (and, conversely, if $\alpha_t > 1$, it will overestimate the impact). The reason is that under the conventional method, if an

---

[12] The coefficient α_t can vary over time as plans change assumptions on g or N changes (either a policy change for the open interval, or, automatically, with the passage of years for a closed interval). Note that α_t cannot be less than one with g = 0, i.e. "level dollar" amortization. Conversely, with open interval amortization, α_t can be less than one indefinitely with sufficiently high values of g and N, and that has been the case for some plans with 30-year rolling horizons, as Munnell, et. al. point out.

[13] They calculate that contribution shortfalls account for 23.7 percent of the rise in UAL's from 2001 to 2013 for the 150 plans in Boston College's Public Plans Database.

[14] The measured impact of the series of contribution shortfalls itself is independent of the counterfactual modeling of AMT_t using (6) or (8) or any other variant. That impact is measured by the difference between the actual $\Delta_{0,T}UAL$, where all drivers take their actual values, and the counterfactual $\Delta_{0,T}UAL$ where $C_t^* = 0$, i.e. $AMT_t = r^*UAL_{t-1}$, or, equivalently, $\alpha_t^* = 1$. This difference does not involve (6) or (8).

investment shortfall occurs, generating an increment to the UAL and to the interest on the UAL the following year, the amortization is implicitly assumed to cover the incremental interest dollar-for-dollar, adding nothing to the contribution shortfall. If, instead, the true amortization process only partially covers additional interest ($\alpha_t < 1$), then there will be an additional contribution shortfall, induced by the investment shortfall.

This can be understood in terms of equation (5). Under the conventional method, investment shortfalls generate UAL impacts directly through the second term of (5), $\Sigma_{t=1}^{T}(r^* - r_t)A_{t-1},$ with no indirect impact through the third term, $\Sigma_{t=1}^{T}(r^*UAL_{t-1} - AMT_t)$, which remains unchanged by implicit assumption (6). But if amortization is actually governed by (8) with $\alpha_t < 1$, then the marginal contribution shortfall is $(1 - \alpha_t) > 0$, so there will be an additional contribution shortfall induced by the investment shortfall, reflected in the third term of (5). <u>In short, there is a significant interaction between investment shortfalls and an amortization regime that generates contribution shortfalls.</u> This disrupts additivity, but that is the way amortization formulas work. To implicitly assume away the interactions by virtue of (6) yields an appealing additivity, but at the price of underestimating the impact of investment shortfalls in a regime of underfunding.

The conventional method's treatment of C_t as exogenous also explains the puzzling time pattern of other drivers' impact. Consider the FY08 COLA enhancement that raised the UAL that year by $1.15 billion. As Figure 1 shows (brown line), under the conventional method, the cumulative impact remains unchanged thereafter. There is no further measured impact from the potential accumulation of interest on the additional UAL because the implicit assumption is that additional amortization covers that interest. If, instead, the amortization regime does not fully cover additional interest ($\alpha_t < 1$), then there would be additional contribution shortfalls, induced by the benefit enhancement. Again, the potential interaction between the benefit enhancement and an amortization regime that underfunds on the margin is assumed away. The same is true for any given year's deviation from liability assumptions.

This logic also explains the one-shot time pattern of POB's impact. As Figure 1 illustrates, under the conventional method, the reduction in UAL from the POB proceeds is recorded in the year of issue ($2 billion in FY08 for CSTRS) and remains unchanged in the cumulative accounting going forward. One might think that the subsequent earnings on the proceeds of the POB contribute to the further diminution of the UAL attributable to the POB. Indeed, that definitely seems implicit in the arguments heard for POB's, namely the arbitrage gains from the returns r* that are assumed to be earned on the proceeds. Again, the explanation of those returns' omission from the subsequent impact of the POB lies in the conventional method's assumption for amortization. As the POB reduces the UAL (by, say $2 billion), the interest on the UAL is also reduced and, under the conventional method, amortization payments are reduced dollar-for-dollar. This cancels out the assumed return on the proceeds of the POB.

There is, however, another aspect to the POB impact that is also missed by the conventional method, due to another problem, unrelated to amortization. That is the inconsistent treatment of asset values in the second term of (5) (investment shortfalls) for the counterfactuals, including the one for POB's. Before continuing the formal analysis of the treatment of amortization, I need to fix this problem, so that the counterfactuals are internally consistent with respect to asset values.

<u>Consistent Counterfactuals with Exogenous Contribution Shortfalls</u>

To construct consistent counterfactuals, we want to examine the impact of each of the UAL-drivers while simultaneously satisfying system (1) – (3), which implies (4) and (5). The counterfactuals answer the question: <u>*"What is the impact of each UAL-driver, assuming the other drivers are unchanged?"*</u>

Specifically, we start with each of the four series of UAL-drivers – (L_t - $L^e_{t|t-1}$), (r^* - r_t), $C_t \equiv$ (r^*UAL_{t-1} - AMT_t), and POB_t, t = 1,…T – set to their actual values, so the baseline UAL_T is the actual (as is its rise from UAL_0). We then take each driver one at a time, setting it to a series of zeros, and compare the resulting counterfactual UAL_T to the actual (or, equivalently, compare its rise from UAL_0) to find that driver's impact.[15]

In doing so, there is a further flaw implicit in the conventional method, and that involves the term (r^* - r_t)A_{t-1} in (5). Although this series – the dollar value of the annual investment shortfall – is implicitly treated as exogenous in (5), the intent behind this component is clearly to assess the impact of the shortfall in the investment return, (r^* - r_t). That is, the "what if" counterfactuals for the non-investment shortfalls compare the UAL if one of those series of drivers equals zero, which means the lagged asset series A_{t-1} cannot be taken as exogenous -- as implied by the term $\Sigma_{t=1}^{T}(r^* - r_t)A_{t-1}$ in (5). That is, in setting (say) the (L_t - $L^e_{t|t-1}$) series at zero and comparing the resulting value of (5) with the actual value of (5), one cannot legitimately hold the second term constant.

Instead, the counterfactual values of A_{t-1} must be modelled as in (3), using the counterfactual values of AMT_t. Here I maintain the conventional method's implicit assumption of exogenous contribution shortfalls, $C_t \equiv$ (r^*UAL_{t-1} - AMT_t), which implies the amortization regime (6). That is, internal consistency is satisfied by modeling the

---

[15] An alternative set of consistent counterfactuals (call them zero-base) would start with each of the four series of UAL-drivers set to a vector of zeros, so there is no rise in UAL, as shown in (5). We would then take each driver one at a time, setting it to the actual series to find the impact on UAL. These counterfactuals answer the question: <u>*"What is the impact of each UAL-driver at its actual value, assuming the other drivers are zero?"*</u> Arguably, the version in the text is the "what if" that policy people have in mind. One might also argue that these comprise a more compelling set of counterfactuals than the zero-base ones, since the zero baseline is so far removed from the actual. The implications of the zero-base counterfactuals are explained in a note below.

counterfactual values of A_t using (3) and (6), and of course the counterfactual series UAL_t that is implied by the assumed values of the drivers.[16]

So how do the UAL-impacts under these consistent counterfactuals compare with the uncorrected conventional method? The answer is that the impact of the investment shortfall equals that of the uncorrected method, but the others do not.

The result can be seen by examining the counterfactuals with (5). For the investment shortfall, denoting the counterfactual by superscript r:C (for exogenous C), we have:

$$(5^{r:C})\quad \Delta_{0,T}UAL^{r:C} = [\Sigma_{t=1}^{T}(L_t - L^e_{t|t-1}) + \Sigma_{t=1}^{T}(r^* - r_t)A_{t-1} + \Sigma_{t=1}^{T}C_t - \Sigma_{t=1}^{T}POB_t]$$
$$- [\Sigma_{t=1}^{T}(L_t - L^e_{t|t-1}) + \Sigma_{t=1}^{T}(r^* - r^*)A^{r:C}_{t-1} + \Sigma_{t=1}^{T}C_t - \Sigma_{t=1}^{T}POB_t] = \Sigma_{t=1}^{T}(r^* - r_t)A_{t-1} = \Delta_{0,T}UAL^{r:\Sigma}$$

The first bracketed term is the actual rise in UAL, using actual values of the drivers, and the second bracketed term is the counterfactual rise in UAL, using actual values for all drivers except $r_t = r^*$. The difference is $\Sigma_{t=1}^{T}(r^* - r_t)A_{t-1}$, with actual values of A_{t-1}, as in the conventional allocation. The endogeneity of the series $A^{r:C}_{t-1}$ under this counterfactual is irrelevant, since it is multiplied by zero. Thus, $\Delta_{0,T}UAL^{r:C} = \Delta_{0,T}UAL^{r:\Sigma}$.

The endogeneity of A_{t-1}, however, does matter for the other drivers, notably the POB:

$$(5^{B:C})\quad \Delta_{0,T}UAL^{B:C} = [\Sigma_{t=1}^{T}(L_t - L^e_{t|t-1}) + \Sigma_{t=1}^{T}(r^* - r_t)A_{t-1} + \Sigma_{t=1}^{T}C_t - \Sigma_{t=1}^{T}POB_t]$$
$$- [\Sigma_{t=1}^{T}(L_t - L^e_{t|t-1}) + \Sigma_{t=1}^{T}(r^* - r_t)A^{B:C}_{t-1} + \Sigma_{t=1}^{T}C_t] \neq - \Sigma_{t=1}^{T}POB_t = \Delta_{0,T}UAL^{B:\Sigma},$$

because the series $A^{B:C}_{t-1} \neq A_{t-1}$. Specifically, in the year of the POB, if it had not been issued, the counterfactual value of $A^{B:C}_t$ would be reduced by the full POB amount. Therefore, in period t+1 the POB's absence would have reduced any investment shortfalls from $(r^* - r_{t+1})$. That is, to state the result directly, the POB's favorable UAL impact in period t is partially offset in t+1 by the investment shortfalls on the proceeds.

CSTRS provides a dramatic example. Immediately following the POB issue of $2 billion, the fund lost 17.84 percent in FY09, 26.34 percentage points below the assumed return. Thus, the investment shortfall on the $2 billion was $527 million that year, offsetting a good portion of the prior year's UAL reduction from the POB proceeds. This can be seen directly on Figure 1, comparing the FY09 point on the dotted orange curve (representing $\Delta_{0,T}UAL^{B:C}$) with the point on the solid orange curve for $\Delta_{0,T}UAL^{B:\Sigma}$, which

[16] I hold $(NC – B)_t$ to their actual values for all of the counterfactuals. This would not be a strictly valid assumption for counterfactuals where the benefit structure differs from the actual, but in cases where benefit changes are minor – as in CSTRS over the period in question – this assumption should not be terribly problematic. Note also that I am assuming the difference $(NC – B)_t$ is fixed at its actual value, a slightly weaker assumption than holding NC_t and B_t fixed separately.

lies along the -\$2 billion line. The result is intuitive, but the point here is that it is *not* reflected in the conventional method.[17]

Similarly (albeit less dramatically), for the liability shortfalls, we have

$$(5^{L:C}) \quad \Delta_{0,T}UAL^{L:C} = [\Sigma_{t=1}^{T}(L_t - L^e_{t|t-1}) + \Sigma_{t=1}^{T}(r^* - r_t)A_{t-1} + \Sigma_{t=1}^{T}C_t - \Sigma_{t=1}^{T}POB_t] - [\Sigma_{t=1}^{T}(r^* - r_t)A^{L:C}_{t-1} + \Sigma_{t=1}^{T}C_t - \Sigma_{t=1}^{T}POB_t] \neq \Sigma_{t=1}^{T}(L_t - L^e_{t|t-1}) = \Delta_{0,T}UAL^{L:\Sigma},$$

because, again, the series $A^{L:C}_{t-1}$ differs from A_{t-1}, due to the endogenous amortization. Specifically, in the year following any liability shortfall, if that shortfall had not occurred the counterfactual assumes that AMT would be reduced by the interest on the UAL that would have been occasioned by that shortfall. This would reduce $A^{L:C}_{t+1}$ and, therefore, in period t+2 it would reduce any investment losses from $(r^* - r_{t+2})$. The difference, however, is nearly imperceptible in Figure 1, between the solid and dotted liability curves. Similarly, there is a slight difference between the impact of contribution shortfalls under the consistent counterfactual and the conventional allocation, as indicated by the dotted purple line in Figure 1.

I consider the dotted lines (where they differ from the solid ones) to be the correct impact of the UAL drivers, under the maintained assumption that C_t is exogenous. It is worth noting that the property of additivity is violated here. That property depended on each term in (5) equaling the difference between the actual $\Delta_{0,T}UAL$ and the counterfactual value when one of the drivers was set to zero. This no longer holds when A_t's endogeneity is recognized. The non-additivity may be empirically small under exogenous C_t (the dotted lines in Figure 1 sum to 101.3 percent of the total rise in UAL). But additivity does require inconsistent counterfactuals and will be more significantly violated when the assumption of exogenous C is replaced.

The construction of consistent counterfactuals and their comparison with the conventional method[18] has been, perhaps, tedious, given the lack of dramatic

[17] Note, however, that if $r = r^*$ the cumulative UAL impact of the POB is again unchanged beyond the initial impact. The result in $(5^{B:C})$ pertains to investment shortfalls, not the interest on the POB proceeds. *Under the assumption of exogenous C_t*, that interest is still assumed to be offset by reduced amortization. This comment applies also to the interest on liability shortfalls.

[18] For the zero-base counterfactuals, introduced in a previous note, the results are the mirror image of the counterfactuals presented in the text: the impact of investment shortfalls differs from the conventional allocation, but the other impacts are the same. Consider, for example, the liability shortfalls. Under the zero-base method (where the baseline $\Delta_{0,T}UAL = 0$), we examine the counterfactual value of (5) where the liability shortfalls, $(L_t - L^e_{t|t-1})$, hold their actual values while $0 = (r^* - r_t) = C_t = POB_t$, t = 1,…T. Denoting the result with the superscript L:z (for zero-base liability counterfactual), we have:
$\Delta_{0,T}UAL^{L:z} = \Sigma_{t=1}^{T}(L_t - L^e_{t|t-1}) = \Delta_{0,T}UAL^{L:\Sigma}$, since the other three terms of (5) are set to zero by construction. That is, the conventional allocation of the rise in UAL to liability shortfalls corresponds to the zero-base counterfactual. The endogeneity of A_t does not affect this result. That is because its value in the second term of (5) is irrelevant to both the baseline and the liability counterfactual, with $0 = (r^* - r_t)$ in both. The same result holds for POB's and the contribution shortfalls. For the investment shortfalls, however, the

differences, except for the investment shortfalls on the POB. However, it is important to get the counterfactual right, not only for its own sake, but also to determine that the inconsistent treatment of asset values is not the main problem. That is the exogeneity of C_t and the treatment of amortization. In the remainder of this paper, analyzing this problem, I adopt the consistent counterfactuals, i.e. A_t as treated as endogenous.

Exogenous AMT_t Contrasted with Exogenous Contribution Shortfall, C_t

I return to my main theme, the counterfactual treatment of amortization. As we have seen, in analyzing the conventional UAL allocation, (5), the implicit assumption is that AMT_t varies dollar for dollar with r^*UAL_{t-1} under the various counterfactuals that generate endogenous UAL's (i.e. the target of the exercise). That is, the implied counterfactuals are of the form, "What would be the impact on UAL if (say) investments had met the assumed return, *and amortization payments had correspondingly adjusted dollar-for-dollar to the interest on the resulting UAL's*?" As I have argued, this will underestimate the impact of investment shortfalls for typical amortization formulas with $\alpha_t < 1$. To get a deeper, more formal understanding of why this is so, in this section, I will consider the case of exogenous AMT_t. To put it in context of (6) and (8), this is the extreme case where the counterfactual AMT_t does not vary at all with the interest on counterfactual changes in UAL. In terms of Figure 2, we are simply rotating the line all the way around the actual values of AMT_t and $r^*UAL_{t-1,}$ from (6) to (8) to the actual AMT_t line itself, where the slope is zero and the marginal shortfall is one. I examine this extreme case in detail because it generates clear analytical results that will serve to clarify the intermediate case (8), illustrated in the next section. As mentioned above, it may also be relevant to the case where budgeters do not fund the amortization formula, but simply contribute "what they can afford."

Specifically, there are two points that will emerge from this analysis. First, under this formulation interest is reflected in the UAL impact of the drivers such as (L_t - $L^e_{t|t-1}$) and POB_t, unlike the conventional allocation. That is, the interest ensuing from the initial UAL impact accrues for additional UAL impact instead of going to amortization payments. This is a direct consequence of assuming counterfactual AMT_t does not vary at all, let alone dollar-for-dollar with interest on the UAL.

Second, there are significant interactions among the UAL-drivers for the impact on the UAL. That is, if we add up the individual counterfactual UAL impacts, the total may far

---

result differs from the conventional allocation: $\Delta_{0,T}UAL^{r:z} = \Sigma_{t=1}^{T}(r^* - r_t)A^{r:z}_{t-1} \neq \Sigma_{t=1}^{T}(r^* - r_t)A_{t-1} = \Delta_{0,T}UAL^{r:\Sigma}$, since the endogenous series $A^{r:z}_{t-1}$ (generated by (3) with $AMT^{r:z}$ instead of actual AMT) will not equal the actual series A_{t-1}. In principle, that can matter, when $r_t \neq r^*$. In practice, this may not matter much. In our adapted CSTRS example, with exogenous C_t, the UAL impact of investment shortfalls is 58.4 percent of the total rise under the zero-base counterfactuals, instead of 59.7 percent in the conventional allocation.

exceed the actual rise in UAL. Thus, if our goal is to allocate the rise in UAL among its drivers, we may face a large adding-up problem. On the other hand, if the policy interest is the impact of each driver taken one at a time, these are the estimates that we want. This issue will persist (in attenuated form) under our preferred alternative (8), as discussed in the next section, so this section explains the problem in its purest form.

I will derive the reduced form expression for $\Delta_{0,T}$UAL. Specifically, I will solve for the series A_t and UAL_t to express $\Delta_{0,T}$UAL solely in terms of the series: (L_t - $L^e_{t|t-1}$), r_t, AMT_t, NC_t, B_t, and POB_t, given the initial conditions A_0 and UAL_0. To do so, the system (3) – (4) is usefully expressed in matrix form:

$$\binom{A}{UAL}_t = \boldsymbol{M}_t \binom{A}{UAL}_{t-1} + \boldsymbol{v}_t \ ,$$

where $\boldsymbol{M}_t = \begin{bmatrix} (1+r_t) & 0 \\ (r^* - r_t) & (1+r^*) \end{bmatrix}$ and $v_t = \begin{bmatrix} AMT + (NC - B) + POB \\ (L - L^e) - AMT - POB \end{bmatrix}_t$.

This implies $\binom{A}{UAL}_T = \boldsymbol{M}_T \boldsymbol{M}_{T-1} \ldots \boldsymbol{M}_1 \binom{A}{UAL}_0 + \boldsymbol{M}_T \boldsymbol{M}_{T-1} \ldots \boldsymbol{M}_2 v_1 + \cdots + \boldsymbol{M}_T v_{T-1} + v_T$.

It can be shown, by induction, that

$$\boldsymbol{M}_T \boldsymbol{M}_{T-1} \ldots \boldsymbol{M}_1 = \begin{bmatrix} \prod_{t=1}^{T} (1+r_t) & 0 \\ \left[(1+r^*)^T - \prod_{t=1}^{T} (1+r_t) \right] & (1+r^*)^T \end{bmatrix}$$

and similarly for $\boldsymbol{M}_T \boldsymbol{M}_{T-1} \ldots \boldsymbol{M}_t$. Thus, we have

$$\binom{A}{UAL}_T = \begin{bmatrix} \prod_{t=1}^{T} (1+r_t) & 0 \\ \left[(1+r^*)^T - \prod_{t=1}^{T} (1+r_t) \right] & (1+r^*)^T \end{bmatrix} \binom{A}{UAL}_0 + \left(\sum_{t=1}^{T-1} \begin{bmatrix} \prod_{\tau=t+1}^{T} (1+r_\tau) & 0 \\ \left[(1+r^*)^{T-t} - \prod_{\tau=t+1}^{T} (1+r_\tau) \right] & (1+r^*)^{T-t} \end{bmatrix} v_t \right) + v_T.$$

Finally, from this expression we have:

$$(9)\ \Delta_{0,T}UAL = \left[(1+r^*)^T - \prod_{t=1}^{T}(1+r_t)\right]A_0 + [(1+r^*)^T - 1]UAL_0$$
$$+ \sum_{t=1}^{T-1}\left[(1+r^*)^{T-t} - \prod_{\tau=t+1}^{T}(1+r_\tau)\right][AMT_t + (NC-B)_t + POB_t]$$
$$+ \sum_{t=1}^{T}(1+r^*)^{T-t}\left[\left(L_t - L^e_{t|t-1}\right) - AMT_t - POB_t\right].$$

For the actual values of all the variables (as denoted by the absence of superscripts), the sum in (9) is equal to the sum in (5), but the counterfactuals (other than the one for contribution shortfall) will differ, because here AMT_t will be taken as exogenous instead of $C_t \equiv (r^*UAL_{t-1} - AMT_t)$.

One can readily find the counterfactual impact from (9) for liability shortfalls and POB's. For liability shortfalls, $(L_t - L^e_{t|t-1})$, simply take the difference between the actual $\Delta_{0,T}UAL$ given in (9) and the value it would take with $(L_t - L^e_{t|t-1}) = 0$. Since everything else in (9) is unchanged (exogenous or pre-determined), the result is straightforward:

$$\Delta_{0,T}UAL^{L:AMT} = \sum_{t=1}^{T}(1+r^*)^{T-t}\left(L_t - L^e_{t|t-1}\right),$$

where the superscript AMT denotes its exogeneity in the counterfactual. Here, the UAL impact of liability shortfalls includes the cumulative interest (at assumed rate r*), unlike the conventional framework where no interest is attributed to liability shortfalls, $\Delta_{0,T}UAL^{L:\Sigma} = \Sigma_{t=1}^{T}(L_t - L^e_{t|t-1})$. That is because in the conventional framework, with exogenous C_t, AMT_t responds to cover the interest on liability shortfalls. Conversely, with AMT_t exogenous, the liability shortfalls have no impact on amortization. In effect, the assumption of exogenous AMT_t reallocates the interest on the UAL impact from amortization (under exogenous C_t), to the UAL instead.

The difference can be seen in Figure 3, which depicts the impact on UAL of selected drivers,[19] under exogenous C_t, exogenous AMT_t, and (for reference in the next section) exogenous α_t. For all cases, I depict the impact under consistent counterfactuals (endogenous asset values). The UAL impact of the CSTRS benefit change in FY08 (brown curves) illustrates the difference between exogenous C_t and exogenous AMT_t. The dotted line, reproduced from Figure 1, shows the cumulative impact after FY08 was essentially unchanged[20] from the initial impact, under exogenous C_t because

[19] Not shown, to maintain legibility in the figure, are the UAL impact of deviations from liability assumptions and of the contribution shortfall. However, the latter is the same under exogenous AMT_t and exogenous C_t (as well as exogenous α_t, discussed below) since, in all cases, we are comparing the actual UAL with the UAL under the assumption that $AMT_t = r^*UAL_{t-1}$.

[20] It is exactly unchanged under the conventional method, but is slightly changed under consistent counterfactuals, where asset values are treated endogenously, as discussed above.

amortization is assumed to cover the additional interest. As the dashed curve shows, under exogenous AMT_t, the impact grows with accrued interest at the assumed rate r*.

[FIGURE 3 APPROXIMATELY HERE]

Similarly, it is straightforward to show that the reduction in the UAL due to the POB will include the investment return, at the actual rate, r_t:

$$\Delta_{0,T}UAL^{B:AMT} = -\sum_{t=1}^{T-1}\left[\prod_{\tau=t+1}^{T}(1+r_\tau)\right]POB_t - POB_T.$$

As we saw in Figure 1, under the conventional method, the initial impact persisted unchanged. Under consistent counterfactuals with exogenous C_t, the overall impact shrunk after FY08, reflecting the investment shortfalls, relative to r*, on the POB proceeds. This is shown in the dotted orange curve of Figure 3, reproduced from Figure 1. Under exogenous AMT_t, depicted in the dashed orange curve of Figure 3, we can again see the strongly negative return in FY09 shrinking the initial impact. Overall, however, the impact (i.e. reduction in UAL) grows over the period, as the actual return, r_t has been generally positive, albeit less than r*. The difference is that under the conventional frameworks' exogenous C_t, AMT_t is reduced to offset the assumed interest on the reduced UAL; with exogenous AMT_t, the POB occasions no such reduction in contributions, so the reduction in UAL is greater.

I turn now to our main focus, the impact of investment shortfalls. The counterfactual value of $\Delta_{0,T}$UAL, is found by setting $r_t = r^*$ in (9). This zeros out the first and third terms and leaves the other terms unchanged, under exogenous AMT_t. Thus, the difference between the actual $\Delta_{0,T}$UAL and its counterfactual value is the first and third terms:

$$\Delta_{0,T}UAL^{r:AMT} = \left[(1+r^*)^T - \prod_{t=1}^{T}(1+r_t)\right]A_0 + \sum_{t=1}^{T-1}\left[(1+r^*)^{T-t} - \prod_{\tau=t+1}^{T}(1+r_\tau)\right][AMT_t + (NC-B)_t + POB_t]$$

This UAL impact is simply the compounded difference between (1 + r*) and (1 + r_t) over the T-year period, as applied to the initial assets and the annual cash flows (all of which are held exogenous at the actual levels).[21]

The UAL impact under exogenous AMT_t is depicted for the CSTRS example by the dashed red line in Figure 3. It is dramatically higher than the conventional estimate,

[21] There is an additional term for the cash flows for period T. This is not affected by r_T, so it is the same under the actual and the counterfactual, and does not enter the impact of the investment shortfalls.

under exogenous C_t, depicted by the dotted red line (reproduced from Figure 1): $12.0 billion vs. $5.0 billion.

Equation (9) can be used to illuminate the difference between the UAL impact of the investment shortfall under exogenous C_t and AMT_t. The difference is in the counterfactual value of AMT_t in the last term of (9). Under exogenous C_t, the counterfactual value of AMT_t is given in (6), specifically here, $AMT^{r:C}_t = r^*UAL^{r:C}_{t-1} - C_t$. Thus, $\Delta_{0,T}UAL^{r:C}$ can be written as the difference between the actual value of (9) and the counterfactual value of (9) with $r_t = r^*$ and $AMT^{r:C}_t$.[22] Specifically, $\Delta_{0,T}UAL^{r:C}$ has the same two terms as $\Delta_{0,T}UAL^{r:AMT}$, but also has a third term:

$$(10)\ \Delta_{0,T}UAL^{r:C}$$
$$= \left[(1+r^*)^T - \prod_{t=1}^{T}(1+r_t)\right]A_0$$
$$+ \sum_{t=1}^{T-1}\left[(1+r^*)^{T-t} - \prod_{\tau=t+1}^{T}(1+r_\tau)\right][AMT_t + (NC-B)_t + POB_t]$$
$$+ \sum_{t=1}^{T}(1+r^*)^{T-t}[AMT_t^{r:C} - AMT_t]$$

The third term is negative, since the counterfactual values of AMT_t are reduced under (6) (exogenous C_t) when $r_t = r^*$ reduces the UAL's. Thus, $\Delta_{0,T}UAL^{r:AMT}$ (exogenous AMT_t) exceeds $\Delta_{0,T}UAL^{r:C}$ (exogenous C_t) by $\Sigma_{t=1}^{T} (1+r^*)^{T-t}(AMT_t - AMT^{r:C}_t) = \Sigma_{t=1}^{T} (1+r^*)^{T-t}\, r^*(UAL_{t-1} - UAL^{r:C}_{t-1})$. This represents the cumulative value (with interest, as of period T) of the amortization payments that are assumed, under exogenous C_t, to cover the additional interest on UAL from the failure of investment to meet the assumed returns, r*. Under exogenous AMT_t these additional payments would not have been made, so the full impact of the investment shortfalls would fall on UAL, instead of being offset, as under the conventional method.

As Figure 3 suggests, adding up the UAL impacts of each driver under exogenous AMT_t well exceeds the conventional estimates and, hence, the actual rise in UAL – about double in the case of CSTRS. Why is there this big adding-up discrepancy? The answer is (by definition) that there are large interaction effects. Concretely, the important interactions here are between the impact of the investment shortfalls and a regime of underfunding. If investment shortfalls raise UAL, but the funding formula does not adjust amortization for the ensuing additional interest (or does not fully adjust, as in (8)), then there is a further rise in UAL from the subsequent underfunding. Thus, if we add up the UAL impacts one at a time (especially investment shortfalls and contribution shortfalls) we will find a total that is much larger than taking them all together.

[22] The result is equivalent to that given in ($5^{r:C}$) $[\Sigma_{t=1}^{T}(r^* - r_t)A_{t-1}]$ but the derivation from (9) is more readily compared with $\Delta_{0,T}UAL^{r:AMT}$.

So what are we to make of the adding-up problem? In my view, the answer is "not too much." The end-goal of a policy-relevant exercise is to evaluate useful counterfactuals, not necessarily to parcel out the total rise in UAL in a fashion that adds up. Indeed, if we get a solution that adds up for uninteresting counterfactuals – especially ones that are misunderstood -- we may in fact be misinforming the user. If the user's interest is the impact of (say) a high assumed rate of return, we should use the right counterfactual to answer that simple question. If the user's interest is the impact of multiple drivers taken together, then one should simply use that joint counterfactual, which will properly handle the interactions.

To repeat, none of this discussion is meant to argue in favor of treating AMT_t as exogenous in our counterfactual analyses, except in cases where funders simply contribute "what they can afford," instead of funding the formula. However, the same issues arise when AMT_t is governed by the formula, albeit in attenuated form. I have examined the exogenous AMT_t case in such detail because it illuminates the issues most clearly.

Exogenous Amortization Factor, α_t

Instead of assuming that C_t or AMT_t is exogenous, it is more natural to assume that the actual amortization formula in use is exogenous.[23] As discussed above, amortization formulas typically result in payments that are proportional to the interest on the UAL, with a factor α_t, as given in (8). If so, then taking the formula as exogenous is equivalent to taking α_t as exogenous. Thus, if one is willing to assume that the formula is of the proportional form, one need not delve deeply into the formula's intricacies, but simply take the actual $\alpha_t \equiv AMT_t/r^*UAL_{t-1}$ as our exogenous series.

The "what if" exercises for investment and liability shortfalls and POB take α_t as exogenous at the actual level and model the counterfactual AMT_t using (8). In this way, the estimated impact of the UAL-driver under consideration factors in a proportional AMT response to changes in the UAL interest, rather than a dollar-for-dollar response or zero response.[24] As shown in Figure 2, if $\alpha_t < 1$ (underfunding), the counterfactual AMT formula lies between the other two counterfactuals examined, exogenous C_t and AMT_t. Thus, it is not surprising that the results of this method lie between the two other methods. Specifically, the solid lines in Figure 3 take as exogenous the series of α_t's

[23] That said, it is not unheard of for plans to change their amortization formula in response to adverse developments, such as investment shortfalls.

[24] The "what if" exercise for contribution shortfalls compares the actual UAL under the actual series α_t with the no-shortfall series, $\alpha^* = 1$ (analogous to comparing the series r_t and r^* for investment shortfalls). As stated earlier, it is no different than for exogenous AMT_t or C_t.

derived from our adapted version of CSTRS. Taken at face value (though see the caveats in the note below), α_t has averaged about 0.6 over the period FY01-FY14.[25]

These results – the solid lines in Figure 3 – represent my preferred method for estimating the UAL impact of each driver, taken one at a time. Most notably, the impact of investment shortfalls (with the superscript notation r:α, to denote exogenous α_t) is $7.3 billion instead of $5.0 billion under the conventional method. Again, this result is due to replacing the implicit counterfactual assumption that AMT_t varies dollar-for-dollar with the interest on the resulting impact on UAL. The more realistic assumption, that AMT_t varies at the actual historical proportion α_t, provides a much more credible and useful estimate for policy-makers of the UAL impact of investment shortfalls.

Adding up the impacts of all UAL-drivers exceeds the actual rise in UAL, due to the interactions discussed earlier. As previously argued, the adding-up goal is not of direct policy relevance. If the goal is to assess the impact of all drivers taken together, then that can be directly modeled: the result will reflect the interactions and will, tautologically, equal the actual rise in UAL. The conventional methodology, in effect assumes away the interactions by virtue of an untenable implicit assumption regarding amortization.[26] The individual impacts add up, but they are not of policy relevance.

The Impact of Shortfalls on Cumulative Contributions and UAL

Another impact of interest from the UAL-drivers is their impact on cumulative contributions. That is, we are not only interested in the impact of these drivers on payments-yet-to-be made – the UAL – but also the impact on payments that have been made – the amortization. This is an important impact and one that differs under exogenous C_t, AMT_t, and α_t. I will compare here the impact on cumulative amortization and UAL of investment shortfalls and other drivers under the different assumptions.

---

[25] My adaptation differs from the actual CSTRS calculations in manners that can be substantive. The differences include the 2-3 year lag between amortization and UAL (due to CSTRS' biennial valuation schedule) and smoothed vs. market valuation. Moreover, the reason $\alpha_t < 1$ for CSTRS is different from the discussion of (7)-(8). Unlike most plans, CSTRS divides up its UAL into separate pieces and amortizes them separately. For the main piece, the remaining amortization period (closed) is now short enough to make $\alpha_t > 1$, but there is also a significant piece of ***negative*** amortization, which is amortized over a shorter period, and that renders the overall ratio of amortization to interest less than one. Thus, the appropriation here of the overall α_t's "as if" the formula was proportional is solely for illustrative purposes. Note also that when the period of negative amortization ends (in FY22), net amortization payments will jump and, barring any change in the formula, the overall α_t will exceed 1. The same feature characterized the Public School Employees' Retirement System of Pennsylvania, and when the period of negative amortization ended, there was a huge (and predicted) spike in amortization payments that has caused great difficulty (see Costrell and Maloney, 2013, especially Figure B-2).

[26] It would be equally misleading to "fix" the α-based impacts to force them to add up. The usual such "fixes" are: (i) scaling down all figures; and (ii) adding drivers one at a time, in an arbitrary order, instead of examining their impacts singly. These are solutions to a problem of appearances, more than policy.

The result is that the total impact is virtually invariant, but the split between paid and unpaid impact varies by assumption.

As we have seen, the counterfactuals embed within them the assumed behavior of AMT_t, depending on what is assumed to be exogenous: C_t, α_t, or AMT_t itself, as illustrated in Figure 2. Thus, for the counterfactual on investment shortfalls (i.e. where r = r* and all other drivers are at their actual values), we have, under the conventional method (exogenous C_t) from (6):

$$(6^{r:C}) \qquad AMT^{r:C}_t = r^*UAL^{r:C}_{t-1} - C_t.$$

This generates a series of impacts of the investment shortfalls on amortization payments, the difference between the actual and the counterfactual, $AMT_t - AMT^{r:C}_t = r^*(UAL_{t-1} - UAL^{r:C}_{t-1})$. To put these amortization impacts on a common footing with the UAL impacts, period by period, one calculates their cumulative value, with the assumed interest r*. That is, one calculates the asset value of the series of amortization impacts, with the following result by year T:

$$\Sigma_{t=1}^{T}(1+r^*)^{T-t}(AMT_t - AMT^{r:C}_t) = \Sigma_{t=1}^{T} (1+r^*)^{T-t}r^*(UAL_{t-1} - UAL^{r:C}_{t-1}).$$

This term appeared (in its negative value) in (10) above, and represented the difference between the UAL impact of investment shortfalls under exogenous C_t and exogenous AMT_t. It is represented in Figure 4 as the $7.0 billion blue bar in the first column. The total impact of investment shortfalls – UAL plus cumulative amortization – is $12.0 billion, the same under exogenous C_t and exogenous AMT_t (the third column in Figure 4). Since (by definition), there is no amortization impact under exogenous AMT_t, the $7.0 billion blue bar in the first column is the difference between the UAL impacts under exogenous C_t and AMT_t, as stated above.

[FIGURE 4 APPROXIMATELY HERE]

For my preferred model, exogenous α_t, the $12.0 billion total impact is split between $7.3 billion impact on UAL (previously depicted in Figure 3) and the $4.7 billion impact on cumulative amortization, represented by the blue bar in the 2nd column of Figure 4. In short, the conventional method -- with exogenous C_t – implicitly attributes a substantial portion ($2.3 billion in this example) of the total impact of investment shortfalls to amortization instead of UAL. That is, the conventional method implicitly assumes that much more of the impact has already been paid for than the amortization formula would imply, so the UAL impact is that much smaller. More precisely, the counterfactual scenario is that without the investment shortfalls, the amortization payments would have been $7.0 billion lower, instead of the $4.7 billion implied by the amortization formula, so the UAL would have been only $5.0 billion lower, instead of $7.3 billion lower.

Similar observations apply to the impacts on the UAL and cumulative amortization of benefit hikes and liability shortfalls. Thus, the 2nd and 3rd triplets in Figure 4 are similar to the 1st triplet, but on reduced scale. The same pattern holds for POB's, but of the reverse sign, as shown in the 5th triplet.

For contribution shortfalls – depicted in the 4th triplet -- the counterfactual amortization assumptions do not matter. They pertain only to the other counterfactual impacts. The UAL impact of the contribution shortfall is offset by the reduction in amortization – by definition – so the total impact is essentially zero.

One might be tempted to dismiss the impact of investment shortfalls and other UAL-drivers on cumulative amortization as a matter of little significance since, after all, this has already been paid. But this impact is not totally benign. It means that some of the costs have been shifted from the cohort that incurred them to a later cohort, creating some generational inequity in doing so. In other words, the split of the impact between paid (cumulative amortization) and yet-to-be-paid (UAL) is really a split between generational inequity already imposed and generational inequity to come.

To summarize, the conventional method understates the impact of investment shortfalls in two ways. Using the CSTRS example, the impact on UAL is understated at $5.0 billion instead of $7.3 billion. In addition, the conventional method ignores the $4.7 billion impact on cumulative amortization, and the attendant inequity that has already been visited upon recent cohorts by the shortfall of investment returns from the assumed return.

Accounting for Pension Obligation Bonds

To complete the picture for CSTRS and other plans that have relied on POB's, one should include the POB debt itself. Although POB's are typically issued by the state, rather than the pension plan, for the taxpayer both the UAL and POB debt are liabilities. Indeed, the full liability is understood in pitching POB's, since they are sold as an arbitrage play, borrowing at one rate and investing in the plan, with assumed return r^*.

In the case of CSTRS, the State Treasurer estimated the interest at 5.88 percent over the 25-year life of the bonds, compared to 8.5 percent assumed return on the proceeds, invested in CSTRS.[27] Since the bonds were issued in FY08[28], I estimate that the return on CSTRS was 6.37 percent through FY14, well below the assumed 8.5 percent, but greater than the 5.88 percent calculated average over 25 years. However, the first two years of interest payments were also borrowed, adding $266 million to the bond issue, for a total of $2.277 billion (including issuing costs of $11 million). In addition, the debt service payments are highly back-loaded, so the outstanding POB debt continues to rise until FY22 – see Figure 5.[29]

[FIGURE 5 APPROXIMATELY HERE]

It is fairly straightforward to add the outstanding POB debt to the UAL and to calculate the counterfactual if the POB's had not been issued. The result is depicted in the orange POB line of Figure 6. Since the initial POB issue was $2.277 billion and only $2.0 billion went to pay down the UAL, the FY08 figure is positive $0.277 billion instead of negative $2.0 billion. The FY09 figure rises further, due to the large investment losses that year. Since then the figure has drifted down, due to generally good investment returns, but remains in positive territory as of FY14: the outstanding POB debt of $2.333 billion outweighs my estimate of the UAL impact, -$2.088 billion.[30]

To summarize, comparing Figures 6 and 1, the main differences from the conventional method are: (1) the inclusion of the POB debt; and (2) the substantially higher estimate of the UAL impact of investment shortfalls, due to the more credible counterfactual of amortization.

[FIGURE 6 APPROXIMATELY HERE]

Conclusion

In this paper I have offered a methodological critique and replacement with substantive implications for assessing the impact of investment shortfalls and other UAL-drivers. The conventional method of calculating the UAL impact over time – simply summing the annual actuarial gain/loss figures -- implicitly assumes that in the counterfactual

---

[27] State Treasurer Denise L. Nappier, as quoted by Alicia Munnell, et al., "Pension Obligation Bonds: Financial Crisis Exposes Risks," Center for Retirement Research, January 2010, p. 3. The actual bond issue was quite complex – 15 current interest bonds (i.e. no principal paid until maturity) and 6 capital appreciation bonds (i.e. no principal or interest paid until maturity) of different maturities and rates.

[28] Authorized by P.A. 07-186, codified in Sec. 10-183qq of Chapter 167a, "Teachers' Retirement System."

[29] The yearly outstanding POB balance is reported in the Annual Information Statement, State of Connecticut, February 28, 2014, revised as of December 8, 2014, p. II-9. The annual sum of debt service payments from the 21 bonds is not disclosed, but can be computed from the information given in the initial POB disclosure, April 16, 2008.

[30] Adding in the cumulative value of amortization payments saved and POB debt service paid makes the result for the total impact of the POB a bit more positive (i.e. more adverse).

exercise, amortization would adjust dollar-for-dollar with the interest on the UAL, unlike the proportional adjustment of typical amortization formulas. As a result, this method implicitly attributes much of the investment shortfall's impact to the cumulative value of amortization instead of the UAL. This can lead to a substantial underestimate of the UAL impact, compared to that implied by actual amortization formulas. In cases where the amortization formula is not funded, and contributions are simply what the budgeters "can afford," the underestimate is likely more pronounced, since the amount budgeters "can afford" is even less responsive to the UAL.

References

Biggs, Andrew G., 2015. "The state of public pension funding: Are government employee plans back on track?" American Enterprise Institute, Economic Perspectives, September 2015.

Connecticut State Teachers' Retirement System, various years, "Actuarial Valuation."

Costrell, Robert M. and Larry Maloney, 2013. "The Big Squeeze: Retirement Costs and School District Budgets. Paying the Pension Price in Philadelphia," Thomas B. Fordham Institute, June 2013.

Munnell, Alicia H., Jean-Pierre Aubry, and Mark Cafarelli, 2015. "How Did State/Local Plans Become Underfunded?" Center for Retirement Research at Boston College, Number 42, January 2015.

Munnell, Alicia H., Thad Calabrese, Ashby Monk, and Jean-Pierre Aubry, 2010. "Pension Obligation Bonds: Financial Crisis Exposes Risks," Center for Retirement Research at Boston College, Number 9, January 2010.

State of Connecticut, 2008, "New Issue. $2,276,578,270.55 State of Connecticut, Taxable General Obligation Bonds," April 16, 2008.

State of Connecticut, 2014, "Annual Information Statement," February 28, 2014, revised December 8, 2014.

Figure 1: ΔUAL, Conventional Method, C_t Exogenous. CSTRS, FY00-FY14

adapted from CSTRS: market asset values, without midyear return on cash flows and accruals

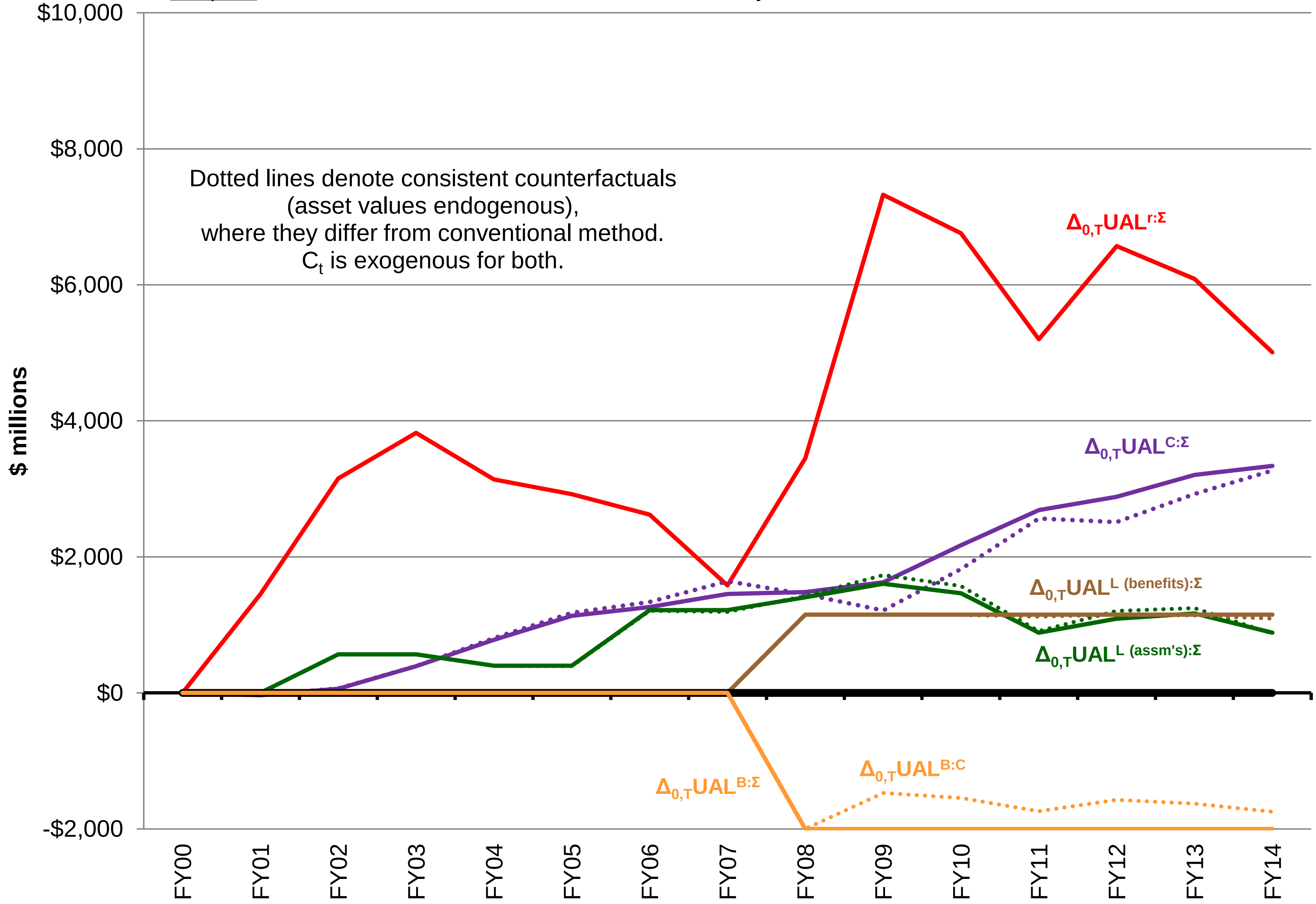


Figure 2: Counterfactual AMT_t as Function of Counterfactual r^*UAL_{t-1}
under Exogenous C_t, Exogenous α_t, and Exogenous AMT_t

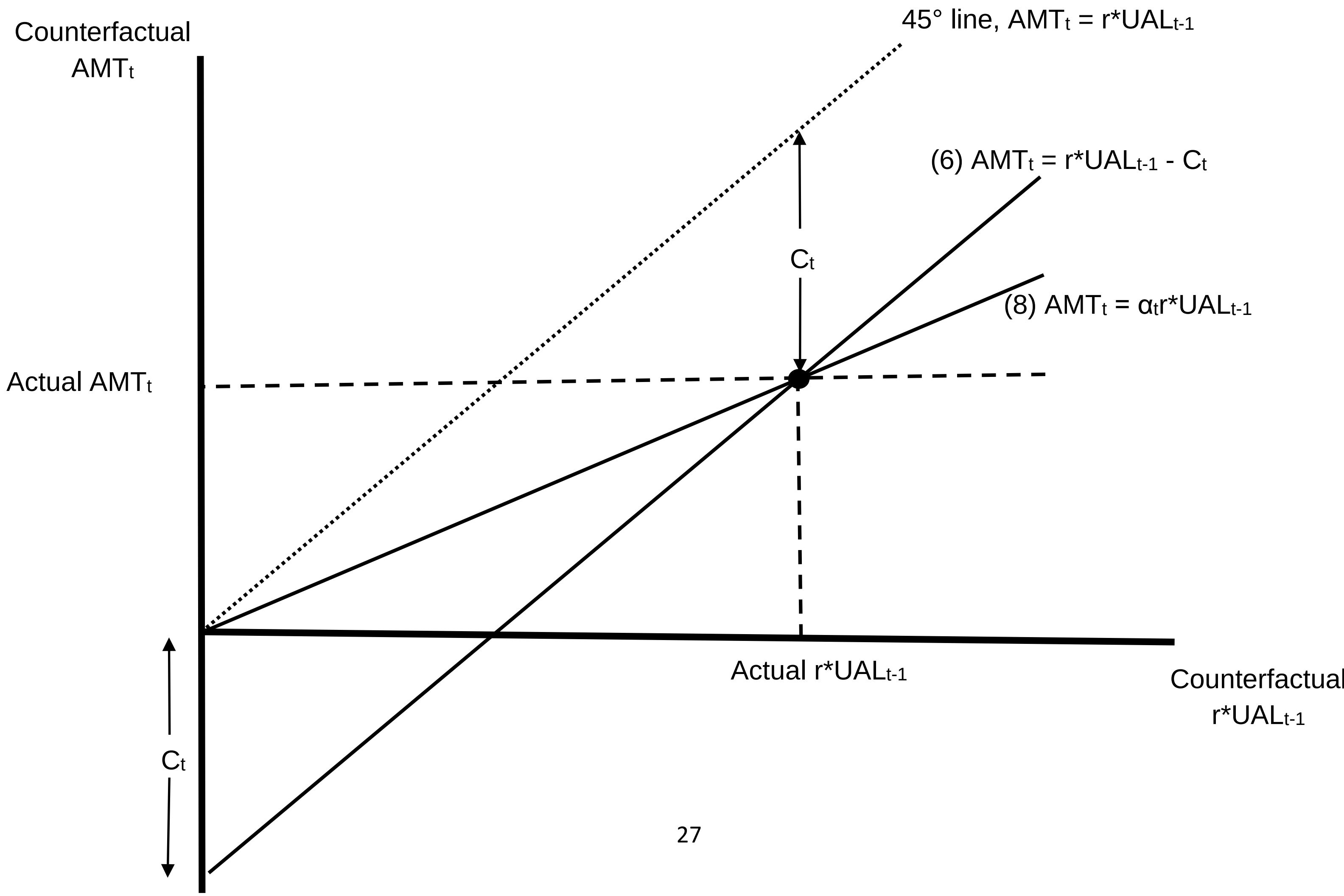


Figure 3: ΔUAL, AMT_t vs. C_t vs. $α_t$ Exogenous. CSTRS, FY00-FY14

adapted from CSTRS. Selected Drivers. Consistent Counterfactuals.

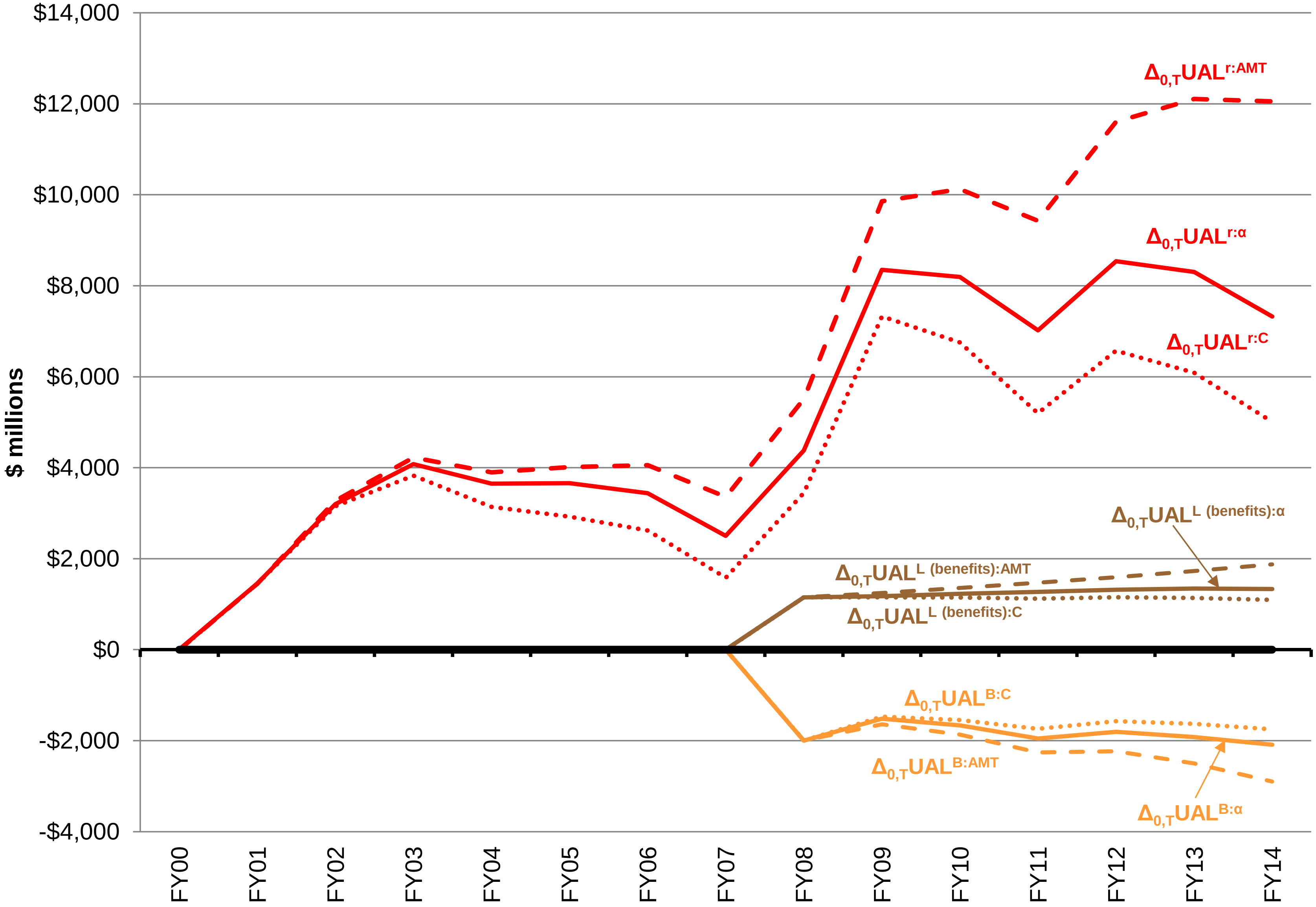
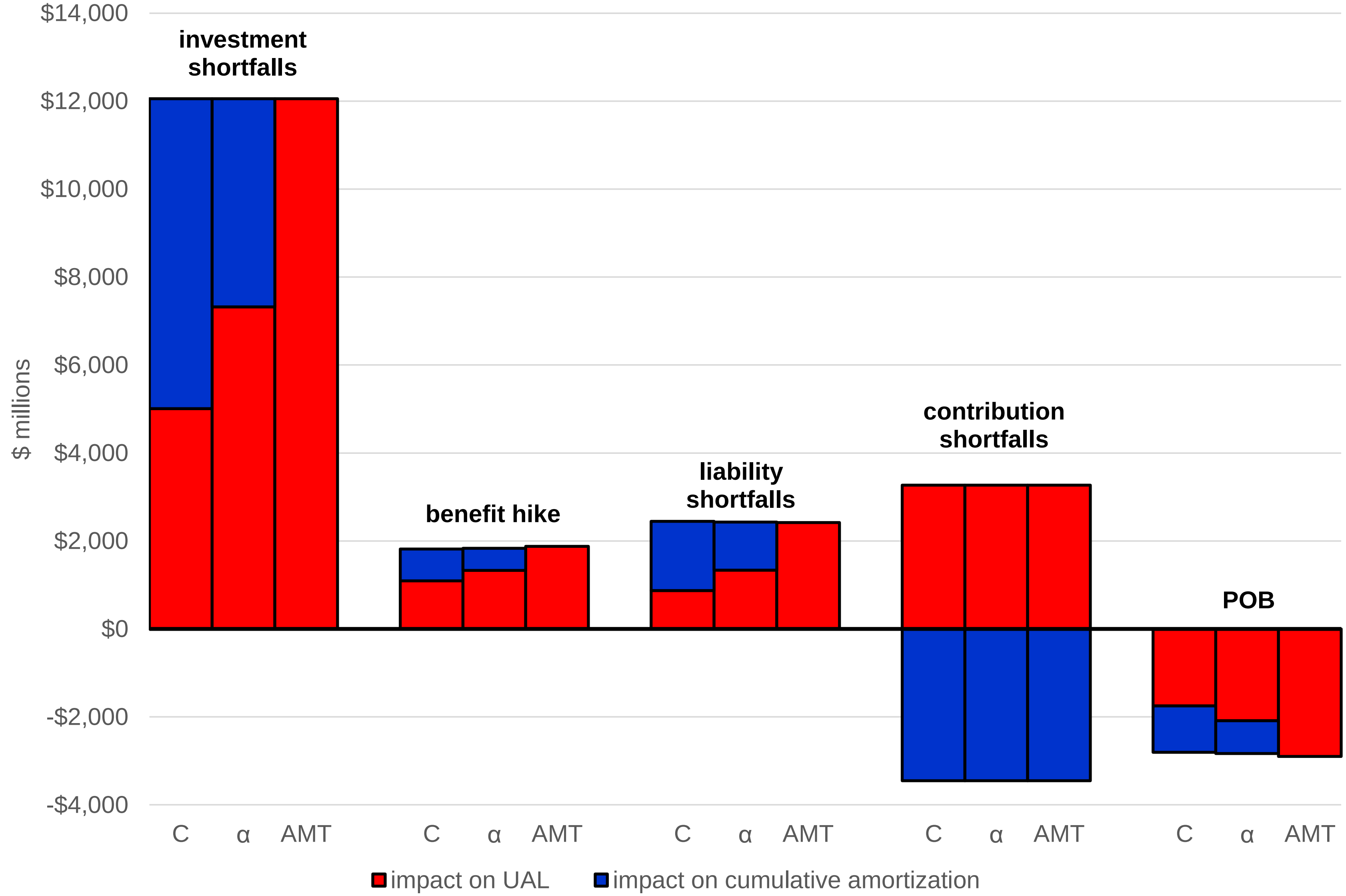
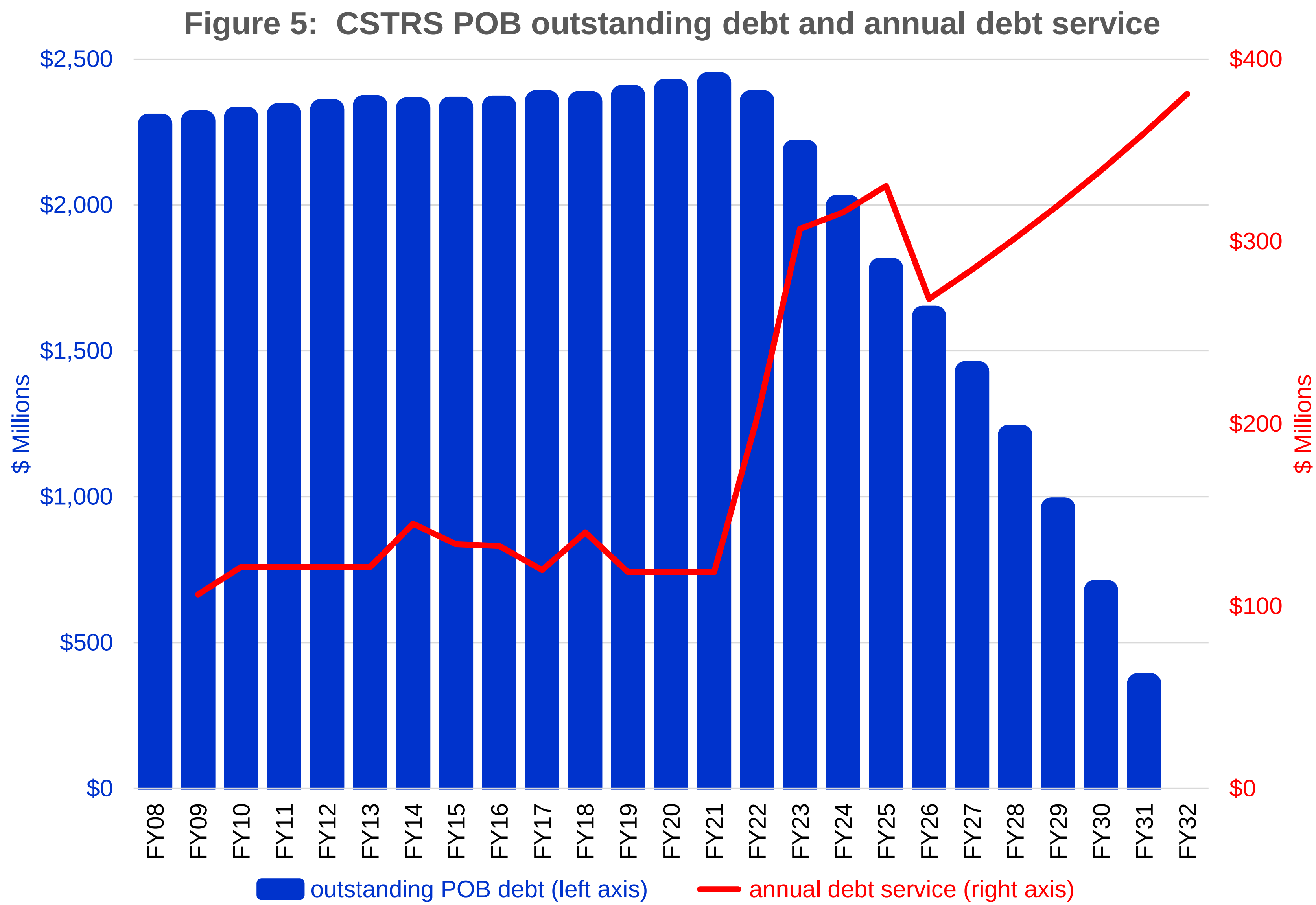


Figure 4: Impact on UAL and Cumulative Amortization. C_t vs. α_t vs. AMT_t exogenous.
adapted from CSTRS, FY00-FY14.





Source: Annual Information Statement, State of Connecticut, February 28, 2014; author's calculations

Figure 6: ΔUAL + POB debt, α_t Exogenous, CSTRS, FY00-FY14

adapted from CSTRS, supplemented by State of Connecticut Annual Information Statement.