Journal of Pension Economics and Finance (2023), 1–29
doi:10.1017/S1474747223000173

ARTICLE

Toward an economic reformulation of public pension funding

Robert M. Costrell and Josh B. McGee

Department of Education Reform, University of Arkansas, Fayetteville, AR 72701, USA
Corresponding author: Robert M. Costrell; Email: costrell@uark.edu

(Received 25 January 2023; revised 10 July 2023; accepted 30 August 2023)

Abstract

We propose an economic reformulation of contribution policy integrating: (1) formalization of sustainability as the steady-state contribution rate, incorporating both the expected return on risky assets and a low-risk discount rate for liabilities; (2) derivation of contribution adjustment policies required for convergence toward the target funded ratio and contribution rate; and (3) a stylized optimization framework for simultaneous determination of the target portfolio return and funded ratio. This analysis provides new theoretical insights into the basis for pre-funding vs. pay-as-you-go, resting on the convexity of the long-run risk–return relationship, and also potentially practical guidelines for contribution policy.

Keywords: public pension finance

JEL Code: H75

1. Introduction and summary

Most state and local employees are enrolled in final-average-salary defined benefit pension plans. These plans rely on taxpayer and member contributions and investment returns to pay for benefits. Under existing actuarial funding policy, the goal is to fully fund benefits over the course of workers' careers. To achieve this goal, plans must make a host of predictions about the future to set adequate contribution rates. If reality falls short of their predictions, the government sponsor must make up the difference with additional contributions. Unfortunately, existing actuarial funding policy has resulted in contributions that have fallen short of the amount needed to cover promised benefit payments resulting in deteriorating funded ratios, steeply rising government contributions, and reduced benefits for new public workers.

Given the scale of public pension promises, pension sustainability not only has big implications for millions of public workers' retirement security but also for government budgets and future generations of workers and taxpayers. Yet, the concept of sustainability has not been clearly defined, nor has the related risk–return tradeoff been well integrated into funding policy. We contend that the current actuarial formulation is part of the problem, and has, as a result, arguably failed to deliver either sustainability or intergenerational equity.

In this paper, we outline an economic reformulation of contribution policy that begins by: (i) conceptualizing sustainability as the steady-state contribution rate, conditional on the target funded ratio; followed by (ii) analysis of contribution adjustment policies required for convergence, in expected value; and concluding with (iii) a stylized optimization framework for simultaneous determination of the target portfolio return and funded ratio. This analysis provides new theoretical insights into the basis for pre-funding, and also potentially practical guidelines for contribution policy.

Specifically, on the theoretical side, we find that the traditional basis for pre-funding vs. pay-as-you-go, which rests on the spread between the risk-free return and the growth rate of payroll, may be greatly augmented by the net benefit of the risk premium chosen by optimizing policymakers. We show that the relative magnitude of these two bases for pre-funding depends on the policymakers' tolerance for risk imposed on future generations and the convexity of the long-run risk–return relationship.

Put differently, the actuarial basis for pre-funding rests on the difference between the pay-go rate and the normal cost rate, which, in turn, rests on the spread between the payroll growth and discount rate. If normal cost is properly discounted, at a low-risk rate, this spread is narrow. However, if the cost of future risk is considered low, relative to the target return, the optimal risk–return play outweighs the true actuarial basis for pre-funding.

On the practical side, our reformulation proposes: (i) replacing the actuarial funding target with a target for the true funded ratio, where liabilities are properly discounted and (ii) replacing actuarial contribution policies of normal cost plus amortization with a two-gap adjustment process. That process adjusts contributions based on the gap with the steady-state contribution rate and the gap with the target funded ratio. We show how the two adjustment parameters can be varied to manage the tradeoff between contribution risk and asset level risk. We also provide an illustrative example (CalSTRS) of how such a policy can be tailored, outside of steady state, to moving targets. We believe the approach sketched out here promises both to provide a more sustainable basis for funding policy and deeper insights into the basis for pre-funding in the first place.

2. The problem with actuarial funding policy and our alternative

At the heart of the issue with actuarial funding is a puzzle it seemingly cannot solve. It is generally agreed among economists that pension liabilities should be discounted at a low-risk rate corresponding to the guaranteed nature of the benefits promised, at least for reporting purposes. But it remains an open question how proper discounting of liabilities should inform funding policy. Actuarial funding policy sets contributions equal to the normal cost plus amortization of any pension debt. However, normal costs must logically be discounted by the same rate as liabilities since they are mathematically linked. Consistently discounting normal cost by the low-risk rate would dramatically raise contributions, compared to standard actuarial practice of discounting by the expected (or assumed) return on risky assets. Unless the pension plan decides to invest only in low-yield, risk-free assets – in which case contributions would indeed have to be dramatically elevated – there seems to be no way to fit the square peg of proper liability discounting into the round hole of actuarial funding policy.[1]

The core problem is that actuarial funding formulas are deterministic[2] and ill-suited to conveying the benefits and costs of investment in risky assets. This requires both the expected return and the risk-free return, as the risk premium is the spread between the two. Our economic reformulation starts (in Section 4) with the steady-state analysis of the fundamental equations of motion for assets and liabilities, where liabilities are properly discounted at a low-risk rate while asset growth carries a risky expected return. For any given target funded ratio, we derive a target contribution rate embedding the expected return on risky assets and the low-risk discount rate for liabilities.

Specifically, the expected steady-state contribution rate is a blend of the pay-go rate and the properly discounted normal cost rate, weighted by the target funded ratio, and partially defrayed by the risk premium between the expected return and the risk-free rate. Among other insights, this relationship shows that the expected steady-state contribution rate can fall well below the true normal cost rate, due to the premium on risky investments. Since the risk premium mirrors the cost of risk, this formulation helps encapsulate the tradeoff between the cost of risk and the benefit of lower expected contributions.

We then formally derive (in Section 5) a family of transition policies for convergence toward the expected steady state. These policies adjust contributions based on the two gaps with the asset and

[1]In principle, GASB now requires a blended discount rate, in the event of projected asset exhaustion, but the admissible projection methods provide so much latitude that a blended rate is almost never invoked.

[2]GASB requires some sensitivity analysis for liabilities, and a few plans do so for contribution rates too.

contribution targets. We illustrate how the adjustment parameters of our proposed policy might be calibrated to manage the tradeoff between long-run contribution rate risk and the speed of adjustment toward the two targets. Simulation of such a policy illustrates the rising risk over time and, importantly, shows the convexity of the long-run risk–return relationship.

Finally (in Section 6), we embed our steady-state results in a stylized optimizing framework for the simultaneous determination of the target funded ratio and asset allocation, based on the tradeoff between risk and return. Our optimizing conditions, together with our steady-state results, generate a new perspective on the basis for pre-funding. Specifically, we find that convexity of the long-run risk–return relationship can lead policymakers to place far greater weight on the net benefits of risky investment (as they evaluate them) than on the narrow difference between the pay-go rate and the properly discounted normal cost rate.

In sum, our analysis offers both a new understanding of the basis for pre-funding and suggests a possible reformulation of funding policy toward that end. Such a policy replaces the standard actuarial formula of normal cost (wrongly discounted) plus an amortization rate, with, instead, a more general steady-state contribution rate, plus two adjustment factors calibrated to manage the tradeoff between contribution rate risk and timely convergence. Our concluding section offers caveats and suggestions for future refinements.

3. Literature review: where does our paper fit?

We see three overlapping strands in the literature on optimal pension funding policy (asset accumulation and asset allocation): (1) corporate pensions; (2) public pensions aimed at full funding; and (3) public pensions where the steady-state funded ratio is an open question addressed by the models. This paper falls in the latter category, but the broader context is useful.

The seminal papers on corporate pension funding policy (Sharpe, 1976; Treynor, 1977) start from an irrelevance proposition. Analogous to Modigliani–Miller, funding policy does not matter under frictionless, fully informed, unregulated complete markets. Specifically, the size and risk of the pension fund can be thought of as creating a put option owned by the firm with a default-contingent claim against the employees. If recognized by the employees (or their bargaining agent), any variation in the risk of default, reflected in the value of the put, would be offset by a wage differential compensating for the pension risk. From that starting point, the literature draws out how funding policy does matter in the 'real' world, based on carefully specified deviations from the ideal, starting with the introduction of ERISA and the implicit pricing – or mispricing – of pension insurance.[3]

Public pension funding policy differs in key respects from that of corporate pensions. Default is generally not an option. The objective function for policymakers is not shareholder value, but rather (ideally) the interests of present and future taxpayers. But here, too, one finds an irrelevance theorem if the Modigliani–Miller conditions hold, along with Ricardian equivalence regarding taxation, so interest focuses on departures from these conditions. Thus, D'Arcy *et al.* (1999) examine the optimal funded ratio over time to minimize the cost of distortionary taxes in a multi-period deterministic model, based on the relationships among the initial funded ratio and the growth rates of pension benefits and the tax base.

This line of thought also extends to the question of asset allocation. While a traditional finance approach would match risk-free income streams to the stream of promised payments, modern asset liability management theory notes that liabilities bear risk (e.g., due to wage fluctuations), which can be hedged to some extent by equity holdings. Lucas and Zeldes (2009) examine the impact of distortionary taxes on the optimal equity holdings, chosen to smooth tax rates over time in a stochastic two-period model, due to the positive correlation between equity returns and liability growth.

Pennacchi and Rastad (2011) consider the polar case where the portfolio can be chosen with stochastic properties that exactly offset those of liabilities. They make the strong argument that such complete 'immunization' would maximize utility of a representative risk-averse taxpayer, given that

[3]See Love *et al.* (2011), for a more recent treatment of the impact of regulation on optimal corporate pension risk.

taxpayers lack the information to achieve those ends through their own portfolio choices. However, Pennacchi and Rastad (2011) also consider the agency problems that may lead to excessive risk when policymakers and fund managers maximize their own utility (which may downplay liability risk) instead of taxpayers' utility (where liability risk ultimately affects the risk of future taxation).

Conversely, factors against risky portfolios include a different agency problem, where upside outcomes may not fully benefit taxpayers but may instead lead to enhanced employee benefits. Van Binsbergen and Brandt (2016) analyze the asset allocation problem through a dynamic programming model, where the objective function represents the preferences (including risk aversion) of an investment manager (arguably analogous to that of a public policymaker). Their goal is to examine the impact of financial reporting rules, pre-emptive constraints to control risk and ex-post penalty payments for underfunding.

All of the public pension policy models above are tethered by the goal of full funding at some future date.[4] However, there is another strand of work, based on overlapping generations (OLG) to perpetuity that poses the question of whether liabilities should be fully funded, even in steady state. This literature, of course, begins with Samuelson's (1958) seminal analysis, where the optimality of pre-funding vs. pay-go turns on whether the rate of return exceeds the growth rate, a condition that is generally assumed to hold. In addition, the traditional interpretation of intergenerational equity holds that each generation of taxpayers should pay for its own full cost of public services, including pre-funding benefits (Munnell *et al*., 2011).

Further contributions in this third strand examine the pros and cons of pre-funding based on additional departures from the Ricardian and Modigliani–Miller conditions. Bohn (2011), for example, constructs an OLG model featuring intermediation costs faced by individual borrowers that exceed those of public entities, so that it is efficient for public pension plans to take on debt on behalf of the taxpayers. As a result, the optimal funding level is less than 100 percent and may well be zero (pay-go). Conversely, various aspects of political economy (transparency, agency problems, distorted political time horizons) may argue for pre-funding, to one extent or another (Brown *et al*., 2011, section 3).[5]

More recently, Lenney *et al*. (2019*a*, 2019*b*, 2021) challenge the intergenerational equity case for pre-funding, observing that rising contributions to pay down pension debt have burdened current generations with the cost of benefits for prior generations, instead of spreading these costs over the indefinite future. Their recommended funding policy is, instead, to simply maintain current pension debt ratios, a policy that effectively takes the properly discounted current funded ratio as given and sets the steady-state contribution rate based on that status quo.[6] Under their preferred scenarios of modest expected returns, aggregate contribution rates would only need a moderate hike to stabilize debt ratios.

Lucas (2021) and Rauh (2021) point out the difficulties that arise from using a risky rate of return in a deterministic model, even as liabilities are discounted at a risk-free rate. As a result, Rauh (2021) argues the debt-stabilizing contribution rate may well rise much higher than Lenney, *et al*. (2021) suggest. He also points out that their paper assumes away the cost of insolvency in their stochastic simulations by unrealistically assuming plans can borrow through periods of negative assets. As he points out, rating agencies factor in the risk of insolvency, and, hence, the risk that contributions would jump to the pay-go rate. Lucas (2021) also places great emphasis on the risk of insolvency, arguing that the goal of (expected) debt-stabilization is a less suitable definition of sustainability than insolvency-minimization.

Where does the present paper's analysis fit in with these preceding literatures? As stated above, this paper (building on Costrell (2018*a*) and Costrell and McGee (2020)), falls in the category of perpetually overlapping generations, where the steady-state funded ratio is an open question, unlike closed interval models necessarily culminating in full funding. The big question we address, as summarized above, is how the target funding decision depends on the risk and return profile that policymakers choose in their asset allocation decision. That said, some of the specific features of our analysis

[4] In the Pennacchi and Rastad (2011) model, taxpayers pay off any shortfall (or recoup any surplus) at time T.

[5] See also the related formal political economy analysis in Glaeser and Ponzetto (2014).

[6] This is a special case of our model, as shown in Section 4 below.

have some surface similarities to other papers summarized above, so it may be helpful to spell out how they differ.

In the first step of our analysis, we derive the steady-state contribution rate, contingent on the target funded ratio and asset allocation. This goal of finding a constant contribution rate is similar to those papers that base their results on equalizing distortionary tax rates over time, *en route* to full funding. Our rationale for analyzing steady states is based instead on the simple notion of sustainability.

The second step of our analysis finds an adjustment process and parameters thereof for the contribution rate that leads to timely convergence. This surprisingly non-trivial exercise – even in the deterministic case – has some surface similarity to the analysis of penalty payments for underfunding but is very different in context and motivation. It is based on securing convergence rather than shaping incentives.

The culminating step of our analysis introduces an optimization framework, similar to the literature cited above, but with a number of key differences in structure and purpose. The purpose is to revisit the theoretical basis for pre-funding, with a focus on the net benefits from public risk-bearing, as evaluated by the policymakers. Specifically, we posit a very simple objective function – a reduced form, so to speak – which may be thought of as capturing all the deviations from Ricardian equivalence and Modigliani–Miller conditions that are more carefully analyzed in the literature above. From this simple setup and estimates of the convexity in the long-run risk–return relationship, we can compare the traditional basis for pre-funding with that derived from the perceived net benefit of risky investments, for any chosen target return.

4. Steady-state analysis

Sustainability – the idea of something being sustained – raises the question of what that something (or more than one something) is for pension policy. It seems natural to identify the contribution rate as the key variable that one would want to stabilize and to do so at a level that would stave off risk of insolvency, rating agency downgrades for the taxing authority, crowd-out of other necessary public services, or some other form of fiscal distress. Judging by the fact that contribution rates have been generally rising since around 2000, it does not seem that current actuarial funding models have succeeded in securing this notion of sustainability.

We propose a return to first principles by formally defining sustainability as a steady state in the contribution rate and funded ratio. There are many such steady states, including the pay-go rate, at zero funding, and, conversely, many degrees of pre-funding, with their corresponding contribution rates. Thus, the nature of the steady state depends on the goals of the policy, as well as the plan parameters and assumptions, most notably the rate of return.

Although the analysis of steady states oversimplifies actual systems – even in expected value – since the parameters themselves never settle into steady states, the approach offers insights akin to other simple economic models. Steady-state analysis arguably lays out the characteristics of the system's trajectory, even if it is aiming at a moving target, as will be illustrated below in Section 5's simulation of CalSTRS. That said, as will be discussed in our Conclusion, further work, building on this baseline, would be useful to analyze more dramatic deviations from steady state, such as plans with rising dependency ratios or falling benefits, and correspondingly trending pay-go rates.

In the next subsections, we lay out our steady-state analysis, derived from the fundamental laws of motion of a plan's assets and liabilities, for a constant pay-go rate and payroll growth rate, to examine how the contribution rate rests on the funding target and portfolio return.

4.1 Steady-state condition for contribution rate and asset targets

The proximate determinant of the steady-state contribution rate is the asset target. As we will show, the asset target determines that portion of benefits to be covered by investment income, leaving the rest for the steady-state contribution rate. Liabilities only enter the picture as a benchmark for the asset target, linked by the target funded ratio. Thus, we consider the asset target first, generating insights of its own, and then bring in liabilities in the next subsection.

There are two sources of pension funding and two uses: contributions and investment income go to cover the payment of benefits and the accumulation of assets. Of these four flow variables, the stream of benefit payments is exogenous to our analysis (determined by the tiered benefit formulas and workforce assumptions), and investment income is governed by the sequentially determined stock of assets and the series of annual returns. This leaves the series of contributions and that of asset accumulation as mechanically linked. That is, the *funding policy is simultaneously a contribution policy and an asset accumulation policy.*

Formally, this relationship is captured in the fundamental asset growth equation:

$$A_{t+1} = A_t(1 + r_t) + \boldsymbol{c_t} W_t - c_t^p W_t, \tag{1}$$

where A_t denotes assets at the beginning of period t, r_t is the return in period t, W_t is payroll, while $\boldsymbol{c_t}$ and c_t^p are the contribution and benefit payment rates, respectively, as proportions of payroll (Table 1 lists notation). Assets grow by investment earnings, plus contributions, net of benefit payments. Equation (1) is simply an accounting identity. To give it economic content, for sustainability analysis, we need to specify a funding policy to drive $\boldsymbol{c_t}$. Given returns and benefit payments, the contribution policy sets asset growth. We will spell out our approach to the choice of contribution policy below, but even before delving into the specifics, equation (1) helps focus on the fundamental tradeoffs among these policies.

It will be useful to re-express equation (1) in terms of the ratio of assets to payroll, $a \equiv (A/W)$. Dividing through (1) by W_t, and denoting the growth rate of payroll by g, we have:

$$a_{t+1}(1 + g_t) = a_t(1 + r_t) + \boldsymbol{c_t} - c_t^p. \tag{1$'$}$$

The big picture here can be illuminated by examining the steady-state relationship between the contribution rate and assets. In steady state, the growth rate of assets must equal that of payroll, so the asset ratio is constant, $a_{t+1} = a_t = a^*$. Removing the time subscript for the steady-state values of the benefit payment rate c^p, the rate of return r, and the payroll growth rate g, we have the relationship between the steady-state contribution rate and asset ratio:

$$\boldsymbol{c}^* = c^p - (r - g)a^*. \tag{1*}$$

The interpretation is straightforward: benefit payments are covered by a mix of contributions and investment income (net of growth), where the mix is determined by the funding policy. Under a policy of pay-go, where no assets are accumulated ($a^* = 0$), the contribution rate must cover the benefits payment rate c^p. Under a policy of pre-funding, to one degree or another, the goal is to accumulate a certain asset level, a^*, so the income from those assets (net of growth) can help fund benefits, ultimately reducing reliance on contributions.

One very modest test of sustainability is to consider whether current contribution rates are sufficient to sustain a steady state at *current* asset levels. To be sure, that is not the goal of current policies,

Table 1. Pension funding notation

A = assets on hand
L = accrued liabilities, the present value of future benefits earned to date
f = funded ratio, A/L (full funding goal is f = 100%)
W = payroll
$a = A/W$ = assets/payroll
$\lambda = L/W$ = liabilities/payroll
$\boldsymbol{c}$ = contribution rate, % of payroll
c^p = benefit payments as % of payroll ('pay-go rate')
c^n = newly accrued liabilities as % of payroll ('normal cost rate')
r = return on assets; $R = (1 + r)$
d = discount rate used to calculate present value of liabilities; $D = (1 + d)$
g = growth rate of payroll; $G = (1 + g)$

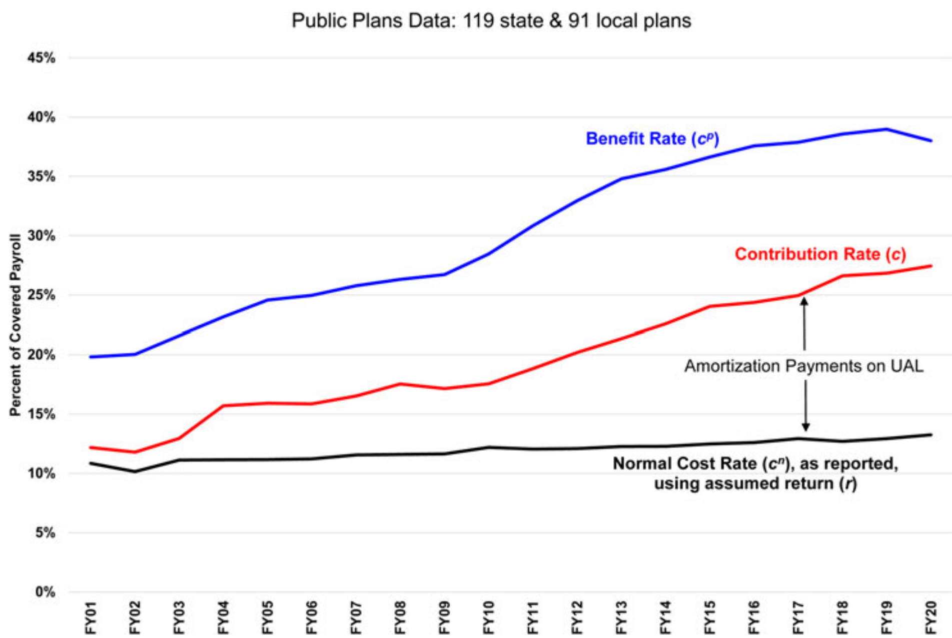

Figure 1. Normal cost, contribution and benefit rates, FY01–FY20.
Source: Center for Retirement Research at Boston College, MissionSquare Research Institute, and National Association of State Retirement Administrators.

which are attempting to raise asset levels to amortize pension debt. But if we consider the minimal target of $a^* = a_0$, would the current contribution rate, c_0, need to rise to sustain the asset level? Is $c_0 < c^*(a_0)$?[7] Let us consider recent trends and magnitudes.

Figure 1 depicts the aggregate values of $\boldsymbol{c_t}$ and c_t^p for FY01–FY20, of the 119 state and 91 local plans in the Public Plans Database (PPD), which account for 95 percent of state and local pension assets and members in the US. The contribution rate, as a percent of payroll, has been steadily climbing since the turn of the century, from about 12 to 27 percent.[8]

The benefit (or 'pay-go') rate has also trended up, from 20 to 38 percent, due in part to benefit increases enacted in the 1990s and early 2000s, but largely due to the aging workforce and falling number of actives per retiree. However, it may now be leveling off.[9]

It is important to note that throughout this period the benefit rate exceeds the contribution rate by a large margin, over 10 percentage points since 2010. That is, the primary cash flow (i.e., excluding investment income) is negative. Thus, if assets were to be depleted, contributions would have to jump to cover benefits. That is arguably the main cost of insolvency.[10]

Figure 2 depicts the asset ratio $a \equiv (A/W)$ from the same dataset. This has fluctuated with market returns, along with trends in benefit payments and contributions, but in recent years assets have hovered around a multiple of five times covered payroll.

[7]Equivalently, using (1*): is $a_0 < a^*(c_0)$? If so, then a policy of holding contributions at the current rate of c_0 would lead to a continual draw-down of assets, following the dynamic of (1′). That is, the current contribution rate would lead to insolvency. This is a special case of the convergence analysis discussed below.

[8]This includes employer and employee contributions. The FY20 mix is 20 and 7 percent, respectively.

[9]Lenney *et al.* (2019*a*, 2019*b*) project that the benefit rate will peak over the next decade and decline thereafter, as recent hires, in less generous tiers, enter retirement and beneficiaries of more generous tiers die off.

[10]It is important to distinguish between insolvency (i.e., reversion to pay-go) and default on benefits, which would violate taxpayer guarantees for a public plan, unlike private plans.

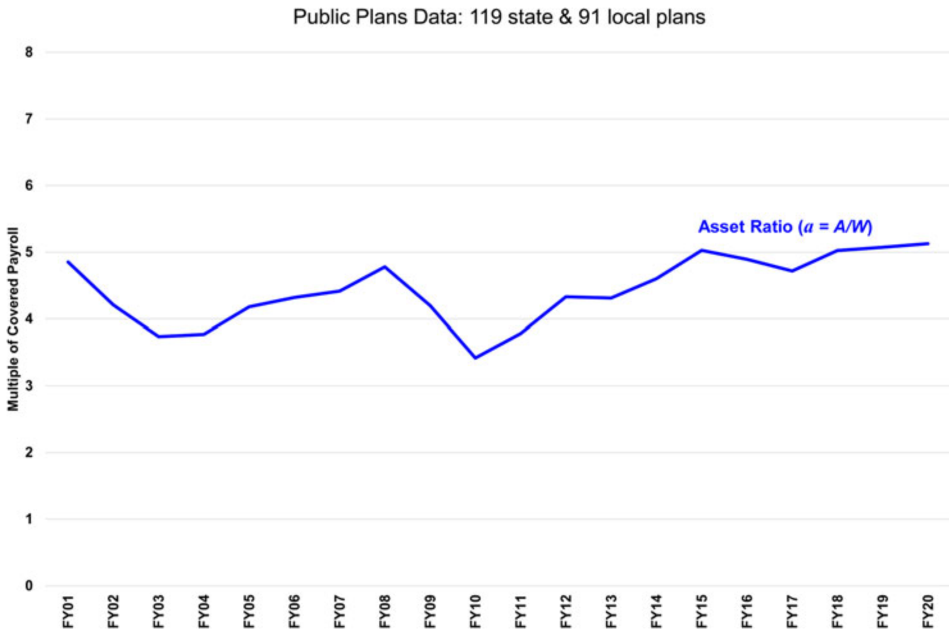

Figure 2. Assets/payroll, FY01–FY20.
Source: Center for Retirement Research at Boston College, MissionSquare Research Institute, and National Association of State Retirement Administrators.

With these data for a_0 and c^p, along with typical plan assumptions of $g = 3$ percent and $r = 7$ percent, we calculate $c^*(a_0) = c^p - (r - g)a_0 = 0.38 - (0.07{-}0.03) \times 5 = 0.18$. This is less than the current contribution rate, $c_0 = 0.27$. Thus, taken at face value, this would suggest that, *in the aggregate*, the current configuration is sustainable, and, indeed, that contribution rates could fall while still supporting current asset ratios. Of course, this depends on a host of assumptions, not least of which are the assumed rate of return and growth rate – specifically, the spread between the two. As long as $(r - g)$ exceeds about 2 percent (e.g., $r > 5\%$ for $g = 3\%$), the current overall contribution rate could sustain a_0, under this simple analysis.

This picture also generally holds for the individual plans in the PPD database. Using each plan's assumed return (the vast majority lie between 7.0 and 7.5 percent for FY20), we find that in 158 of the 188 plans for which $c^*(a_0)$ can be calculated, the contribution rate c_0 exceeds that value.[11] This also holds for 69 of the 79 largest plans, with assets exceeding $10 billion.

This result, however, is sensitive to the assumed return, or, more precisely, the assumed spread between r and g. Reducing each plans' assumed return to 5 percent (while holding $g = 3$ percent) changes the picture. Under this assumption, the contribution rate for most plans (107 of the 188 plans, and 48 of the largest 79 plans) is too low to sustain the current asset ratio. As this exercise illustrates, it is important to bear in mind that the steady states we examine are, at best, steady states in the expected value of contributions, with significant risk in actual outcomes.

Indeed, looking back over the time series depicted, even though c_0 has consistently exceeded $c^*(a_0)$ under the assumed spread between r and g, the asset ratio has not risen. Despite ever-rising contribution rates, the attempt to raise the asset ratio has generally failed. This not only indicates faulty assumptions; more fundamentally, it points to a failure of contribution policies to self-correct – a

[11]The assumed growth rate for payroll is only available in the PPD for 76 plans. Of those, the vast majority lie between 2.75 and 3.5 percent, so we set the growth rate at 3.0 percent for the calculation of $c^*(a_0)$ in all plans.

topic we return to below under the subject of convergence. But first, we turn to the role of liabilities in setting asset targets.

4.2 Steady-state condition for liabilities

We begin with the fundamental growth equation for liabilities:

$$L_{t+1} = L_t(1+d) + c_t^n W_t - c_t^p W_t, \tag{2}$$

where L_t denotes accrued liabilities at the beginning of period t, d is the discount rate, and c_t^n is the normal cost rate, the rate at which new liabilities accrue, as a percent of payroll. Liabilities grow by the interest on past liabilities, plus newly accrued liabilities, net of benefit payments that extinguish prior liabilities. Equation (2) is analogous to the asset growth equation (1), but with some key differences:

First, the formulation in (2) allows for a distinction between the discount rate d and the assumed (or expected) rate of return on assets r. Standard actuarial practice, of course, has traditionally discounted by r. By contrast, finance economics has consistently made the case that guaranteed benefits should be discounted by interest rates of correspondingly low-risk bonds, at least for accounting purposes (Brown and Wilcox, 2009; Novy-Marx and Rauh, 2009, 2011; Biggs, 2011). If asset accumulation and projections thereof reflect actual and assumed returns on a higher-risk pension fund portfolio, this raises the question of how a dual rate system should play out in contribution policy. In the previous subsection, focused on asset accumulation, the steady-state contribution policy depended only on r and not on d. We consider below how the consideration of liabilities, discounted at $d < r$, should factor into contribution policy.

The second difference between the liability growth equation (2) and the asset accumulation equation (1) is the role of c_t^n, the normal cost rate, vs. c_t, the contribution rate. The normal cost rate is independent of the contribution policy. It is determined by the benefit formula, the cohort's assumed separation probabilities over its members' careers, and (importantly) the discount rate.[12] This means equation (2) is logically prior to the asset accumulation and contribution equation (1). In our model, equation (2) will feed into (1*) by tying the asset target, a^*, to liabilities.

It will be useful to express (2) in the state variable $\lambda \equiv L/W$ = liabilities/payroll, using the same steps as in the derivation of (1′):

$$\lambda_{t+1}(1+g_t) = \lambda_t(1+d) + c_t^n - c_t^p. \tag{2′}$$

If we take the benefit formula and demographic/work-life assumptions as exogenous, along with d and g, then so are c^n and c^p. Thus, we can readily derive the steady-state liability ratio:

$$\lambda^* = (c^p - c^n)/(d-g). \tag{2*}$$

This expression has a simple interpretation. First note that the present value of future payroll in steady state is $W_t/(d-g)$, using the formula for a growing perpetuity. The PV of future benefit payments and liability accruals (normal costs) are, respectively, fractions c^p and c^n of the PV of future payroll. Thus, equation (2*)'s steady-state liability ratio, λ^*, represents the difference between the PV of future benefit payments and normal costs, scaled to current payroll.[13]

Figure 3 depicts the aggregate liability ratio, drawing again on the PPD, where the liabilities are reported based on each plan's assumed return, r. That ratio (depicted by the red curve) has gradually risen from about 4.6 in FY01 to about 7.2 in FY20, a rise of 56 percent. Several factors have

[12]It also depends on the specific actuarial cost method for allocating liabilities between past and future accruals. To fix ideas, we have in mind the standard entry age normal cost method.

[13]This follows from the basic identity that the PV of all future benefits equals the PV of benefits yet to be accrued (the PV of future normal costs) *plus* the PV of benefits previously accrued but not yet paid. The latter term is the accrued liability, so it equals the difference between the PV of all future benefits and the PV of future normal costs.

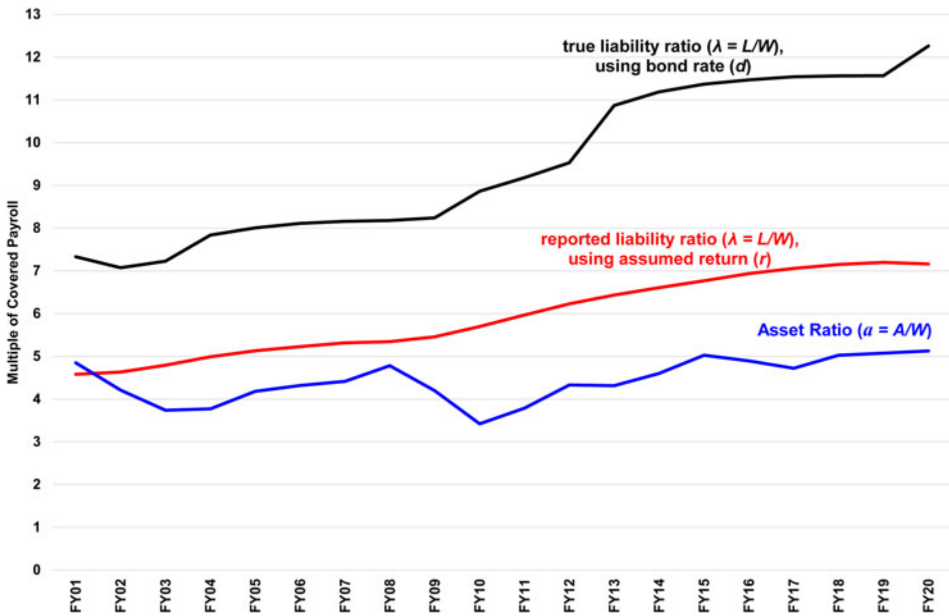

Figure 3. Assets and liabilities, true and reported, FY01–FY20.
Sources: Center for Retirement Research at Boston College, Federal Reserve Board of Governors & authors' calculations. Both series use PPD payroll.

contributed to this trend, including reductions in the assumed return and a rise in the ratio of retired to actives.[14]

Liabilities are much higher when discounted at a low-risk rate d, instead of r. Estimates vary regarding the magnitude of the impact. Here, we consider the liability estimates of the Federal Reserve Board of Governors, depicted by the black curve in Figure 3.[15] Comparing these estimates with the reported values of the PPD suggest that properly discounted liabilities are about 60 percent higher.[16] By this measure, the liability ratio has risen from about 7.3 in FY01 to about 12.3 in FY20, a rise of 67 percent.

Tying this together with the previous subsection, on asset accumulation, the actuarial goal of fully funding reported liabilities would raise the asset ratio from about five times payroll to seven. Although this goal has proven challenging, it still falls well short of matching true liabilities. The true funded ratio, upon accumulating assets of seven times payroll, would be about 7/12, or 58 percent. Considering the rise in contributions that would be needed to achieve this goal (examined in the next section), this may well be the limit of what is politically feasible or socially optimal under an objective function of the type we introduce in Section 6.

In any case, a more accurate label for the current policy would be something like '60 percent funding' (of true liabilities) rather than 'full funding'. More generally, as we will formalize below, the way to integrate risk-free discounting of liabilities into a policy of accumulating risky assets

[14]Benefit changes have also affected the trends, but in no simple fashion, as many plans raised benefits in the early 2000's and then cut them for new hires in the 2010's. Comparing the liability ratios with the calculated values of λ^* for FY01, FY10, and FY20, we find these values match for FY01 (4.6 vs. 4.5), but for FY10 and FY20, the liability ratios exceed the calculated values, 5.7 vs. 4.3 and 7.2 vs. 6.1, respectively. There are many potential explanations for these gaps, but they would be consistent with plans that are beyond mature, rather than in steady state.

[15]Federal Reserve series Z1/Z1/FL224190043. The denominator in the ratio depicted is the PPD payroll series.

[16]Lenny *et al.*'s estimates of the funded ratio, using reported liabilities (2021, Table 1) and rediscounted liabilities (2021, Table A7) imply that the latter is about 80 percent higher.

with higher expected returns is to set the target funded ratio relative to true liabilities. That target may well be less than 100 percent, but it would have the virtue of being accurate and, as we will show, such a policy will integrate the costs of risky investment with the benefits of reduced expected contributions.

4.3 Target funded ratio and the steady-state contribution rate

The natural link between our steady-state analysis of asset accumulation and liabilities is to tie the asset goal, a^*, to liabilities. We here consider the general goal of a target funded ratio, f^*, including both ‘full’ actuarial funding, and such putative standards as ‘the 80 percent rule’.[17] Setting the asset goal of $a^* = f^*\lambda^*$, and, for the moment, following the actuarial convention of $d = r$, we find, from (1*) and (2*):

$$\boldsymbol{c}^* = c^p - (r-g)f^*\lambda^* = c^p - f^*(c^p - c^n) = (1-f^*)c^p + f^*c^n. \tag{3}$$

As the funded goal varies from zero to full funding, the steady-state contribution rate varies from the pay-go rate to the normal cost rate, with a weighted average of the two for intermediate funding targets. Thus, full actuarial funding is a special case, where the steady-state contribution rate is c^n (discounted at r), reached upon completion of the amortization schedule.

Let us now consider the steady-state implications of a dual rate system: discount rate d for liabilities and assumed return r on assets. We then have:

$$\boldsymbol{c}^* = c^p - (r-g)f^*\lambda^* = c^p - [(r-g)/(d-g)]f^*(c^p - c^n). \tag{3$'$}$$

As before, if the funding goal f^* is zero, the contribution target is pay-go, and as f^* is set higher, the contribution target falls. Our question here is the impact on c^* of reducing d below r. We have already seen from (1*) that the only avenue for a drop in d to affect c^* is through its impact on the asset target a^*. Since we are considering asset goals of the form $a^* = f^*\lambda^*$, this means that a drop in d below r would raise the target contribution rate through a rise in the liability ratio λ^* unless it is offset by a reduction in the target funded ratio f^*.

If, for example, our funding goal is to merely maintain the current asset ratio, $a^* = a_0$, then the rise in λ^* from revaluation at d would, in effect, be completely offset by an implicit drop in the target funded ratio f^*.[18] Under this simple goal, the distinct role of d drops out of (3′), and we are back at (1*) with $\boldsymbol{c}^* = c^p - (r-g)a_0$. Setting d to a low-risk rate for the valuation of liabilities is here purely an accounting and reporting measure, unrelated to contribution policy, as discussed in Costrell and McGee (2020).

More generally, let us consider the implications of setting $d < r$ when f^* is a deliberately chosen target (as discussed in a later section), rather than an artifact of maintaining the status quo. The first implication is that under a full-funding policy, $f^* = 1$ (or anywhere near full), $c^* < c^n$: contributions will not cover normal costs (properly discounted). Formally, (3′) implies

$$\boldsymbol{c}^* - c^n = (c^p - c^n)(1 - [(r-g)/(d-g)]f^*) < 0, \text{ for } f^* > [(d-g)/(r-g)]. \tag{3$''$}$$

To fix magnitudes here, consider the values we have been using, $r = 0.07$ and $g = 0.03$, along with $d = 0.04$ (a typical discount rate used in private pension accounting). The critical value of f^* in the expression above is then 25 percent. For any target funded ratio exceeding 25 percent of true liabilities, steady-state contributions need not cover the true normal costs (discounted at d). This contrasts starkly with standard actuarial funding schedules, under which contribution rates drop to (but not below) reported c^n (discounted at r) upon reaching full funding.

[17]See Costrell (2018*a*), where equation (3) was previously derived.

[18]Specifically, the rise in λ^* effectively reduces f^* to a_0/λ^*, so (3′) simplifies to $\boldsymbol{c}^* = c^p - (r-g)a_0$. This is implicit in the model of Lenney *et al.* (2019*a*, 2019*b*) and Lenney *et al.* (2021).

The point can be illuminated by re-writing ($3'$) and simplifying to obtain:[19]

$$\boldsymbol{c}^* = c^p - (d-g)f^*\lambda^* - (r-d)f^*\lambda^* = (1-f^*)c^p + f^*c^n - (r-d)f^*\lambda^*. \tag{3*}$$

Comparing with (3), where $d = r$, we have a higher (rediscounted) normal cost rate, but the third term, which is negative, is new. Under a deterministic interpretation of c^* and r, this represents the implicitly assumed arbitrage profits between the return on accumulated assets and interest on covered liabilities. These assumed arbitrage profits help defray the higher normal costs, in lieu of contributions that might otherwise be required.

Alternatively, if c^* and r are understood to be expected values of risky variables, the last term may be interpreted as the risk premium on the portfolio. While this reduces the expected contribution rate, it simultaneously mirrors the implicit cost of risk borne by the sponsoring government. This expression nicely captures the tradeoff between risk and return, to which we return in Section 6.

5. Contribution policy for convergence toward steady state

Although steady-state calculations are instructive, they are not compelling unless there is a dynamic process that converges toward a steady state. Moreover, that dynamic process, once determined, informs us of the expected costs along a transition path to the target asset ratio a^*. Of course, the steady state is always a moving target, as the parameters c^p, r, and g vary over time, but we can analyze whether the system moves in the right direction, taking these parameters as constants, at their projected values.

In assessing the path to a^*, we need not concern ourselves for the moment with the dynamics of the liability ratio, λ_t, given in ($2'$), although we will return to this below when we illustrate with CalSTRS data. Once we determine the steady-state liability ratio, λ^* from (2^*), and choose the target funded ratio f^*, we have the target asset ratio $a^* = f^*\lambda^*$. This is all we need to map out a transition path for a_t, using equation ($1'$) and a contribution policy c_t to be specified. As we will show, even with these simplifications, the determination of a smoothly convergent contribution policy is non-trivial.

Convergence is not automatically assured, as can be discerned by considering the asset accumulation equation ($1'$) alone (before adding in a contribution policy equation). To simplify notation, let $R = 1 + r$, $G = 1 + g$, and re-express ($1'$) as:

$$a_{t+1} = a_t(R/G) + (c_t - c^p)/G. \tag{1''}$$

For $R > G$ (as usually assumed), the coefficient on the prior value of the state variable a exceeds one, which is destabilizing.

For example, consider a simple policy that sets the contribution rate to some target rate and holds it constant.[20] Unless that target rate corresponds to the steady-state value for maintaining the *current* asset ratio, the system will diverge. Stated alternatively, suppose one aims at an asset ratio $a^* \neq a_0$, and immediately sets $c = c^*$ (using (1^*)), jumping up or down from c_0, and holding it there. The system will then move away from a^*, rather than toward it. If a^* is set greater than a_0, then a_t, will shrink further away from a^*, and conversely if a^* is set lower than a_0.[21]

The reason is straightforward. Setting a higher a^* means setting a lower c^* (see equation (1^*)), since one expects to rely on higher investment income, in lieu of contributions, to cover benefits. But since assets are not yet at that higher level of a^*, the investment income falls short of that which would obtain in the steady state to which one aspires. Thus, by prematurely setting contributions at the correspondingly low level, c^*, one embarks on a path of asset decumulation. And conversely for $a^* < a_0$.

[19]We can drop the assumption of steady state in λ and obtain an expression with the same pieces and the same interpretation: $\boldsymbol{c}_t = (1-f^*)c^p + f^*c^n - (r-d)f^*\lambda_t$. This generalization of (3^*) loosens the condition that assets and liabilities grow at the rate g; we require only that they grow at the same rate as each other, so that f is constant at f^*.

[20]Many states set a fixed rate in statute, instead of an actuarially varying rate. Similarly, the Lenney *et al.* (2019*a*, 2019*b*) policy simulation sets c equal to a steady-state value and holds it there.

[21]Formally, the solution is $a_t = a^* + (R/G)^t(a_0 - a^*)$, which continually magnifies any initial deviation from a^*.

So, what would a contribution policy look like that converges to a steady state targeted at a^* with contributions c^*? It might be thought that an adjustment process that gradually closes the *contribution gap* between c^* and c_t, rather than a sudden jump to c^*, would do the job, but as we shall see below, it will not. The reason, as would be suggested by the discussion above, is that the contribution required to cover benefits depends on the *asset gap* between a^* and a_t. Alternatively, then, one might suppose that an adjustment process for contributions based on the *asset gap* would fit the bill. However, as we shall see, that will not suffice either. For a convergent path, we will show that the policy should adjust contributions based on *both* gaps,[22] between c^* and c_t and between a^* and a_t, in combinations to be derived below.

Before doing so, note that the policy we are deriving differs not only from a discrete jump to c^*, but also from the trajectory of actuarial funding policy. The actuarial payment schedule typically sets either a constant percent of payroll, or ramps up to such a rate, and then falls off a cliff at the end of the amortization period, once full funding is expected to be achieved.[23] The policy we derive below aims to converge smoothly on a steady state.

Specifically, consider a contribution policy that starts by specifying a target asset ratio, a^* (more on how that might be chosen, in Section 6), and calculates the corresponding steady-state contribution rate c^*, using (1*) above. We then annually adjust the contribution rate based on the gaps between the target and actual values for assets and contributions:

$$c_{t+1} = c_t + \beta(c^* - c_t) + \gamma(a^* - a_t), \text{ where } \beta \in (0, 1). \tag{4}$$

Along with (1″), we have a simple system of two linear difference equations to be analyzed using standard methods. We derive the bounds on β and γ needed for convergence in the Appendix.

The first convergence condition (see Appendix) is $\gamma > \beta(R - G) \equiv \gamma_{min}$. This is positive for $R > G$, thereby showing formally what was asserted above: a piece of the adjustment mechanism must be based on the asset gap, not just that of the contribution rate. The logic is straightforward. Suppose the contribution rate is already at its target c^*, but the asset level is below the target a^*. Then contributions will have to rise in the short run to accumulate more assets, before eventually dropping back down toward c^*.

The second convergence condition, $\gamma < G - R(1 - \beta) \equiv \gamma_{max}$, implies that the adjustment mechanism must include the contribution gap, too, $\beta > 0$. Formally, since we must have $\gamma_{max} > \gamma_{min}$, this requires $\beta > (R - G)/G > 0$. The logic here is also straightforward. If assets are at their target ratio, but the contribution rate is below c^*, then c_t needs to rise.

As our discussion above suggests, the convergence toward steady state may not be monotonic. Indeed, it may not only reverse direction once (asymptotically monotonic), it may be oscillatory. The condition for non-oscillatory behavior, also given in the Appendix, is $\gamma < G[(R/G) - (1 - \beta)]^2/4 \equiv \gamma_{m/o}$, where the subscript m/o denotes the boundary between monotonic and oscillatory. It can be shown that for $\gamma_{max} > \gamma_{min}$, $\gamma_{m/o}$ falls between the two.

Thus, the asymptotic behavior of the system varies with the range of γ as given in Table 2. Figure 4 illustrates the combinations of β and γ that correspond to these asymptotic behaviors for $r = 7$ percent and $g = 3$ percent. In general, it seems reasonable to presume that policymakers would prefer non-oscillatory convergence. Thus, the relevant combinations of β and γ would lie between γ_{min} and $\gamma_{m/o}$, depicted by the black and blue curves in Figure 4.

5.1 Deterministic simulations: representative plan

Armed with these analytics, we illustrate the dynamic paths for contributions and assets under plausible policies. Taking the representative FY20 plan assumptions given above, $R = 1.07$, $G = 1.03$, $c^p = 0.38$, $c_0 = 0.27$, and $a_0 = 5$, we set the target ratio $a^* = 7$. As discussed earlier, this increase of 40 percent above a_0 would accumulate approximately the assets needed for 'full' actuarial funding (discounted at

[22]Interestingly, the Taylor rule for optimal monetary policy is also a two-gap rule.

[23]This refers to 'closed interval' amortization. 'Open interval' amortization extends the payoff date each year. Although commonly used in times past, open interval is no longer recommended by GASB.

Table 2. Convergence conditions for adjustment parameters β and γ: $c_{t+1} = c_t + \beta(c^* - c_t) + \gamma(a^* - a_t)$

Range of γ	Asymptotic behavior of (1″)–(4)
$\gamma < \gamma_{min} \equiv \beta(R - G)$	Monotonic divergence
$\gamma_{min} < \gamma < \gamma_{m/o} \equiv G[(R/G) - (1 - \beta)]^2/4$	Monotonic convergence
$\gamma_{m/o} < \gamma < \gamma_{max} \equiv G - R(1 - \beta)$	Oscillatory convergence
$\gamma_{max} < \gamma$	Oscillatory divergence

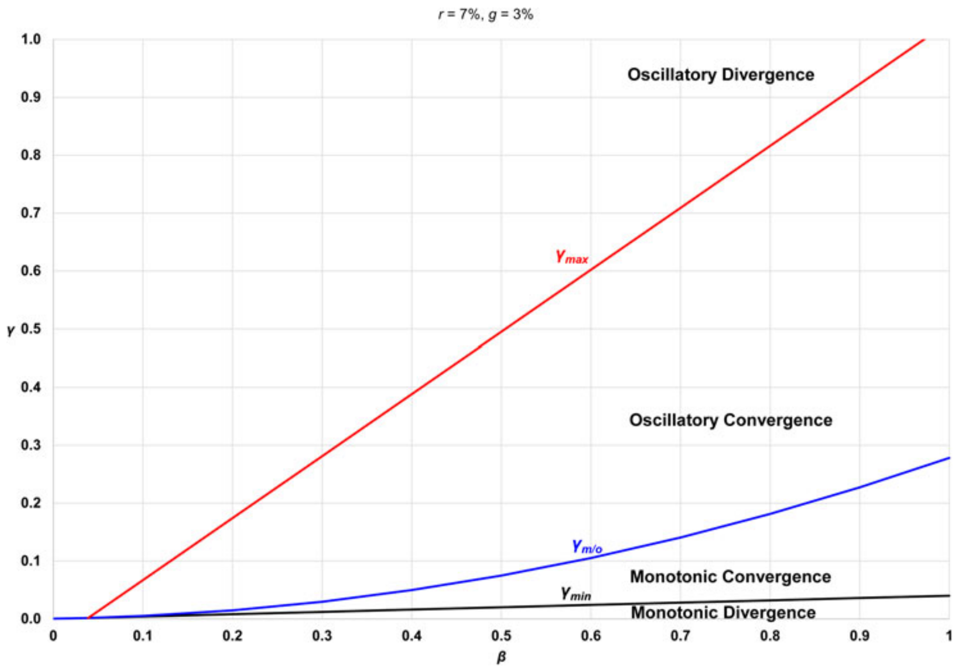

Figure 4. Asymptotic behavior of asset accumulation and contribution rate.

the expected return), or about 60 percent of true liabilities (discounted at a low-risk bond rate). Equation (1*) gives us $c^* = 0.38 - (0.07 - 0.03) \times 7 = 0.10 < c_0 = 0.27$, thus allowing eventually for a dramatically lower contribution rate.[24]

The choice of adjustment parameters β and γ must navigate an intertemporal policy tradeoff. Contributions need to rise in the short run to accumulate the assets required for the long-term reduction in c^*. Thus, the tradeoff is between speed of reaching c^* vs. tempering the short-term rise in c required to reach a^*.

Suppose we set a target of approaching c^* by year 30 (corresponding to a somewhat conventional time horizon for actuarial amortization schedules) and set the contribution adjustment parameter β equal to 0.5 (half speed). Then we find that the tradeoffs are plausibly managed by choosing the asset adjustment parameter γ near the maximum value for monotonic convergence, $\gamma_{m/o} = 0.075$, on the blue curve in Figure 4.

Figure 5 depicts the corresponding paths for the contribution rate (red curve, on the right scale) and asset ratio (blue curve, on the left scale). This path raises the contribution rate for about seven years to a maximum of 36 percent (a nine-point hike), before ultimately dropping down to

[24]This estimate of c^* is close to the average normal cost rate of about 13 percent, as reported using plans' assumed return (depicted in Figure 1), and consistent with the steady-state result in (3) for 'full' actuarial funding $c^* = c^n(r)$.

https://doi.org/10.1017/S1474747223000173 Published online by Cambridge University Press

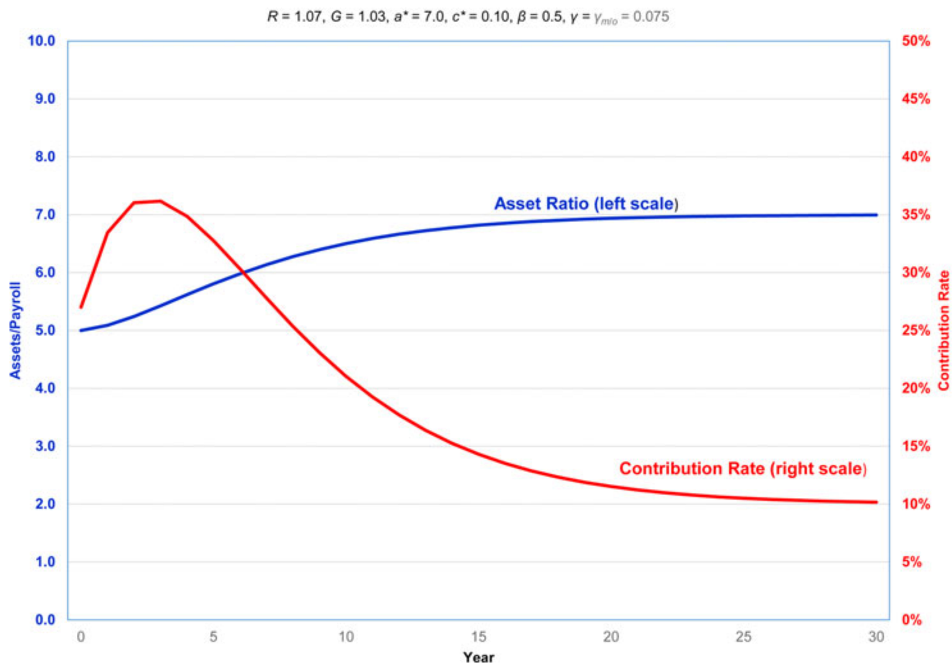

Figure 5. Simulation of trajectory to 'full' actuarial funding.

approximately ten percent by year 30. Setting β any faster requires a sharper short-term rise in contributions and setting it any slower fails to approach c^* that closely in 30 years.

Our dynamic analysis shows how to generate a smooth adjustment path to 'full' funding, unlike the actuarial scenario of the contribution cliff envisioned upon completion of a closed interval amortization schedule.[25] The path depicted is challenging: it calls for a substantial rise in contributions over the near term.

5.2 Deterministic simulations: CalSTRS

We consider the case of CalSTRS to illustrate the contrast with actuarial funding schedules for a plan where the steady state is a moving target. Not only is CalSTRS of general interest as one of the country's largest plans, but CalSTRS also provides useful cash flow projections out to the end of its amortization schedule in 2046. Importantly, this allows us to replace a constant benefit rate c^p, with its projected rate, which rises significantly, from 47 to 57 percent over the projection period, before starting to edge back down.[26] We use these time-varying cash flows in our simulation, which has important impacts on the trajectory of asset accumulation and, accordingly, our adjustment path for contributions.

In addition, the CalSTRS projections include liability figures. These imply the liability ratio λ rises and falls over the projection period, as the plan gets over the hump in benefit payments. Our simulation sets the target asset ratio to this moving target of 'full' funding, $a^*(t) = \lambda(t)$ to keep assets from falling too far behind the ultimate target.

The steady-state contribution rate is the reported normal cost rate $c^n(r)$, both under actuarial funding schedules and in our model, when the asset target is 'full' actuarial funding (as shown in (3)). Here, too, we let the target contribution rate $c^* = c^n(r)$, vary over time as provided by CalSTRS in its

[25]Open interval amortization has no such cliff, but often never approaches the funding target (see, e.g., Costrell, 2018*b*, Figure 3).

[26]This pattern reflects the rising ratio of retirees to actives and the retirement of later cohorts with lower benefits.

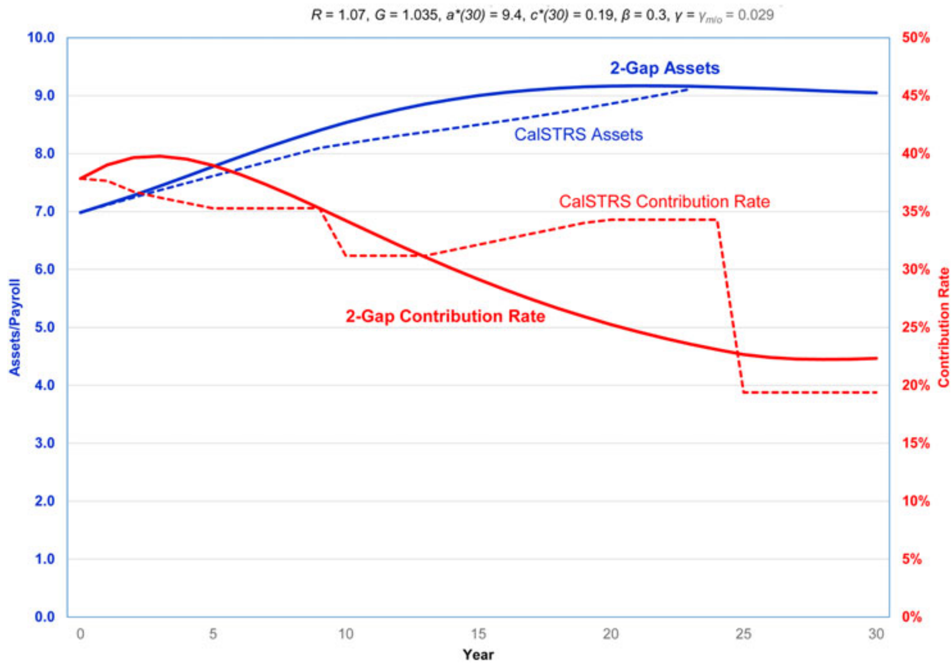

Figure 6. Contribution trajectory, CalSTRS vs. two-gap adjustment.

projections, although the variation is minimal, from 20.4 to 19.4 percent. Finally, CalSTRS sets $R = 1.07$, and $G = 1.035$, both of which we take as constants.

Our two-gap adjustment process (4), with the convergence bounds depicted in Figure 4, is designed for constant parameters, rather than the time-varying ones we see in practice. So it is of some interest to see how well (or poorly) our process converges. Given the rise in projected benefit payments discussed above, we chose somewhat slower adjustment parameters (e.g., β equal to 0.3 instead of 0.5) to keep the contribution amplitude manageable. The resulting trajectories for contributions and asset accumulation are depicted in Figure 6.

For comparison, CalSTRS' projected trajectories, under their actuarial funding policies, are also depicted. As with most actuarial schedules, the contribution rate is scheduled to drop precipitously at the conclusion of the amortization period in 2046, from 34 percent to the normal cost rate of 19 percent.[27] The CalSTRS schedule also features a notable drop in year 9, followed by a gradual rise through the end of the 24-year schedule. Such an idiosyncratic feature is not typical of actuarial schedules but does seem to illustrate how plans can arbitrarily schedule contribution relief after a good year in the market or other such developments.[28]

Our alternative two-gap schedule rises gently for a few years and exceeds the CalSTRS schedule for 13 years, leading to more rapid asset accumulation, before dropping below the last 11 years of CalSTRS' backloaded schedule. By 2046, under the parameter values we have chosen for the two-gap adjustment, contribution rates continue to smoothly decline, but would not yet drop to their steady-state value, unlike the CalSTRS schedule. In short, the two-gap adjustment process traces out a smoother trajectory, eschewing opportunistic backloading and a less-than-credible future funding cliff, for a more generationally rational contribution path.

[27]Under current provisions, the state will provide an additional contribution of 3.5 percent, over and above the member and employer normal cost.

[28]CalSTRS features a volatile state contribution, on top of member and employer contributions, that fluctuates with asset gains and losses in a convoluted formula that amortizes past unfunded liabilities but generates new ones.

Even so, the potential long-term gain is by no means certain. As noted previously, c^* is highly sensitive to the assumed return. Returning to our example of the representative fund depicted in Figure 5, at $R = 1.05$, instead of 1.07, $c^* = 0.24$ instead of 0.10. In this case, the accumulation of assets requires a much larger hike in short-run contributions (over 20 points, not shown), for very little gain in the long-run. Moreover, even if the assumed return is accurate in expected value, the distribution of outcomes can be very wide under stochastic returns.

5.3 Stochastic simulations

Let us consider a stochastic model of the two-gap path to 'full' actuarial funding for our representative plan. We ran Monte Carlo simulations of the adjustment path, with R distributed as lognormal, mean 1.07, and standard deviation of 0.15.[29] Figure 7 depicts the trajectories for the contribution rate and asset ratio at the median, 25th, and 75th percentiles of their distributions.

The first point to note is that the median of these distributions is indistinguishable from the deterministic trajectories depicted in Figure 5. The mean (not depicted) differs, due to the asymmetry of the lognormal distribution, but would also track the deterministic case if the distribution were symmetric; this gives us some basis for interpreting the steady-state values we derived analytically as expected values, which will be useful in Section 6.

That said, Figure 7 clearly shows the huge risk in these trajectories, as illustrated by the spread between the 25th and 75th percentiles. Such risks are unavoidable, with investment in risky assets. Moreover, the risk rises (the spread widens) over time, contrary to the popular notion that the good years and the bad years 'average out' over time.[30] Given the assumed asset allocation, the only latitude in managing that risk is the degree to which it falls on the contribution rate or the asset ratio. Under the adjustment parameters depicted, which seemed reasonable for the deterministic case ($\beta = 0.5$, $\gamma = 0.075$), quite a bit of the risk falls on the contribution rate: the 25th–75th percentile spread widens to over 50 percentage points by year 30.[31,32]

The contribution risk can be reduced, pushing it instead onto asset risk, by dampening the adjustment parameters. For example, if we cut γ in half to 0.0375, choosing a slower trajectory toward the targets of $a^* = 7$ and $c^* = 0.10$, we find a somewhat smaller contribution risk, as depicted in Figure 8. The 25th–75th percentile spread is narrower than in Figure 7 (about 35 percentage points, instead of 50) and the spread for the asset ratio (not shown, for purposes of clarity) is somewhat wider. That said, as in the previous case, the 25th-percentile asset ratio never dips as low as 4, and the risk of insolvency is negligible (unlike policy simulations with a constant contribution rate in papers mentioned above).

To reiterate, these simulations are not meant to be policy prescriptions, but rather to illustrate how contribution policy might be reformulated, using adjustment parameters toward asset and contribution targets. We believe this approach holds promise, compared to the current actuarial approach,

[29]We estimated the standard deviation values associated with specific target returns using the publicly available, forward looking capital market assumptions published by Callan in early 2020 (pre-pandemic). We estimated the portfolio allocation that would generate each target return across a diversified portfolio including large cap US equities (e.g., S&P 500), small/mid Cap US equities (e.g., Russell 2500), Global ex-US Equity (e.g., MSCI ACWI ex USA), real estate (e.g., NCREIF ODCE), private equity (e.g., Cambridge Private Equity), and aggregate US bonds (e.g., Bloomberg Barclays Aggregate). We then applied that allocation using Callan's estimated standard deviation and asset class correlations to calculate the associated standard deviation values for each return.

The risk–return profiles generated by Callan and other vendors, used by plan managers, build on the insights of Merton's (1973) three-fund theorem. Portfolios that incorporate bonds (and other assets) negatively correlated with equities are superior to those of the basic CAPM asset mix of risk-free Treasuries and stock market funds.

[30]The law of large numbers applies to the average annual rate of return, but not to the total return, or the assets on which the return is earned. Figures 7 and 8 visually illustrate the Fallacy of Time Diversification.

[31]The high probability of contributions going negative represents the chance of a run of good returns leading to asset accumulation far beyond the target, so excess assets are drawn down to pay benefits.

[32]Similar results hold for our CalSTRS simulation, depicted in Appendix Figure A. Here we use CalSTRS' (2023*b*) assumed standard deviation of 11.6 percent. The 25th–75th percentile contribution rate spread is wider under the Callan assumption of 15.1 percent.

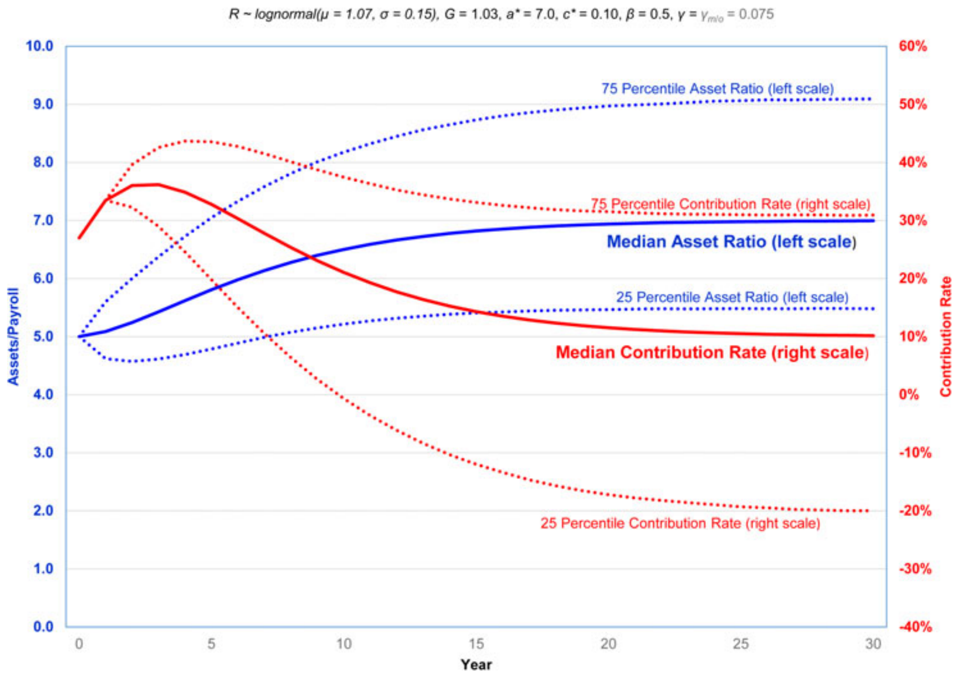

Figure 7. Stochastic simulation of trajectory to 'full' actuarial funding.

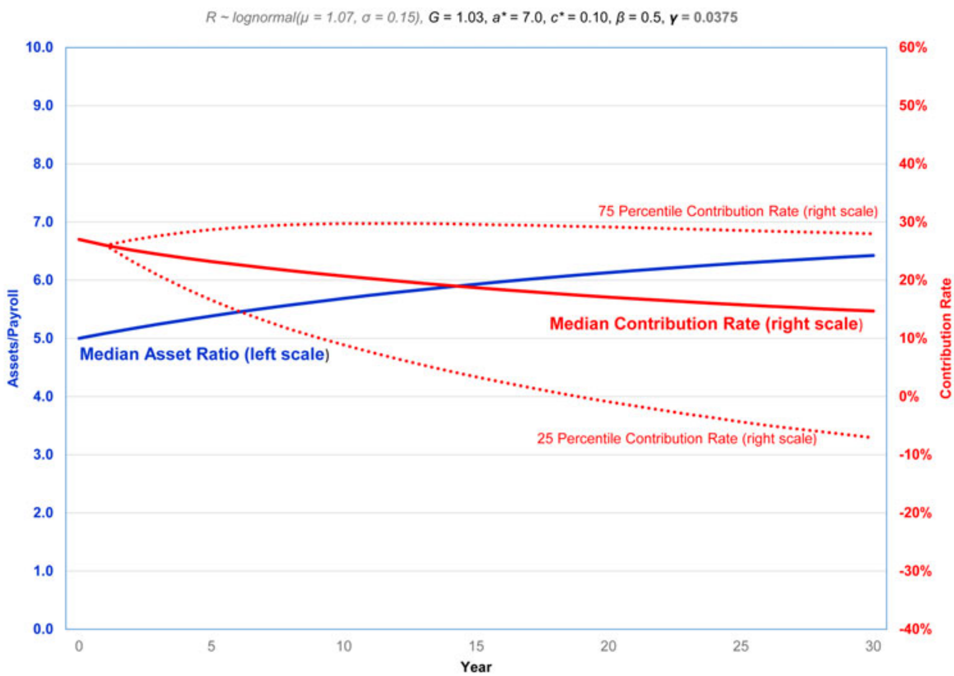

Figure 8. Stochastic simulation of slower trajectory to 'full' actuarial funding.

even as it would require further refinement. Such a policy would not mitigate risk – only a change in asset allocation would do that (as discussed in Section 6) – but illustrates how it can be deliberatively used to apportion risk between assets and contributions.

It is, perhaps, noteworthy that the approach set out here is consistent with a formal result from stochastic control theory. In a model with one state variable and one control variable (here, the asset ratio and contribution rate, respectively), and a quadratic loss function in the two variables, the optimal control is of the type we have considered: linear in the two gaps. Moreover, as shown by Turnovsky (1974), the introduction of stochastic elements dampens the adjustment in the optimal control, resulting in slower response in the control variable, consistent with the example we have illustrated here in Figure 8 vs. Figure 5.

Finally, we need an anchor for the asset accumulation goal. That anchor has traditionally been based on liabilities, and reasonably so, but, as we have argued above, the target should be based on true liabilities, appropriately discounted. Whether that target should be 100 percent of true liabilities or 60 percent, as in the $a^* = 7$ simulations depicted (corresponding to 100 percent of actuarial liabilities) or something less, and how to approach that question, is the subject to which we now turn. Our goal is, first, to integrate the insights from our previous analysis of the steady-state (expected) contribution rate, and the risk thereof, into a simple decision framework for the target funded ratio and asset allocation. Finally, using the asset allocation condition and the risk–return tradeoff, we can provide new insight into the bases for asset accumulation.

6. Choosing the target funded ratio: a stylized optimization framework

The approach we sketch out here begins with a very simple (and, hence, only semi-formal) objective function representing the two main tradeoffs: short-run vs. long-run costs and expected future contributions vs. risk. For a normative interpretation, this may be considered a social welfare function, or, alternatively, for a positive interpretation, we may consider it as the policymakers' stylized objective function (which may depart from the taxpayers' interest).

We abstract from specific features highlighted in the literature above, such as distortionary taxation, developed to model departures from Ricardian and MM irrelevance. Instead, our setup can be thought of as a 'reduced form' incorporating these various frictions or, instead, representing policy-makers' implicit assumption that taxpayers and investors simply do not offset funding policies by their individual actions, as the irrelevance theorems maintain.

Given this setup, we analyze the joint optimization of the funded ratio and asset risk. Using our results from (3*) in this framework allows us to present the marginal benefits and costs of these choices in readily understood terms and helps clarify the logic of pre-funding, incorporating the public's (or policymaker's) tolerance for risk imposed on future taxpayers.

Specifically, our optimization analysis exploits the insights from our steady-state result that distinguishes between r and d. In so doing, we arguably help resolve a bit of schizophrenia in the debate over actuarial discount rates. It is increasingly (if grudgingly) recognized, even by non-economists, that liabilities should be discounted at a low-risk rate corresponding to the guarantee of promised benefits. Yet finance economists typically restrict their conclusion to reporting requirements, and not necessarily to funding policy. Our approach integrates dual rates – specifically, the risk premium $(r - d)$ – into a framework for funding policy that simultaneously represents the benefits and costs of funding future pension payments from risky assets.

Here is our framework. In general terms, the asset risk and target funded ratio should be based on the preferences (social or policymaker's) for intergenerational sharing of cost and risk. Thus, we posit a stylized objective function:

$$- V[(a^* - a_0), E(c^*), \sigma(c^*)],$$

where $(a^* - a_0)$ is a shorthand measure of the costs required (non-discounted) over some period to reach the asset target; $E(c^*)$ is the expected value of the steady-state contribution rate at the asset target;

and $\sigma(c^*)$ is the risk of c^* from relying on asset income. Since these arguments to V are 'bads', we preface V with a minus sign,[33] so the partials V_1, V_2, and V_3 are positive.

More precisely, let us think of c^* as the contribution rate c_t as t gets large (e.g., $t = 30$), by which point $E(c_t)$ approaches the steady-state value $E(c^*)$ given by (3*). Similarly, we can think of $\sigma(c^*)$ as the corresponding contribution risk, $\sigma(c_t)$, but it is not a steady-state value, since risk continually rises, as discussed above and shown below. We may instead think of it as the future contribution risk as the expected contribution rate approaches its steady-state value.

Specifically, for analytical purposes we posit:

$$\sigma(c_t) = s(\sigma(r); t, \beta, \gamma)f^*\lambda^* \equiv S(r; t, \beta, \gamma)f^*\lambda^*. \tag{5}$$

This formulation lets the contribution risk rise with time, as illustrated in our simulations below. At any given time (e.g., $t = 30$), we express the contribution risk as a general function s of the annual risk of r, $\sigma(r)$, which itself is a function of the annual risk premium, $(r - d)$, where d is considered a parameter. As we shall see, the shape of the composite function $S(r; t = 30, \beta, \gamma) \equiv s(\sigma(r); t = 30, \beta, \gamma)$ will be important below but is left unspecified here.

The one substantive assumption in (5) is that the long-term contribution risk is multiplicative in the target asset ratio $a^* = f^*\lambda^*$. This assumption is analytically convenient, as it mirrors the fashion in which $a^* = f^*\lambda^*$ enters $E(c^*)$ in (3*). Fortunately, this assumption appears to be a very close approximation for our simulations of $\sigma(c_{30})$ – which we take to illustrate $\sigma(c^*)$ – as we vary a^* in the relevant range.[34]

The key take-away from this formulation is the parallel role of $(r - d)$ in $E(c^*)$ and $\sigma(c^*)$ in (3*) and (5) (via $\sigma(r)$). Thus, we simultaneously represent both the benefit and cost of the chosen degree of risk.

6.1 Optimal asset risk

The optimization problem requires a joint decision on two instruments: (i) the risk profile of the asset allocation, formally represented by the target return r (for given d); and (ii) the target funded ratio, f^*. Taking these in turn, we first optimize $-V[(a^* - a_0), E(c^*), \sigma(c^*)]$ over r, conditional on the funding target f^*. The choice variable r enters $E(c^*)$ and $\sigma(c^*)$ through the risk premium, $(r - d)$. Thus, from (3*), we have $dE(c^*)/dr = -f^*\lambda^*$, and from (5), we have $d\sigma(c^*)/dr = S'(r)f^*\lambda^*$, where, for notational simplicity, we omit the parameters, $(t = 30, \beta, \gamma)$. Consequently,

$$-dV[(a^* - a_0), E(c^*), \sigma(c^*)]/dr = -V_2 dE(c^*)/dr - V_3 d\sigma(c^*)/dr = [V_2 - V_3 S'(r)]f^*\lambda^*. \tag{6}$$

The bracketed term simply represents the balance of weights between additional risk and return. We assume there is an interior optimum for (6) with $r > d$,[35] where $V_2 = V_3 S'(r)$.

Figures 9a and 9b illustrate the tradeoff between the future contribution rate and risk for selected target rates of return. These simulations are similar to those above but depict risk with the standard deviation of the contribution rate over time, rather than the 25th–75th percentile spread. As above, we use the Callan (2020) capital market assumptions, but vary the composition of the portfolio to obtain the target returns depicted. These assumptions generate $\sigma(r)$, which feed into the simulations that generate outcomes for c_t.

Figure 9a depicts the trajectories for the median contribution rate, as r varies from 4 to 7 percent.[36] The lower trajectories for higher target returns illustrate the benefit, in lower future contributions, from more aggressive asset allocations.

[33]Equivalently, we could cast the problem as minimizing $V[(a^* - a_0), E(c^*), \sigma(c^*)]$.

[34]Specifically, we find that as we raise a^* from 7 to 9, the ratio of $\sigma(c_{30})$ to a^* varies by less than 1 percent.

[35]This would hold, for example, with a quadratic objective function.

[36]As noted earlier, the mean contribution rate is lower due to the asymmetry of the distribution of r under the lognormal assumption. This deviation between mean and median contribution rate is fairly minimal by year 30 for $r = 4$ and 5 percent. It widens notably by year 30 for $r = 6$ percent and quite substantially for $r = 7$ percent.

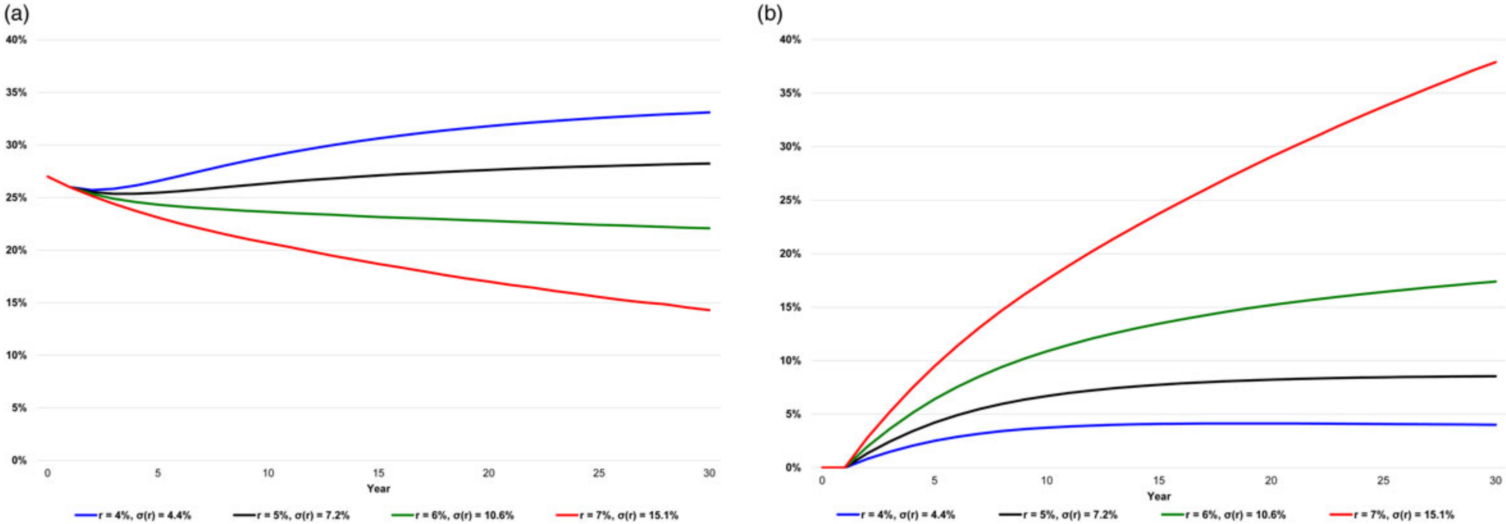

Figure 9. (a) Median contribution rate, varying risk and return. (b). Standard deviation of contribution rate, varying risk and return.

https://doi.org/10.1017/S1474747223000173 Published online by Cambridge University Press

Figure 9b depicts trajectories for the standard deviation of the contribution rate, as r varies.[37] The higher trajectories for higher target returns illustrate the cost from riskier asset allocations. Moreover, the rate at which the risk rises (the gaps between the curves in Figure 9b) exceeds that of the benefit (the drop between the curves in Figure 9a). That is, the 'price' of seeking higher returns rises as the plan gets more aggressive.[38] This corresponds to the convexity of the composite function $S(r)$ in (5), as will be important below.[39]

6.2 Optimal target funded ratio

We now turn to our main focus, the optimization of $-V[(a^* - a_0), E(c^*), \sigma(c^*)]$ over the choice variable f^*. We consider the marginal impact of raising the funded ratio. Qualitatively, the benefit is the reduction in the long-run expected contribution rate, $E(c^*)$, and the two costs are the short-run rise in contributions to accumulate more assets, $(a^*–a_0)$, and the increased risk of c^*, from raising the portion of future benefit payments defrayed by risky investment income.

Formally, we consider the three pieces of

$$-dV[(a^* - a_0), E(c^*), \sigma(c^*)]/df^* = -V_1 da^*/df^* - V_2 dE(c^*)/df^* - V_3 d\sigma(c^*)/df^*. \tag{7}$$

The first piece is the marginal cost in the short run to accumulate more assets, $a^* = f^*\lambda^*$:

$$-V_1 da^*/df^* = -V_1\lambda^*. \tag{8}$$

The second piece is the marginal benefit of reducing the expected long-run contribution rate, $E(c^*)$, by raising f^*. From (3*), we have:

$$-V_2 dE(c^*)/df^* = V_2[(c^p - c^n) + (r - d)\lambda^*]. \tag{9}$$

Note that the magnitude of this benefit depends on how aggressive the asset allocation is, $(r - d)$.

Taking these first two pieces of (7) together and using (2*) for λ^*, we find the net benefit (ignoring the cost of risk for the moment) of raising the target funded ratio is positive if $V_2(r - g) > V_1$. This condition is just a simplified version of the usual intergenerational tradeoff. Suppose, to take the simplest example, the accumulation of additional assets is immediate. The subsequent reduction in the contribution rate, as a percent of payroll, represents a perpetuity that grows at rate g. Then, if social (or policymakers') cost is simply the present value of current and future contributions, discounted at rate δ, we find $V_1 = 1$ and $V_2 = 1/(\delta - g)$. Consequently, the net benefit is positive if and only if $(r - g)/(\delta - g) > 1$, i.e., $\delta < r$, a standard result.[40]

Finally, the third term in (7) is the marginal cost of the increased risk from relying on the additional income generated by asset accumulation. Using (5) for $\sigma(c^*)$, we have:

$$-V_3 d\sigma(c^*)/df^* = -V_3 S(r)\lambda^*. \tag{10}$$

Pulling these three pieces together, we have:

$$-dV[(a^* - a_0), E(c^*), \sigma(c^*)]/df^* = -V_1\lambda^* + V_2[(c^p - c^n) + (r - d)\lambda^*] - V_3 S(r)\lambda^*. \tag{7'}$$

[37]Note that the low-risk portfolio, $r = 4$ percent, is not risk-free. That is because the fixed rate bond portfolio ($r = 2.75\%$) is not risk-free ($\sigma(r) = 3.75\%$) and, to reach the target return of 4 percent, one must add an equity component. Thus, although our simulations illustrate the analysis below, they do not accord exactly with the risk-free assumption.

[38]The change in year-30 risk per unit change in median contribution rises from 0.9 to 2.6 as r varies from 4 to 7 percent.

[39]We also use CalSTRS' (2023*b*) capital market assumptions (pp. 10–11) to generate similar simulations. See Appendix Figures B and C. Note that these assumptions generate $\sigma(r = 7\%) = 10.5\%$, which differs from CalSTRS' (2023*a*) valuation report assumption (p. 51) that $\sigma(r = 7\%) = 11.6\%$, as used in Appendix Figure A.

[40]Adding convexity to the annual disutility of contributions is straightforward.

As discussed earlier, the risk premium, $(r-d)$, simultaneously conveys the benefit and the cost of risky investment. These are reflected in the second and third terms above, respectively, since $S(r) \equiv s(\sigma(r))$ and $\sigma(r)$ is a function of $(r-d)$.

In the second term above, it is important to reiterate that c^n is discounted at d, not r. Thus, the gap between c^p and c^n is much narrower than under actuarial accounting. The benefit from asset accumulation is smaller in that regard but enhanced by the risk premium. We can further clarify how the costs and benefits of risky investment net out here by regrouping terms in (7′) and substituting from (6), the optimality condition for the target return, $V_2 = V_3S'(r)$:

$$\begin{aligned} &-dV[(a^* - a_0), E(c^*), \sigma(c^*)]/df^* \\ &\quad = -V_1\lambda^* + V_2(c^p - c^n) + [V_2(r-d) - V_3S(r)]\lambda^* \\ &\quad = -V_1\lambda^* + V_2(c^p - c^n) + V_2[(r-d) - S(r)/S'(r)]\lambda^*. \end{aligned} \tag{7″}$$

Our analysis assumes d is risk-free, so $S(r=d) = s(\sigma(r=d)) = 0$. Thus, $S'(r) > S(r)/(r-d)$, if $S(r)$ is convex, as alluded to above (and discussed further below). This means the bracketed term is positive: the marginal benefit from reducing future contributions by the extra income from additional risky assets outweighs the cost of the extra risk.[41]

Heuristically, we may think of this convexity result as follows. If policymakers assess the convexity of $S(r)$, then they willingly accept the relatively high marginal impact on risk, $S'(r)$, of their asset allocation decision. Based on (6), this decision implies a low aversion to future risk, V_3, relative to the value they place on the benefits of expected return, V_2. Carrying this inference from (6) over to the asset accumulation decision, (7″), the bracketed term on the second line shows how the relatively low degee of risk aversion implies a greater willingness to accumulate assets for the net benefit of the risk premium. As we will show below, this can play a significant role in the basis for pre-funding vs. pay-as-you-go.

6.3 Convexity of long-run risk-return and the bases for pre-funding

Convexity of the composite function $S(r) = s(\sigma(r))$ depends on the convexity of both $s(\sigma)$ and $\sigma(r)$, where $s(\sigma)$ relates the risk of the long-run contribution rate to the risk of the annual return and $\sigma(r)$ gives the risk of the annual return as a function of its expected value. Under basic portfolio theory, where assets are simply a mix of the risk-free asset and the market portfolio, $\sigma(r)$ is linear. Under the more complex capital market assumptions we use for our simulations – based on the insights of Merton (1973) – $\sigma(r)$ embeds some convexity in the annual risk. However, that convexity appears to be swamped by the long-run convexity of $s(\sigma)$.

As Figure 10 (depicting $\sigma(c_t) = S(r; t)a^*$, by equation (5)) shows $S(r) \equiv s(\sigma(r); t, \beta, \gamma)$ is notably convex for $t = 30$. This contrasts with the near-linear appearance of $\sigma(r)$, superimposed on the same graph.[42] Thus, we infer that the lion's share of $S(r)$'s convexity is due to that of $s(\sigma)$, the non-linear impact on the long-run contribution risk of the risk in annual returns, rather than the non-linearity of the annual risk–return relationship, $\sigma(r)$. This convexity of $S(r)$ is apparently due to compounding of risk over time, as illustrated by the comparison between the curvature of $S(r)$ for $t = 30$ and $t = 15$.

To formalize the convexity of $\sigma(c_t) = s(\sigma(r); t)a^* \equiv S(r; t)a^*$ would be intractable, but perhaps the following approximation provides some insight. Note first, from (1*) that $\boldsymbol{c}^* = c^p - (r-g)a^*$. We conjecture that the vast majority of the deviation of c_t from c^* is due to that of a_t from a^*, where the latter

[41]This result would be mitigated by consideration of the impact of higher r on short-term contributions, assumed away here by the undiscounted formulation of our first term in V. Expanding (6) to incorporate the favorable impact of higher r on short-term (as well as long-term) contributions would imply $V_3S'(r) > V_2$ (instead of holding with equality), thereby representing a greater willingness to tolerate risk of future contributions. This would reduce the bracketed term in (7″). That said, the convexity of the long-run risk–return relationship would still be an important factor in the basis for pre-funding, examined below, if not in quite so stark a form.

[42]The results are similar for our simulations of CalSTRS; see Appendix Figure D.

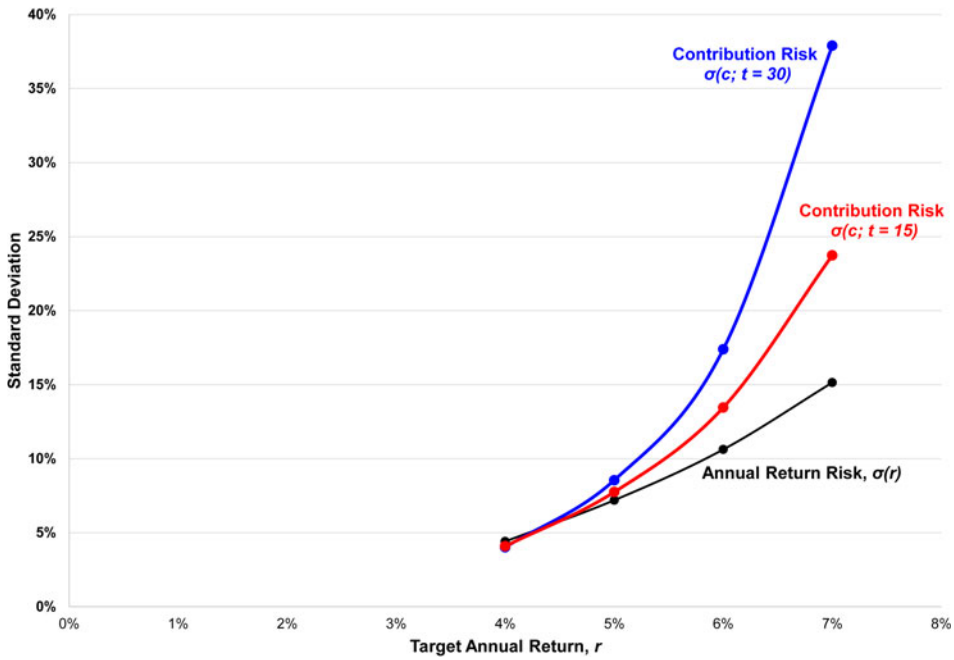

Figure 10. Convexity of contribution risk and annual return.

spread is depicted in Figure 7. If so, then $\sigma(c_t) \approx (r - g)\sigma(a_t)$. Taking R as distributed log-normal, then the basic result, $\sigma(a_t) = \sigma(r)a^*\sqrt{t}$ implies

$$\sigma(c_t) \approx (r-g)\sigma(r)a^*\sqrt{t}, \text{ so } d^2\sigma(c_t)/dr^2 = [(r-g)\sigma''(r) + 2\sigma'(r)]a^*\sqrt{t}.$$

Thus, even if $\sigma(r)$ is linear ($\sigma''(r) = 0$), we have convexity, especially as t gets large.

Returning to (7″), at the joint optimum the marginal cost of accumulating more assets (the first term in (7″)) is balanced by two marginal benefits (the second and third terms in (7″)). The second term represents the reduction in steady-state contributions from pre-funding at $r = d$ instead of pay-go, $(c^p - c^n)$. The third term represents the net benefit from the additional income from risky assets, $[(r - d) - S(r)/S'(r)]\lambda^*$, where the latter term is governed by the convexity of $S(r)$.

We can gain some purchase on the potential relative magnitude of these two benefits by further analysis of these two terms. Substituting from (2*) for λ^*, and rearranging, we find:

$$\begin{aligned}(c^p - c^n) &+ [(r-d) - S(r)/S'(r)]\lambda^* \\ &= (c^p - c^n)\{1 + [1 - (S(r)/(r-d))/S'(r)](r-d)/(d-g)\} \\ &= (c^p - c^n)\{1 + [1 - \mathit{arc\ slope/tangent}](r-d)/(d-g)\}.\end{aligned} \tag{11}$$

Here, *arc slope* is the slope of the arc of $S(r)$ over the interval (d, r) and *tangent* is the slope at r (see Figure 10). By convexity, the slope of the tangent exceeds that of the arc to the left of r, so (1 – *arc slope/tangent)* is a positive measure of the convexity of $S(r)$ at r, over the interval (d, r).

To illustrate magnitudes, let $d = 0.04$ and $g = 0.03$ and consider the *arc slope* and *tangent* from Figure 10. For $r = 0.07$, we estimate *arc slope/tangent* at about 11.3/25.5 = 0.44, which implies the net benefit from risky investments is about (1 – 0.44)(0.07 – 0.04)/(0.04 – 0.03) = 1.67 times the straight pre-funding benefit, $(c^p - c^n)$. At $r = 0.06$, we estimate that multiple drops to about 1.09 times $(c^p - c^n)$, i.e., still doubling the straight pre-funding benefit. At $r = 0.05$, the multiple drops to

https://doi.org/10.1017/S1474747223000173 Published online by Cambridge University Press

about 0.48. These estimates are by no means dispositive; they merely illustrate the potential relative size of the two marginal benefits from accumulating more assets.[43]

This exercise leads us to re-examine the basis for pre-funding in light of the dual rates. By properly discounting liabilities at d instead of r, the normal cost rate is dramatically elevated, and $(c^p - c^n)$ is greatly diminished; it rests on the gap between d and g instead of the much wider gap between r and g.[44] It is the latter gap that is often taken as the basis for pre-funding. That is also what underlies the actuarial goal of 'full' funding contributions dropping to the normal cost rate (discounted at r) instead of the much higher pay-go rate. Our analysis suggests that the much narrower gap between c^p and c^n (discounted at d) might well be a less powerful motive for pre-funding than the exploitation of risky investments.[45] We show how the relative weight of these two bases for pre-funding depends on policy-makers' choice of target return.

Finally, we should emphasize that the object of our analysis, the optimal funded ratio f^*, is applied to a much higher liability ratio, λ^*, discounted at d. Thus, even if the above analysis is taken to suggest less-than-full funding, it does not resolve the issue of whether current funding targets (100% of reported liabilities or 60% of true liabilities) are too high or too low. A low target funded rate for true liabilities may well exceed current funding targets.

7. Conclusions and future research

Standard actuarial practice pursues intergenerational equity and sustainability by employing funding rules that seek to ensure each generation pays for the services it receives. These rules operate through the concepts of normal cost and amortization, which, together, aim to fully fund benefits for each cohort of workers and taxpayers. Normal cost is meant to pre-fund the full cost of benefits earned by a cohort of employees over their careers, while amortization is meant to close funding gaps that result from payment shortfalls and unrealized assumptions.

In practice, these rules have failed to achieve intergenerational equity or sustainability. The true market cost of earned benefits and of asset risk have been understated, leading to the accumulation of large pension debt and steeply rising contributions to amortize that debt. These payments are crowding out spending in other areas like infrastructure and education, as current generations pay for past benefits and new debt accrues (McGee, 2016; Biggs *et al.*, 2022; Costrell and McGee, 2022). Going forward, neither intergenerational equity nor sustainability is embedded in a deliberative policy choice framework that adequately considers the risks involved for future generations of public workers and taxpayers.

The shortcomings of the actuarial approach lie not only in practice, but in theory, as discounting by the expected return fails to convey the cost of risk. The reformulation we propose goes back to fundamentals, properly discounting liabilities and incorporating the portfolio's risk premium into contribution policy to simultaneously represent the benefits and costs of risky investment.

Our approach is organized around steady-state analysis, to operationalize the concept of sustainability. The result of our analysis can be thought of as replacing both pieces of actuarial contributions: normal cost and amortization, in a way that explicitly recognizes long-run contribution risk. These pieces are embedded in a framework where the properly discounted target funded ratio is a choice that depends on both the risk and the return from investment in risky assets, and which may well differ from full funding at the risk-free rate.

Specifically, normal cost is effectively replaced by the steady-state expected contribution rate, derived from the laws of motion for assets and liabilities and the target funded ratio. As a result, the steady-state contribution rate: (i) broadens the normal cost piece (wrongly discounted) to a blend of normal cost (properly discounted) and pay-go, weighted by the target funded ratio; and

[43]Our estimates for CalSTRS, using Appendix Figure D, are much higher, since CalSTRS assumes $g = 0.035$.

[44]Lenney *et al.* (2021, Table A9) rediscount normal cost higher than the pay-go rate, apparently assuming $d < g$.

[45]In the limit, as $d \to g$, $c^n \to c^p$, so the second term in (7″), $V_2(c^p - c^n)$, vanishes. By L'Hôpital's rule, as $d \to g$, $\lambda^* = (c^p - c^n)/(d - g) \to -c^{n\prime}(d)$, so it does not vanish from the first term of (7″). Thus, in this limiting case, as $d \to g$, there will be no interior solution for f^* in the absence of the third term.

(ii) folds in the benefit of excess returns (r–d) in exchange for the risk borne by the sponsoring government and future taxpayers.

Instead of an amortization schedule, our approach specifies an adjustment process to the target contribution rate and asset ratio. We have shown how to set the adjustment parameters, starting with the deterministic case. However, as we also show, in a stochastic world, the risk widens over time, even as the contribution rate approaches a steady state in expected value. We show how the adjustment parameters might be modified from the deterministic case to better manage risk. That choice of parameters must navigate the tradeoff between lower contribution risk and speed of adjustment toward the targets. We also illustrate how such a policy can be tailored, outside of steady state, to moving targets, with the example of CalSTRS.

Finally, we sketch out a stylized optimization framework for choosing a target funded ratio, f^*, as applied to true liabilities (i.e., discounted risk-free at d), and choosing a target expected return on assets, r, with its associated risk. This simple framework balances intergenerational equity, the quest for returns, and investment risk based, ideally, on the policymakers' assessment of public preferences, or, more likely, their own incentive structure.[46] Incorporating our steady-state results into this simple framework sheds new insight on the costs and benefits of asset accumulation. The standard rationale for pre-funding vs. pay-go ($c^n < c^p$), is attenuated by properly discounting normal costs but is augmented by the net benefit of the excess returns from risky investments.[47]

The net benefit of those excess returns depends on the combination of policymakers' tolerance for risk to be imposed on future taxpayers, on the one hand, and, on the other hand, the convexity of long-run risk with respect to the expected return. The relevant measure of convexity varies with the optimal target return and asset risk. We illustrate the menu of such risk–return profiles to show that this second rationale for pre-funding – the net benefit of excess returns – can be quite substantial and even outweigh the traditional rationale for pre-funding.

To be sure, the optimization framework we present is nice, but not necessarily descriptive of current practice. Indeed, under actuarial practice, the asset target decision (a^*) does not appear to be based on any optimization framework, implicit or explicit. Rather, a^* is simply set at 100 percent of the actuarial calculation of liabilities, corresponding to about 60 percent of true liabilities. However, given the latitude plans seem to exercise in choosing the discount rate and other assumptions, one might interpret these decisions as, in effect, reverse engineering asset and contribution targets to satisfy policymaker preferences over some heuristic objective function.

To speculate along these lines, we could characterize common critiques of pension funding policy as: (i) understating V_2 relative to V_1 – excess time preference; (ii) under-estimating the social cost of risk borne by future taxpayers, V_3 relative to V_2 – insufficient risk aversion; or (iii) underestimating the amount of risk, $\sigma(c^*)$, perhaps due to misplaced confidence in time diversification, excessive self-confidence in investment acumen based on past good luck (Andonov and Rauh, 2022), and/or the distorted incentives from US public pension accounting rules (Andonov *et al.*, 2017).

Would a proper evaluation of the true social costs and benefits lead us to raise or reduce the target asset accumulation? Equation (7′) shows that excessive time preference reduces the target (the second term) and insufficient risk aversion raises it (the third term). This leads us to an inconclusive assessment of whether the target funded ratio is too high or too low.

What does our analysis say about how the steady-state expected contribution rate compares with the actuarial rate, namely the wrongly discounted normal cost rate? The question ultimately comes down to the accuracy of the expected return. As we saw in discussion of (1*), if the asset target

[46]Unlike much of the literature reviewed above, we do not start from the irrelevance theorems of Ricardian equivalence and Modigliani–Miller, whereby taxpayers offset public decisions on debt and risk by their own spending and portfolio decisions. Our stylized framework might be interpreted as a reduced form incorporating the various frictions (e.g., distortionary taxes) that undermine the irrelevance theorems, as carefully examined in prior literature, or, alternatively, as representing policy-makers' belief (rightly or wrongly) that such offsetting private behavior is not important.

[47]Although Bohn's (2011) main result is zero pension funding, he alludes to the possibility that if voters believe active fund management can beat the market, this may help explain the practice of pre-funding.

$a^* = 7$ is about right, then for $r = 7$ percent and $g = 3$ percent, the steady-state contribution rate of about 10 percent (depicted in Figures 5 and 7) is in the same ballpark as the reported normal cost rate of about 13 percent (depicted in Figure 1). If, however, $r = 6$ percent, then $c^* = 17$ percent, exceeding the reported normal cost rate, and at $r = 5$ percent, $c^* = 24$ percent, much higher yet.

In any case, the prospect of any steady-state relief from the current contribution rate of 27 percent is small consolation, given the transition paths to $a^* = 7$ depicted in these figures, even for $r = 7$ percent. These paths exhibit substantial short-term hikes, dramatically widening risk, and quite possibly never declining at all. Choosing a less aggressive portfolio and judicious adjustment parameters reduces the risk, but also reduces the prospect of substantial long-run decline in contributions.

In short, there are no perfect choices, but there may be better and worse ones. Our hope is that the analysis provided here helps elucidate the tradeoffs involved in pursuit of pension funding sustainability and intergenerational equity, and how these tradeoffs might inform the approach to contribution policy we propose. It is an approach that integrates proper liability discounting with clearer consideration of the benefits and costs of risks in setting asset and contribution targets, while pointing the way to a deliberative process of adjustment toward those targets and managing the risks better than current policies.

Our analysis and proposed funding approach lead us to several potentially fruitful areas for future work. While we use CalSTRS as a real-world example in this paper, applying our approach to historical and/or projected data for a broader set of public plans and comparing our approach to actual/expected performance will provide a more comprehensive picture of how well it might perform. Relatedly, this will afford the opportunity to delve further into the impact of non-steady-state dynamics due to changing dependency ratios, benefit levels, payroll growth, etc., which remains a significant area of uncertainty in the current analysis.[48] We hope to do more to develop rules for choosing (and potentially modifying) adjustment parameters in a dynamic environment. Finally, we hope to develop simple quantitative measures of the tradeoffs underlying our approach that could practicably inform policymakers' choice of target funded ratio as applied to true liabilities.

Supplementary material. The supplementary material for this article (Appendix Figures A-D) can be found at https://doi.org/10.1017/S1474747223000173.

Acknowledgements. An earlier version of this paper was presented at the 47th Annual Conference of the Association for Education Finance and Policy (AEFP), Denver CO, March 17, 2022, where we received helpful comments from Cory Koedel, Mike Podgursky, and Dan Goldhaber. We presented an antecedent paper at the Brookings Municipal Finance Conference, July 14, 2020, where we also benefited from participants' comments. Finally, we very much appreciate the comments and encouragement of Josh Rauh, as well as the external reviewers. The viewpoints expressed, let alone any errors, are those of the authors alone.

References

Andonov A and Rauh JD (2022) The return expectations of public pension funds. *The Review of Financial Studies* **35**, 3777–3822.

Andonov A, Bauer RMMJ and Cremers KJM (2017) Pension fund asset allocation and liability discount rates. *The Review of Financial Studies* **30**, 2555–2595.

Biggs A (2011) An options pricing method for calculating the market price of public sector pension liabilities. *Public Budgeting & Finance* **31**, 94–118.

Biggs A, Boyd D, McGee JB and Monahan AB (2022) Addressing and avoiding severe fiscal stress in public pension plans, Urban Institute, Washington, D.C.

Bohn H (2011) Should public retirement plans be fully funded? *Journal of Pension Economics and Finance* **10**, 195–219.

Brown J and Wilcox D (2009) Discounting state and local pension liabilities. *American Economic Review* **99**, 538–542.

Brown JR, Clark R and Rauh J (2011) The economics of state and local pensions. *Journal of Pension Economics and Finance* **10**, 161–172.

California State Teachers' Retirement System (2023*a*) Defined benefit program of the California State Teachers' Retirement System, June 30, 2022 Actuarial Valuation.

[48]Other important risks include inflation, longevity, and work-life assumptions. One might also explore how to pair our proposed approach with other levers for managing risk (e.g., COLA adjustments).

California State Teachers' Retirement System (2023*b*) Investment policy and management plan, CalSTRS Investment Branch, January 2023.
Callan Institute (2020) 2020 Capital market assumptions, webinar slide deck, January 2020.
Costrell RM (2018*a*) The 80 percent pension funding target, high assumed returns, and generational inequity. *Contemporary Economic Policy* **36**, 493–504.
Costrell RM (2018*b*) Accounting for the rise in unfunded public pension liabilities: faulty counterfactuals and the allure of simple gain/loss summations. *Journal of Pension Economics and Finance* **17**, 23–45.
Costrell RM and McGee J (2020) Sins of the past, present, and future: alternative pension funding policies, EdWorkingPaper: 20–272, Annenberg Institute at Brown University: August 7, 2020; conference paper for Brookings Municipal Finance Conference, July 13–14, 2020.
Costrell RM and McGee J (2022) Recent research on teacher pension funding, benefits, and policy debates. In *Recent Advancements in Education Finance and Policy*, Downes T and Killeen K (eds), Charlotte, NC: Information Age Publishing, pp. 389–422.
D'Arcy SP, Dulebohn JH and Oh P (1999) Optimal funding of state employee pension systems. *The Journal of Risk and Insurance* **66**, 345–380.
Glaeser EL and Ponzetto GAM (2014) Shrouded costs of government: the political economy of state and local public pensions. *Journal of Public Economics* **116**, 89–105.
Lenney J, Lutz B and Sheiner L (2019*a*) The sustainability of state and local government pensions: a public finance approach, Brookings Municipal Finance Conference, July 14, 2019.
Lenney J, Lutz B and Sheiner L (2019*b*) The sustainability of state and local government pensions: a public finance approach, University of Chicago Policy Forum: the pension crisis, November 7, 2019.
Lenney J, Lutz B, Schule F and Sheiner L (2021) The sustainability of state and local pensions: a public finance approach. *Brookings Papers on Economic Activity*, Spring 2021, 1–48. Online Appendix.
Love DA, Smith PA and Wilcox DW (2011) The effect of regulation on optimal corporate pension risk. *Journal of Financial Economics* **101**, 18–35.
Lucas D (2021) The sustainability of state and local pensions: a public finance approach, comment. *Brookings Papers on Economic Activity*, Spring 2021, 49–57.
Lucas DJ and Zeldes SP (2009) How should public pension plans invest? *American Economic Review* **99**, 527–532.
McGee JB (2016) *Feeling the squeeze: pension costs are crowding out education spending*. Report 22, Manhattan Institute.
Merton RC (1973) An intertemporal capital asset pricing model. *Econometrica* **41**, 867–887.
Munnell AH, Aubry J-P and Quinby L (2011) Public pension funding in practice. *Journal of Pension Economics and Finance* **10**, 247–268.
Neusser K (2021) Difference equations for economists, online March 15, 2021, http://www.neusser.ch/downloads/DifferenceEquations.pdf.
Novy-Marx R and Rauh JD (2009) The liabilities and risks of state-sponsored pension plans. *Journal of Economic Perspectives* **23**, 191–210.
Novy-Marx R and Rauh JD (2011) Public pension promises: how big are they and what are they worth? *Journal of Finance*, **66**, 1211–1249.
Pennacchi G and Rastad M (2011) Portfolio allocation for public pension funds. *Journal of Pension Economics and Finance* **10**, 221–245.
Rauh J (2021) The sustainability of state and local pensions: a public finance approach, comment. *Brookings Papers on Economic Activity*, Spring 2021, 57–61.
Samuelson P (1958) An exact consumption loan model of interest with or without the social contrivance of money. *Journal of Political Economy* **66**, 467–482.
Sharpe WF (1976) Corporate pension funding policy. *Journal of Financial Economics* **3**, 183–193.
Treynor JL (1977) The principles of corporate pension finance. *The Journal of Finance* **32**, 627–638.
Turnovsky SJ (1974) The stability properties of optimal economic policies. *American Economic Review* **64**, 136–148.
Van Binsbergen JH and Brandt MW (2016) Optimal asset allocation in asset liability management. In Veronesi P (ed.), *Chapter 8 in Handbook of Fixed-Income Securities*, Hoboken, NJ: John Wiley & Sons, pp. 147–167.

Appendix

Convergence conditions

We can usefully express the system (1″) and (4) in matrix form:

$$\begin{bmatrix} a \\ c \end{bmatrix}_{t+1} = \begin{bmatrix} (R/G) & (1/G) \\ -\gamma & (1-\beta) \end{bmatrix} \begin{bmatrix} a \\ c \end{bmatrix}_t + \begin{bmatrix} (-c^P/G) \\ (\gamma a^* + \beta c^*) \end{bmatrix}.$$

Denote the transition matrix above by A. The asymptotic stability condition (see Neusser (2021), equation (3.18), p. 84) is: $|tr\ (A)| < 1 + det(A) < 2$. In the present case, this implies

(i) $\gamma > \beta(R - G) \equiv \gamma_{min} > 0$, and
(ii) $\gamma < G - R(1 - \beta) \equiv \gamma_{max}$.

The condition for asymptotic oscillation is $[tr\ (A)]^2 < 4 \cdot det(A)$, or, in the present case:

(iii) $\gamma > G[(R/G) - (1 - \beta)]^2/4 \equiv \gamma_{m/o}$.

Cite this article: Costrell RM, McGee JB (2023). Toward an economic reformulation of public pension funding. *Journal of Pension Economics and Finance* 1–29. https://doi.org/10.1017/S1474747223000173