Tail Risk in Pension Funds: an Analysis using ARCH Models and Bilinear Processes

Iqbal Owadally

This paper is published in *Review of Quantitative Finance and Accounting*, 43, 2, 301-331

Abstract Pension funding rules and practice contain implicit smoothing and counter-cyclical mechanisms. We set up a stylized model to investigate whether this may give rise to tail risk, in the form of large but rare losses, when pension liabilities are imperfectly but optimally hedged by pension fund assets. We find that pension losses follow a nonlinear dynamic process, and we derive a complete description of the stochastic properties of this process using Markov chain and bilinear stochastic process theory. The resulting pension dynamics resembles that of a modified ARCH model, which suggests that bursts in volatility may occur, and tail risk may be present. Simulations confirm that pension losses exhibit skewness, leptokurtosis and heavy tails, specially when cash flow smoothing is pronounced. Regulators and investors should be aware of the total amount of smoothing in pension funds as this may contribute to extreme losses, which may adversely affect the security of employee benefits as well as the valuation of firms with corporate pension plans.

Keywords Pensions · Risk · Heavy-tailed distribution · LARCH process

JEL G23 · G28 · G32

I. Owadally

Cass Business School, City University London, 106 Bunhill Row, London EC1Y 8TZ, United Kingdom

Tel.: +44-20-70408478

Fax: +44-20-70408572

E-mail: m.i.owadally@city.ac.uk

1 Introduction

Risk management in pension plans, as in other financial institutions, requires a careful assessment of tail risk. Tail risk is the occurrence of extreme losses with a higher probability than may be normally expected. In the 2008 financial crisis, a large number of pension funds around the world reported severe difficulties with falling asset values and increasing liabilities.[1] Large losses occurred, and higher contributions were then required from pension plan sponsors. Regulators and legislators worldwide responded by easing the funding requirements on sponsors. The purpose of this paper is to investigate the presence of tail risk in corporate pension plans resulting from pension funding practice and the application of funding rules.

We consider 'defined benefit' pension funds. In 2011, such funds had assets totaling $6.6 trillion in the U.S. and $14.8 trillion globally.[2] In this paper, we set up a stylized model to investigate the mechanism by which extreme losses might arise. Our starting point is that pension liabilities are not completely marketable and therefore not fully hedgeable. Losses arise from imperfect, albeit optimal, hedging. However, pension funding rules and practice seek to smooth these cash flows inter-temporally and counter-cyclically. This delays the emergence of possible funding problems, which may then accumulate with adverse dynamic effects.

We show in our stylized model that pension losses follow a nonlinear dynamic process, which we analyze using Markov chain theory and the theory of bilinear stochastic processes. The resulting pension dynamics resembles that of a modified GARCH process. We therefore posit that the pension

[1] Global pension funds experienced losses of about 25% of their asset value, as a percentage of pension liability, in 2008 (Towers Watson 2011).

[2] Data from Towers Watson (2011). In 'defined benefit' pension plans, employees receive pensions at retirement that are a function of their service and salary, with the employer taking on the investment and longevity risk associated with these benefits. This is in contrast with 'defined contribution' plans where employees take on all the risks. Contributions from firms and employees to corporate defined benefit funds amounted to about 0.5% of U.S. GDP in 2009 (OECD data reported in Yermo and Severinson 2010). Our analysis is focused on corporate pension plans, but may be extended to plans covering government employees. In the U.S. in 2008, state government defined benefit plans had assets of $2.3 trillion (Novy-Marx and Rauh 2011).

fund loss distribution exhibits leptokurtosis as well as similar extremal behavior to GARCH, and we investigate this further through stochastic simulations using data on U.S. pension fund returns. Our key finding is that tail risk—transpiring in the form of a leptokurtic distribution of pension losses as well as Pareto-like slow decay in the tail of this distribution—is present when pension cash flows are deferred and smoothed.

An important policy implication is that, as with pension accounting rules, pension funding rules should be carefully designed to limit the extent of smoothing that is permissible, taking into account normal corporate practice regarding the management of cash flows, operating leverage and tax liability. In particular, pension fund managers and regulators should be aware that smoothed funding methods must be monitored as they may generate risk that may not be observable in normal financial circumstances, but may manifest itself in the form of rare but large losses. Our research also informs the policy debate in the European Union concerning the application of the Solvency II capital adequacy standard to pension funds.

This article is organized along the following lines. The background and motivation for our research, as well as related papers in the literature, are discussed in section 2. A pension fund model is developed in section 3 and its stochastic properties (strict and weak stationarity, existence of moments) are investigated in section 4. An analogy with ARCH-type models is also made, which leads us to the presumption that tail risk may exist in pension losses. Simulations are carried out in section 5 demonstrating tail risk, and we conclude the paper in section 6 with some policy recommendations.

2 Background and Related Literature

2.1 Pension accounting

This research coincides with recent developments in the institutional setting of corporate pension plans, in terms of accounting, funding, as well as regulatory response to the 2008 financial crisis. This paper is also related to a number of concomitant strands of the research literature.

The first key development is in pension accounting. There is a significant literature on the value relevance of pension accounting information, and also on smoothing in pension expense (as opposed to pension funding). Jiang (2011) shows that the corridor amortization mechanism used in income statements under Financial Accounting Standard 158—hereafter FAS 158, see FASB (2006)—is ineffective and introduces biases leading to overestimation of plan sponsors' earnings in the long term. FAS 158 took effect in late 2006. It amends and updates previous accounting rules, notably Financial Accounting Standard 87—hereafter FAS 87, see FASB (1985)—which incorporated smoothing and deferred recognition of pension gains and losses on both the balance sheet and income statement. Mitra and Hossain (2009) examine the implication for valuation of transition adjustments in accounts when FAS 158 was initiated and full recognition of gains and losses was required, rather than disclosure in footnotes. They find that investors react, and stock prices fall, when the adjustment is significant, thereby showing that markets do process pension accounting information on the balance sheet. Both Jiang (2011) and Mitra and Hossain (2009) conclude that the elimination of FAS 87-style smoothing from the FAS 158 balance sheet is beneficial to investors. On the other hand, Hann et al. (2007) find that non-smoothed fair-value accounting does not improve, and may impair, the value relevance of balance sheet and income statements, as compared to smoothed pension accounting under FAS 87. They suggest that the amortization mechanism of FAS 87 separates more persistent income from highly volatile gains and losses, thereby helping investors assess value.

Earlier accounting research, based on the older accounting standard FAS 87, shows that investors fail to price corporate pension liability efficiently (Franzoni and Marín 2006), with firms exploiting the latitude in assumption-setting and smoothing mechanisms embedded in FAS 87 to manage earnings (Picconi 2006; Bergstresser et al. 2006). In particular, Franzoni and Marín (2006) show that firms with significantly underfunded pension plans are overvalued by investors. They suggest that there are two reasons for this, one being related to accounting (the amortization corridor for pension expenses) and the other being related to funding. Specifically, mandatory pension contributions—required by law to enable pension plans to recover to full funding—are spread out over time and impact earnings and cash flows over several years.

2.2 Pension funding

The second key development for corporate pension plans in the U.S. has been in the area of pension funding. Pension funding contributions differ from pension accounting expense, in general. First, there are mandatory contributions imposed by regulators for minimum funding purposes. U.S. legislation changed in 2006 with the introduction of the Pension Protection Act 2006 (hereafter PPA 2006), which accelerates the remediation of deficits. Broadly, firms had 30 years to fund 90% of liability before PPA 2006, but now have 7 years to fund 100% of liability. Secondly, there are voluntary or discretionary contributions, made by pension plan sponsors, which are tax-deductible up to certain limits. PPA 2006 permits contributions that are deductible up to 150% of the pension liability, a larger amount than before, therefore encouraging firms to over-fund their pension plans. See Franzoni (2009), Campbell, Dhaliwal and Schwartz (2010, 2012) and Shivdasani and Stefanescu (2010) for details.

Research in the area of pension funding hinges around the interdependence between sponsors' financial policy and the investment and funding policies of their pension plans. Rauh (2006) finds a significant negative association between a firm's mandatory pension contributions and its capital investment. Required contributions to underfunded pension plans appear to force companies to divert cash away from investment. Pension funding rules and practice are therefore critical to the investors in the firm. Indeed, Franzoni (2009) demonstrates that increased mandatory pension contributions are followed by depressed stock returns. Campbell, Dhaliwal and Schwartz (2012) find that financially constrained firms facing an increase in mandatory contributions also face an increase in cost of capital. In an event study surrounding the date of the introduction of PPA 2006, when minimum funding rules were tightened, Campbell, Dhaliwal and Schwartz (2010) observe a fall in the stock market valuation of sponsors with large unfunded pension liabilities and capital expenditures.

PPA 2006 also increased the level of voluntary contributions qualifying for tax-deductibility. Accordingly, Campbell, Dhaliwal and Schwartz (2010) examine the data to see whether the equity valuation of corporate sponsors with higher marginal tax rates responds to the adoption of the new legislation. Their study reveals a positive effect on the stock returns of such firms. Funding rules matter to investors not just because of mandatory contributions but also because of their effect on discretionary contributions. Shivdasani and Stefanescu (2010) establish that the existence of a pension fund affects the capital structure of firms and their total amount of financial leverage, particularly because discretionary pension contributions are a source of tax savings. On the other hand, Petersen (1994) considers operating leverage, but also concludes that greater flexibility in pension contributions enables a firm to lower its operating leverage, which may add value in the presence of market imperfections.

2.3 Financial crisis and counter-cyclical measures

The third key event which has had an impact on pension funds has been the financial crisis in 2008 and the subsequent recession. Pension plan assets worldwide fell by about 25%, as a percentage of pension liability, in 2008 (Towers Watson 2011). Firms that were financially distressed in the subsequent recession could not meet funding rules. In the U.S., new legislation was passed in 2008 and 2010 to ease the statutory funding requirements introduced only in 2006 in the Pension Protection Act (Love, Smith and Wilcox 2011; Yermo and Severinson 2010). The Organization for Economic Cooperation and Development (OECD) also studied the response of regulators in Europe, Canada and Japan to the financial crisis: regulators relaxed the exigencies of minimum funding rules, allowing cash-constrained plan sponsors to bring their plans back to full funding over longer recovery periods, and partly suspending strict market valuations of pension liabilities in some countries (Yermo and Severinson 2010).

The OECD has since promulgated counter-cyclical measures and issued guidelines for OECD member countries which state that "Funding rules should aim to be counter-cyclical, providing incentives for the build-up of reserves against market downturns" (OECD 2007). Yermo and Severinson (2010) are particularly critical of funding rules which compel sponsors to ramp up contributions at a time when they may be cash-constrained. Such funding rules have systemic pro-cyclical effects in that they require numerous pension funds to sell assets at the same time to limit deficits, thereby exacerbating market falls. Yermo and Severinson (2010) also contend that poorly designed funding rules can have second-order macroeconomic effects in that large numbers of firms may curtail business investment to fund pension deficits, as documented by Rauh (2006), which slows down economic recovery and worsens business conditions for firms.

Counter-cyclical measures are also a key part of the Solvency II regime (EIOPA 2012) for insurance regulation in the European Union (E.U.). Proposals to apply this regime to pension funds are under consideration. Solvency II comprises explicit counter-cyclical measures. For example, a counter-cyclical or liquidity risk premium is used with the effect that capital requirements are temporarily reduced in times of financial crisis. An 'equity dampener' adjustment weakens the stress test on the equity portion of a portfolio when large market falls occur. Large market falls are defined relative to an averaged value of equity prices.

To some extent, counter-cyclicality is built into pension funding. For example, the U.S. Pension Protection Act 2006 permits both asset values and liability discount rates (based on the corporate bond yield curve or at least three different maturity segments) to be averaged over up to 24 months (OECD 2007). Likewise, Japanese pension liabilities are also discounted using an average of 10-year bond yields over the previous 5 years, while suitable swap rates are used in the Netherlands with smoothing permitted for contribution calculation (Yermo and Severinson 2010). In other countries, counter-cyclical funding may be made explicit through a separate contingency reserve. For example, Norwegian tax authorities allow payments into a "premium fund", over and above the payment of regular premiums, and Swiss pension funds are allowed "asset fluctuation reserves" and "employer contribution reserves" with typical amortization periods of 5–7 years (OECD 2007). Indeed, virtually all major economies afford a recovery or amortization period of between 3 to 7 years to their pension funds to make up funding deficiencies, these periods having been extended temporarily in the aftermath of the 2008 financial crisis. This is analyzed by Broeders and Chen (2010) who model pension fund insolvency using Parisian options and use stochastic simulations to investigate the optimal recovery period which regulators should allow in order to maximize employees' utility. They suggest recovery periods of between 1 and 5 years depending on pension plan features and other assumptions.

2.4 Other related literature

We briefly mention three other strands of research which are tangential to ours. First, in the subsequent analysis, we exploit some novel results from the econometrics literature on a variant of the GARCH model called LARCH (Linear-ARCH). This model is extensively studied by Giraitis et al. (2000, 2004) with further advances by Kristensen (2009) which link LARCH to stochastic bilinear processes (Pham 1993).

Second, pension insurance is an important aspect of pension funding but is not directly analyzed here. In the U.S., firms with a defined benefit pension plan pay a premium to the Pension Benefit Guarantee Corporation (PBGC) which then takes over pension promises (subject to a cap) if a plan sponsor becomes bankrupt. There is a flat premium per participant plus an additional premium dependent on the unfunded pension liability, but the premium is not risk-based. This creates a put option for plan sponsors which encourages a more aggressive pension fund investment policy. See, for example, Rauh (2009), Love, Smith and Wilcox (2011) and McGill et al (2004, p. 801) for further details.

Third, the issue of optimal pension funding and investment is not addressed in this paper. The literature on this subject (see e.g. Berkelaar and Kouwenberg 2003; Rudolf and Ziemba 2004; Owadally and Haberman 2004a; Josa-Fombellida and Rincón-Zapatero 2006) employs broad utility and objective functions on measures of fund levels and contributions, and does not focus on tail risk. This indicates that optimal funding solutions may need to be re-assessed in the light of our findings. Blake (2006b, p. 83–91) summarizes the issues affecting pension investment policy, including tax and pension insurance effects.

3 Pension Fund Model

3.1 Pension liabilities

A corporate pension fund consists of the assets held against the pension liabilities of a firm. First, we describe the pension liabilities. We assume a simple pension plan where all members enter at age E and leave the plan at retirement age R (where $E, R \in \mathbb{N}$ and $R > E$), whereupon they receive a retirement benefit which consists of a life annuity. (For example, $E = 27$ and $R = 67$.) No other benefit is payable.

We also assume that there is no demographic or longevity risk and that survival proceeds exactly according to a life table $\{\ell_x\}$, where ℓ_x is the number of individuals alive at age $x \in \{E, E+1, \dots\}$. The survival probability from age x to age y is therefore ℓ_y/ℓ_x (Dickson et al. 2009, p. 42). The age profile of the pension plan population is stable.

A suitable spot rate curve, which may be used to discount and value pension liabilities, is assumed to exist.[3] Denote the $(\tau - t)$-year continuously compounded spot rate at time t as $s(t, \tau)$. That is, $s(t, \tau)$ is the yield-to-maturity at time t of a zero-coupon bond maturing at time τ. The present value at time t of a deferred life annuity paying \$1 yearly in advance from retirement age R till death for an individual aged $x \leq R$ at time t is

$$\alpha(t, x) \;=\; \sum_{j=0}^{\infty} \exp\left(-(R - x + j)\, s(t, t + R - x + j)\right)\, \ell_{R+j}/\ell_x. \tag{1}$$

An annuity is purchased at retirement at which point an individual leaves the plan. The actual pension or annuity income received every year in retirement is usually a function of the number of years of service and the projected pensionable salary of the individual. The pensionable salary could be the final salary of the individual or his career-average salary. See for example Sundaresan and Zapatero (1997), McGill et al (2004, p. 235), Blake (2006b, p. 194), and Dickson et al. (2009, p. 306–312) for more details.

We assume that salary proceeds according to a deterministic age-based scale (Dickson et al. 2009, p. 292). Since the age profile of the plan membership is stable, payroll is constant. The pension benefit accrues to a plan member as he works and contributions are made to the plan. We may therefore define a benefit accrual function m_x where m_x is the benefit that has accrued to an individual aged x

[3] This is typically based on the yields on high-quality corporate bonds to reflect the credit risk present in pension benefits. Practical examples of such yield curves, updated monthly, are supplied by Citigroup (2010) and Mercer (2012). The U.S. Treasury Department also implements a high-quality corporate bond yield curve for use with the pension funding rules in the U.S. Pension Protection Act 2006 (Girola 2011).

(where $E \leq x \leq R$). From retirement age R until death, an individual receives a pension or annuity income of m_R every year. We assume that, at entry into the plan at age E, an employee is not immediately endowed with benefit rights, so that $m_E = 0$. The accrued liability for an individual aged x (where $x \leq R$) at time t may therefore be valued as $m_x \alpha(t, x)$.

While individuals are working and are members of the plan, their employer contributes to the pension fund, which accumulates and then pays out for the annuity when they retire. How much to contribute during the working lifetime is the subject of various actuarial cost methods. These provide a systematic way of accumulating funds to provide retirement benefits. Associated with these methods are different measures of the pension liability. A straightforward discussion of these measures is given by Novy-Marx and Rauh (2011) with more details given by McGill et al (2004, p. 617).

We assume that the pension liability is measured using the Projected Benefit Obligation (PBO). Under the PBO, m_x is a function of the projected salary at retirement of an individual aged x. For accounting purposes in the U.S., pension liability is reported using the PBO under FAS 158. For funding purposes, the U.S. Pension Protection Act 2006 requires underfunding to be measured using a metric for which the PBO is a very close proxy (Shivdasani and Stefanescu 2010; Campbell, Dhaliwal and Schwartz 2010).[4]

The pension liability is denoted by Y_t. When measured as the PBO, it is the sum of accrued liabilities for every individual in the plan. There are ℓ_x members aged x (where $E \leq x \leq R$), so

$$Y_t = \sum_{x=E}^{R} m_x \ell_x \alpha(t, x). \tag{2}$$

To simplify the liability dynamics, we shall assume that the spot rate curve $\{s(t, \tau), \tau = 1, 2, \dots\}$ is subject to small, parallel shifts only. Associated with the spot rate curve at time t is a single,

[4] The PBO is also the preferred measure of pension liability under several international accounting standards. An alternative measure of pension liability is the Accumulated Benefit Obligation (ABO) which is defined in FAS 87. It is also used by pension funding regulators in various countries because it is a termination measure, as compared with the PBO which is a broader, going-concern measure (Broeders and Chen 2010). Our subsequent analysis holds for the ABO rather than the PBO, as long as m_x is re-defined as a function of the current salary of an individual, rather than his projected salary at retirement.

term-independent discount rate δ_t at which we can discount the liability cash flows to obtain Y_t as in equation (2). We may regard δ_t as the yield-to-maturity or internal rate of return, at time t, on a notional dedicated bond portfolio whose cash flows match the pension liability cash flows as closely as possible.

We assume that the spot rate curve is subject to stochastic shocks such that it mean-reverts to a long-run curve $\{\widehat{s}(\tau), \tau = 1, 2, \dots\}$, corresponding to which we have a single, term-independent discount rate δ. We may therefore view δ as approximately equal to the average yield-to-maturity on the aforementioned notional bond portfolio. Further, define Y to be the liability calculated using $\{\widehat{s}(\tau)\}$, or equivalently using δ. Shocks to the term structure translate into shocks to δ_t (centered approximately around δ), and consequently into shocks to Y_t (centered approximately around Y).

Finally, we denote by D_Y the volatility or modified duration of the pension liability when discounted at rate δ. That is, $D_Y = -Y^{-1}\partial Y/\partial\delta$. We may then employ a first-order approximation centered around δ.

Assumption 1 The spot rate curve $\{s(t, \tau), \tau = 1, 2, \dots\}$ is subject to small, shape-preserving stochastic perturbations. The dynamics of the pension liability Y_t is approximately given by

$$Y_t = Y_{t-1} \exp\left(-D_Y(\delta_t - \delta_{t-1})\right). \tag{3}$$

3.2 Normal pension cash flows

Actual pension fund cash flows may differ from reported cash flows, in general. In particular, contributions may be different from pension accounting expense. We separate the pension cash flows into a normal cash flow, denoted by Z_t, and a supplementary cash flow. The latter is required to make up shortfalls in the fund or to reduce contributions in the event of surpluses, and is defined later in section 3.4. The normal cash flow Z_t is further decomposed into three cash flows. The first is the benefit paid out from the pension fund at time t. This is $Z_t^{(1)} = m_R \ell_R \alpha(t, R)$, since there are ℓ_R individuals at retirement age R.

The second cash flow is the required contribution under the actuarial cost method to be paid at time t. For a plan member aged x (where $E \leq x < R$), this is equal to the present value of the benefit that the member will accrue during the following year, that is $(m_{x+1} - m_x)\alpha(t,x)$. The total required contribution paid into the fund at time t is obtained by summing over all workers, noting again that there are ℓ_x workers aged x.

$$Z_t^{(2)} \; = \; \sum_{x=E}^{R-1} (m_{x+1} - m_x)\ell_x \alpha(t,x). \tag{4}$$

$Z_t^{(2)}$ is equivalent to the service cost in the calculation of pension expense in FAS 158 and its predecessor FAS 87.

The final component of normal cash flow Z_t is an additional interest charge on the liability, given by $Z_t^{(3)} = e^{-\delta}(\delta_t - \delta)Y_t$. The term $\delta_t Y_t$ corresponds to the interest cost in FAS 158 pension expense (and FAS 87), and is a finance charge on the liability (Blake 2006a, p. 62). This term is discounted here because of the modeling assumption that cash flows occur at the start of the year. The term δY_t is analogous to the expected return on plan assets in FAS 158 (and FAS 87). This is a negative item in that a positive expected return reduces pension cost (Grant et al. 2007; McGill et al 2004, p. 729). We consider an economic valuation of liabilities, without the FAS 158 objective of pension expense smoothing, and plan assets are assumed to closely hedge liabilities (see Assumption 2 below), so no long-term expected rate of return assumption is required here.

The normal cash outgo from the fund at time t is therefore $Z_t = Z_t^{(1)} - Z_t^{(2)} + Z_t^{(3)}$. Further cash flows may be required to pay off prior-service cost representing benefit enhancements attributable to plan members' past service or benefit rights obtained at the inception of the plan (Grant et al. 2007; McGill et al 2004, pp. 729–731). However, these are one-off, amortized payments and are ignored here.

We show in Appendix A that the normal cash outgo is approximately a constant percentage of the liability.

$$Z_t/Y_t \; \approx \; 1 - e^{-\delta}. \tag{5}$$

Equation (5) holds approximately because of the simplified liability dynamics of Assumption 1 enabling us to ignore, to first order, the convexity of the liability cash flow stream. Additionally, equation (5) follows from the earlier assumption of a stable age profile. The benefit payout and liability vary stochastically only with interest rates. In practice, they will also depend on other stochastic variables such as inflation, mortality, employee turnover etc.

Since $Z_{t-1} \approx Y_{t-1}(1 - e^{-\delta})$ from equation (5), we may re-write the liability dynamics in equation (3) as

$$Y_t \; = \; (Y_{t-1} - Z_{t-1}) \exp\left(r_t^L\right). \tag{6}$$

In the above, $r_t^L = \delta - D_Y(\delta_t - \delta_{t-1})$ and may be described as the 'liability return', this term being borrowed from the one-period models of Leibowitz and Henrikssan (1988) and Sharpe and Tint (1990), and their multi-period extension by Rudolf and Ziemba (2004).

3.3 Pension fund assets

We now turn to the pension fund assets. In the light of equation (5), it is natural to express the assets as a proportion of liabilities. The funding status (difference between liabilities and assets, scaled by liabilities) or funded ratio (ratio of plan assets to liabilities) are key quantities and are arguably more important than the PBO or fair value of plan assets by themselves: see for example Rauh (2009), Franzoni (2009) and Rudolf and Ziemba (2004).

We denote by F_t the market value of pension fund assets at time t as a percentage of liability Y_t. It is also convenient to define $\overline{F}_t = 1 - F_t$ as the unfunded liability at time t as a percentage of Y_t. The unfunded liability represents the deficit in the fund. We also define C_t to be a supplementary contribution, again as a percentage of liability Y_t, paid into the fund. The cash flow C_t is additional to the normal cash flow and is paid to reduce any deficit. In the case of a surplus, C_t may be negative so that the overall contribution is reduced.

The budget constraint on the pension fund is $F_tY_t = (F_{t-1}Y_{t-1} - Z_{t-1} + C_{t-1}Y_{t-1}) \exp\left(r_t^A\right)$, in dollar terms, where r_t^A is the stochastic return on pension fund assets in time interval $(t-1,\, t)$. With the help of equations (5) and (6) we may rewrite the budget constraint as

$$F_t \; = \; (F_{t-1} + C_{t-1} - Z) \exp\left(r_t^A - r_t^L + \delta\right). \tag{7}$$

Since r_t^A is the asset return and r_t^L is the liability return, the term $r_t^A - r_t^L$ may be described as a continuously compounded 'surplus return' (Leibowitz and Henrikssan 1988; Rudolf and Ziemba 2004). This brings us to a discussion of the relationship between the pension fund assets and pension liabilities.

In general, a perfect hedge for pension liabilities cannot be synthetized because these liability cash flows are not marketable. Some pension fund risk is therefore unhedgeable, at least partially. For example, credit risk, related to insolvency of the firm sponsoring the pension fund, may be hedged through credit default swaps, but these may be unavailable in the required volume or may be substituted for credit spread risk (Blake 2006b, pp. 269–271). The pension fund may also be restricted by regulation from taking excessive long, short or complex derivative positions in the securities issued by the firm to avoid agency conflicts and to encourage diversification. The very long duration of pension liabilities and scarcity of long-duration securities create duration mismatch; pension liabilities cannot be perfectly immunized despite the use of derivatives overlays (Adams and Smith 2009).

Background unhedgeable risks may also be present, chief among which is wage inflation risk (Berkelaar and Kouwenberg 2003; Blake 2006b, p. 268). Because retirees want to maintain their living standard in retirement, pensions are a function of salary before retirement. The market in inflation-indexed securities and inflation swaps may not be deep enough for all pension funds to hedge inflation risk, and basis risk will remain as local wages deviate from consumer prices. A more subtle background risk, which is difficult to hedge, is covenant risk. This pertains to the strength of the sponsor's commitment to the plan and the risk that the sponsor decides to terminate the plan or freeze new accrual.[5]

[5] The Pensions Act 2004 in the United Kingdom requires explicit consideration of the strength of the sponsor covenant for regulatory funding purposes (Yermo and Severinson 2010; OECD 2007, p. 205).

Returning to the surplus return $r_t^A - r_t^L$, we assume that a perfect hedge is not possible, because of market incompleteness as discussed above. The pension fund manager pursues an imperfect but optimal hedge instead, which entails a noisy replication or hedging error.[6]

Assumption 2 The stochastic surplus return $r_t^A - r_t^L$ satisfies $\mathbb{E}\left[\exp\left(r_t^A - r_t^L\right)\right] = 1$ and $\{r_t^A - r_t^L, t \in \mathbb{Z}\}$ is a sequence of independent and identically distributed (i.i.d.) random variables.

Note that we make no assumption about specific probability distributions in Assumption 2.

3.4 Supplementary pension cash flows

We discussed the normal pension fund cash flow (denoted by Z_t) in section 3.2. The total cash flow from the plan sponsor to the pension fund may differ from Z_t and we refer to this variation as a supplementary cash flow or supplmentary contribution (denoted by C_t). Supplementary cash flows may occur because of mandatory contributions imposed by regulators, under the U.S. Pension Protection Act 2006 (PPA 2006), for example. They may also occur because of voluntary contributions to take advantage of the tax-deductibility of pension contributions, or to maximize operational leverage. See the discussion in section 2.2 and references therein.

We propose to model the supplementary cash flow C_t in a generic way as follows. First, we define the pension fund loss as the unexpected change in pension fund assets relative to pension liabilities.

$$L_t \;=\; \mathbb{E}\left[F_t \mid \mathcal{F}_{t-1}\right] \;-\; F_t, \tag{8}$$

where $\mathcal{F}_t$ is information available at time t. (A gain is merely a negative loss.)

[6] In theory, this may be achieved using a quadratic (variance-minimizing) hedge or an entropy-minimizing hedge (Duffie and Richardson 1991; Rouge and El Karaoui 2000). In practice, this is consistent with asset-liability management strategies, including Liability-Driven Investment (LDI) (Blake 2006b, p. 260–274). These are growing in popularity with pension plans worldwide, specially in the UK and in the Netherlands (OECD 2007, p. 61). Love, Smith and Wilcox (2011) discuss the asset-liability mismatch prevailing in most U.S. corporate pension plans but report that 51% of them were instigating LDI programs in 2009.

Assumption 3 The supplementary cash flow C_t that is paid by the plan sponsor into the pension fund at time t is given by

$$C_t \;=\; \boldsymbol{\pi}'\,\mathbf{L}_t. \tag{9}$$

In the above, $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_p)'$ satisfying $\sum_{j=1}^{p} e^{-(j-1)\delta}\pi_j = 1$, and $\mathbf{L}_t = (L_t, L_{t-1}, \ldots, L_{t-p+1})'$, where $p \in \mathbb{N}$.

The motivation behind Assumption 3 is that supplementary cash flows tend to be retrospective in nature. In equation (9), C_t depends on present and past losses contained in the loss vector $\mathbf{L}_t$. For example, mandatory contributions are a function of the unfunded liability in the pension plan, and therefore on the history of losses. The choice of funding parameter vector $\boldsymbol{\pi}$ in equation (9) is flexible enough to capture the complex intertwining of funding and fiscal rules, and firms' incentives to contribute to their pension funds. Thus, an economic downturn may lead to pension losses in years which coincide with sponsor distress, leading to a reduced funding contribution (within minimum funding rules). In subsequent years, however, plan sponsors may increase contributions for tax or operational reasons when their balance sheets recover. See the discussion in section 2.2, and particularly Rauh (2006), Shivdasani and Stefanescu (2010) and Petersen (1994).

Assumption 3 also reflects the counter-cyclical measures and cash flow-smoothing behavior described in section 2.3. Smoothing in liability discount rates and asset values—as permitted in the U.S. under the Pension Protection Act 2006 (PPA 2006) and as detailed in Actuarial Standards of Practice Nos. 4, 27 and 44 (see ASB 2007a,b, 2009)—leads to a funding position that implies a smoothed average of market values. The condition in Assumption 9 that the funding parameters or filter weights $\{\pi_j\}$ satisfy $\sum_{j=1}^{p} e^{-(j-1)\delta}\pi_j = 1$ signifies that losses are fully paid off. Note that this is a generalized setting which encapsulates amortization over p years if π_j (for $j = 1, \ldots, p$) is set equal to the reciprocal of the present value of a p-year annuity-due. In this case, we envisage that the dimension p of the loss vector equals the maximum recovery period (for example, $p = 7$ years under PPA 2006). However, p might be greater than the typical amortization period if funding practice also comprises smoothed liability and asset values.

Proposition 1 *The loss L_t satisfies the stochastic recurrence relation*

$$L_t \;=\; \varepsilon_t \left(\boldsymbol{\beta}' \, \mathbf{L}_{t-1} - 1\right), \tag{10}$$

where $\varepsilon_t = \exp(r_t^A - r_t^L) - 1$ *and* $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$, *where* $\beta_j = e^{j\delta} - \sum_{k=1}^{j} e^{(j-k+1)\delta}\pi_k$ *for* $j \in [1, p]$.

The proof of Proposition 1 is given in Appendix B. Proposition 1 exhibits the dynamics of pension losses $\{L_t\}$ as a function of the sequence of asset-liability hedging or mismatch errors $\{\varepsilon_t\}$. If pension liabilities were liable to perfect hedging, then $\varepsilon_t = 0 \; \forall t$ and no pension loss would occur.

4 Bilinear Processes and ARCH Models

4.1 Stochastic properties

In order to make further progress with our analysis of the pension fund, we need to consider the probabilistic properties of $\{L_t\}$ in equation (10). In this section, we find the conditions under which $\{L_t, F_t, C_t\}$ is ergodic, strictly stationary, weakly stationary and has finite moments.

It is helpful to rewrite the stochastic process $\{L_t, t \in \mathbb{Z}\}$ in equation (10) as follows:

$$L_t \;=\; \varepsilon_t \, v_t, \tag{11a}$$

$$v_t \;=\; \sum_{j=1}^{p} \beta_j L_{t-j} \;-\; 1. \tag{11b}$$

We assume throughout that the initial conditions on equation (10), or equation set (11), consist of known finite values $L_l, \ldots, L_{l-p+1}$ for some time $l < 0$ in the distant past.

4.2 Weak stationarity and first two moments

Some elementary results about $\{L_t\}$ may be obtained in a straightforward way and we collect them in the following proposition. Under Assumption 2, $\{\varepsilon_t, t \in \mathbb{Z}\}$ is a sequence of i.i.d. random variables with $\mathbb{E}\varepsilon_t = 0$. Note that, by construction in equation (10) and under Assumption 2, ε_t is independent of L_τ, $\tau < t$.

Proposition 2

(i) $\{L_t\}$ is a martingale difference sequence, $\mathbb{E}L_t = 0$ and $\mathrm{Cov}\left[L_t, L_\tau\right] = 0$ for $t \neq \tau$.

(ii) Assume that $\mathbb{E}\varepsilon_t^2 = \sigma^2 < \infty$. If and only if $\sigma^2\boldsymbol{\beta}'\boldsymbol{\beta} < 1$, then $\{L_t\}$ is weakly stationary and

$$\mathrm{Var}L_t \;=\; \sigma^2 \Big/ \Big(1 - \sigma^2\boldsymbol{\beta}'\boldsymbol{\beta}\Big)\;. \tag{12}$$

The proof of Proposition 2 is straightforward and is relegated to Appendix C. The asset-liability hedge works 'on average', since the hedging error satisfies $\mathbb{E}\varepsilon_t = \mathbb{E}\left[\exp(r_t^A - r_t^L) - 1\right] = 0$ under Assumption 2. Part (i) of Proposition 2 thus shows that the pension loss is zero on average. Part (ii) shows that the filter $\boldsymbol{\pi}$ (or equivalently $\boldsymbol{\beta}$) in the funding contribution calculation under Assumption 3 must be carefully chosen to satisfy $\sigma^2\boldsymbol{\beta}'\boldsymbol{\beta} < 1$ if one wishes to avoid a pension loss with an infinite variance. The poorer the hedge is, the more stringent this condition becomes. In other words, if the hedging or replication error volatility σ is larger, the choice of funding parameter $\boldsymbol{\pi}$ (or equivalently $\boldsymbol{\beta}$) is more restricted.

4.3 State-space representation

The loss process $\{L_t\}$ in equation (10) can be written in state-space form as follows.

$$\mathbf{L}_{t+1} \;=\; \mathbf{A}_{t+1}\mathbf{L}_t \;+\; \mathbf{B}_{t+1}, \tag{13a}$$

$$L_t \;=\; \mathbf{C}_t\,\mathbf{L}_t \;+\; \mathbf{D}_t. \tag{13b}$$

In the state equation (13a), the coefficient matrix $\mathbf{A}_t$ is a $p \times p$ companion matrix which, in block form, is given by

$$\mathbf{A}_t \;=\; \begin{pmatrix} \boldsymbol{\xi}\varepsilon_t & \beta_p\varepsilon_t \\ \mathbf{I}_{p-1} & \mathbf{0}_{p-1} \end{pmatrix}, \tag{14}$$

where $\boldsymbol{\xi} = (\beta_1, \ldots, \beta_{p-1})$, $\mathbf{I}_{p-1}$ is the $(p-1) \times (p-1)$ identity matrix, and $\mathbf{0}_{p-1}$ is a $(p-1) \times 1$ column vector of zeros. $\mathbf{B}_t$ is a $p \times 1$ row vector,

$$\mathbf{B}_t \;=\; (-\varepsilon_t, 0, \ldots, 0)'. \tag{15}$$

The observation equation (13b) is trivial, with $\mathbf{C}_t = (1, 0, \ldots, 0)$ and $\mathbf{D}_t = 0$ (a scalar). Note that $\{\mathbf{A}_t\}$ and $\{\mathbf{B}_t\}$ are sequences of i.i.d. matrices and vectors respectively.

Equation (13a) is in the form of a first-order vector stochastic difference equation. Assuming the initial conditions stated for equation (11), $\mathbf{L}_l \to \mathbf{0}_p$ as $l \to -\infty$, and equation (13a) is solved by

$$\mathbf{L}_t = \sum_{j=0}^{\infty} \mathbf{A}_t \cdots \mathbf{A}_{t-j} \mathbf{B}_{t-j-1} + \mathbf{B}_t. \tag{16}$$

A theorem from Bougerol and Picard (1992a) states the conditions for strict stationarity of generalized autoregressive processes, such as $\{\mathbf{L}_t\}$ in equation (16), in terms of the top Lyapunov exponent,

$$\gamma = \lim_{t \to \infty} \frac{1}{t} \mathbb{E} \left[\log \|\mathbf{A}_t \cdots \mathbf{A}_1\| \right]. \tag{17}$$

Theorem 1 (Bougerol and Picard 1992a) *Suppose that $\{\mathbf{L}_t\}$ is irreducible and that $\mathbb{E} \log^+ \|\mathbf{A}_t\| < \infty$ and $\mathbb{E} \log^+ \|\mathbf{B}_t\| < \infty$. Then $\{\mathbf{L}_t\}$ is strictly stationary if and only if $\gamma < 0$.*

The notation $\log^+(x) = \max(0, \log x)$ is used in the above. The requirement that $\gamma < 0$ in Theorem 1 is analogous to the requirement that the discount factor be less than one for the present value of a non-random perpetuity to be finite.

4.4 Bilinear processes

In order to apply Theorem 1, we must elucidate whether the pension fund losses, expressed as $\{\mathbf{L}_t\}$, form an irreducible Markov chain. A bilinear representation is helpful in this regard, and also gives us access to conditions for the existence of moments.

Definition 1 (Bilinear process) (Pham 1993) The process $\{x_t, t \in \mathbb{Z}\}$ is a bilinear process if it satisfies

$$x_t = \sum_{j=1}^{q_1} a_j x_{t-j} + \sum_{j=0}^{q_2} c_j \epsilon_{t-j} + \sum_{j=1}^{q_3} \sum_{k=1}^{q_4} b_{jk} x_{t-j} \epsilon_{t-k} \tag{18}$$

where $a_j, c_j, b_{jk} \in \mathbb{R}$, $\epsilon_t \sim$ i.i.d. and $\mathbb{E}\epsilon_t = 0$.

The descriptor 'bilinear' refers to the fact that equation (18) is linear in $\{x_t\}$ and, separately, linear in $\{\epsilon_t\}$, but not linear in both because of the product term $x_{t-j}\epsilon_{t-k}$. Pham (1985, 1986, 1993) shows that a specific class of bilinear process, for which $b_{jk} = 0$ for $j < k$, can be written in state-space form with the coefficient matrices of the state equation being polynomials in the noise ϵ_t, these polynomials being of order 2 at most. Denote the coefficient matrices in the relevant state equation by $\widetilde{\mathbf{A}}_t$ and $\widetilde{\mathbf{B}}_t$, where $\widetilde{\mathbf{A}}_t$ is the leading coefficient. The following theorem gives a simple condition for strict stationarity with finite second moments.

Theorem 2 (Pham 1985) *Suppose that $\mathbb{E}\epsilon_t^4 < \infty$. The bilinear process $\{x_t\}$ in equation* (18)*, with $b_{jk} = 0$ for $j < k$, is strictly stationary with finite second moments if and only if the matrix equation*

$$\mathbf{Q} \;=\; \mathbb{E}\left[\widetilde{\mathbf{A}}_t \mathbf{Q} \widetilde{\mathbf{A}}_t'\right] \;+\; \mathrm{Var}\widetilde{\mathbf{B}}_t \tag{19}$$

admits a positive solution in $\mathbf{Q}$.

Pham (1986, 1993) also establishes a sufficient condition for the existence of even moments in the bilinear process $\{x_t\}$, in terms of the spectral radius (denoted by ρ) of the expectation of the Kronecker product of the leading coefficient matrix with itself n times (denoted by $\otimes n$ in superscript).

Theorem 3 (Pham 1986, 1993) *If $\mathbb{E}\epsilon_t^{2m} < \infty$ and $\rho\left(\mathbb{E}\left[\widetilde{\mathbf{A}}_t^{\otimes 2m}\right]\right) < 1$, where $m \in \mathbb{N}$, then $\mathbb{E}x_t^{2m} < \infty$.*

Returning to Theorem 1, we must address the issue of irreducibility before being able to apply Theorem 1. Kristensen (2009, Theorem 5) considers this issue for a somewhat different class of bilinear processes to which he refers as Class 2:

$$y_t \;=\; a_0 + \sum_{j=1}^{q_1} a_j y_{t-j} + \sum_{j=1}^{q_3} b_{jj} y_{t-j}\epsilon_{t-j}. \tag{20}$$

An irreducible Markov chain is never reduced, in terms of its transitions, to a strict subset of its state space: it can 'travel' to any point of its state space given enough time. If the Markov chain is viewed as a deterministic system with its noise process being a sequence of decision variables, then it should be possible to control this system and guide it to any point of its state space. Kristensen

(2009) thus considers the controllability of the bilinear process $\{y_t\}$ in equation (20) in terms of the polynomials

$$\Phi(z) \; = \; 1 - \sum_{j=1}^{q_1} a_j z^j, \qquad \Theta(z) \; = \; \sum_{j=1}^{q_3} b_{jj} z^{j-1}, \tag{21}$$

formed from the coefficients in equation (20). The noise process, viewed as a control variable, must also be unconstrained and Kristensen (2009) shows that it must be absolutely continuous.

Theorem 4 (Kristensen 2009) *The Markov chain in the state-space representation of the bilinear process* $\{y_t\}$ *in equation* (20) *is irreducible if and only if (i)* ϵ_t *has an absolutely continuous component (wrt. Lebesgue measure) in the neighborhood of zero, and (ii) the polynomials* $\Phi(z)$ *and* $\Theta(z)$ *have non-coincident zeros.*

4.5 Stochastic properties of the pension fund

We state the conditions for strict stationarity, ergodicity and the existence of even moments of the pension fund process $\{L_t, F_t, C_t\}$ in the following proposition, which we prove by reference to the theorems cited above. We also derive the first two moments explicitly.

Proposition 3 *(i) Suppose that* ε_t *has an absolutely continuous component in a neighborhood of zero. A necessary and sufficient condition for* $\{L_t\}$ *in equation* (10) *to be strictly stationary and ergodic is that* $\gamma < 0$*, where* γ *is defined in equation* (17) *with* $\mathbf{A}_t$ *defined in equation* (14)*.*

(ii) Assume that $\mathbb{E}\varepsilon_t^2 = \sigma^2 < \infty$*. Then* $\sigma^2 \boldsymbol{\beta}'\boldsymbol{\beta} < 1$ *is a necessary and sufficient condition for* $\{L_t\}$ *to be strictly stationary with finite second moments.*

(iii) If $\mathbb{E}\varepsilon_t^{2m} < \infty$ *and* $\rho\left(\mathbb{E}\left[\mathbf{A}_t^{\otimes 2m}\right]\right) < 1$*, where* $m \in \mathbb{N}$*, then* $\mathbb{E}L_t^{2m} < \infty$*.*

(iv) Stationary first and second moments (provided they exist) are as follows:

(a) $\mathbb{E}L_t = \mathbb{E}\overline{F}_t = \mathbb{E}C_t = 0$ *and* $\mathbb{E}F_t = 1$*.*

(b) $\mathrm{Cov}\left[L_t, L_\tau\right] = 0$ *for* $t \neq \tau$*.*

(c) $Q := \mathrm{Var}L_t = \sigma^2/(1 - \sigma^2\boldsymbol{\beta}'\boldsymbol{\beta})$*.*

(d) $\mathrm{Var}C_t = Q\boldsymbol{\pi}'\boldsymbol{\pi}$ *and* $\mathrm{Var}F_t = \mathrm{Var}\overline{F}_t = Q\boldsymbol{\lambda}'\boldsymbol{\lambda}$*, where* $\boldsymbol{\lambda} = \boldsymbol{\pi} + e^{-\delta}\boldsymbol{\beta}$*.*

(e) If $0 \le \tau \le p-1$*, then* $\text{Cov}\,[C_t, C_{t-\tau}] = Q \sum_{j=1}^{p-\tau} (\boldsymbol{\pi}\boldsymbol{\pi}')_{(j,j+\tau)}$

and $\text{Cov}\,[F_t, F_{t-\tau}] = Q \sum_{j=1}^{p-\tau} (\boldsymbol{\lambda}\boldsymbol{\lambda}')_{(j,j+\tau)}$.

If $\tau \ge p$*, then* $\text{Cov}\,[C_t, C_{t-\tau}] = \text{Cov}\,[F_t, F_{t-\tau}] = 0$.

The proof of Proposition 3 is given in Appendix D. Propositions 2 and 3 supply a complete description of the conditions required on the asset-liability hedging errors $\{\varepsilon_t\}$ and on the funding parameter vector $\boldsymbol{\pi}$ (or equivalently $\boldsymbol{\beta}$ and $\boldsymbol{\lambda}$) to achieve stationarity and finite moments in the pension fund process.

4.6 ARCH-type models and tail risk

In this section, we draw an analogy between the pension loss process $\{L_t\}$ in equation (10) and ARCH-type models. This tells us about the behavior of the tail of the distribution of the pension loss. Heavy tails signify that extreme unfavorable losses occur more frequently than expected compared to the normal distribution, and suggest that there is significant tail risk in the pension fund.

Definition 2 (GARCH) (Bollerslev 1986) The GARCH(q_2, q_1) process $\{z_t, t \in \mathbb{Z}\}$ satisfies

$$z_t \;=\; \sigma_t \, \epsilon_t \,, \qquad \sigma_t^2 \;=\; a_0 + \sum_{k=1}^{q_2} d_k \sigma_{t-k}^2 + \sum_{j=1}^{q_1} c_j z_{t-j}^2, \tag{22}$$

where $a_0 \in \mathbb{R}_{++}$, $c_j, d_k \in \mathbb{R}_+$, and ϵ_t is as in Definition 1.

The link between GARCH models and vector stochastic difference equations and Markov chain theory is well-known. It is exploited by Bougerol and Picard (1992b) and, latterly, by Kristensen (2009) to derive conditions for strict stationarity and ergodicity of GARCH processes. The link between GARCH and bilinear processes is less well-known but is noted by Tong (1990, pp. 116, 136). Rewriting equation (22) as

$$\sigma_t^2 \;=\; a_0 + \sum_{k=1}^{q_2} d_k \sigma_{t-k}^2 + \sum_{j=1}^{q_1} c_j \sigma_{t-j}^2 \epsilon_{t-j}^2 \tag{23}$$

and comparing the above to equation (18) or (20) makes the point that the GARCH squared-volatility σ_t^2 follows a non-negative bilinear process.

Indeed, both GARCH and bilinear models were developed with the aim of capturing non-constant volatility. Simulated data from both GARCH and bilinear models show that: (i) sample paths make large excursions from the mean, (ii) their distributions exhibit leptokurtosis, and (iii) QQ-plots indicate heavy-tailedness. See Fan and Yao (2003, pp. 153, 183) for example. This suggests that the pension loss process $\{L_t\}$ of equation (10) may also exhibit heavy tails.

A recent variant of the GARCH process is the LARCH (Linear-ARCH) which is extensively studied by Giraitis et al. (2000, 2004).

Definition 3 (LARCH) (Giraitis et al. 2000) The process $\{w_t, t \in \mathbb{Z}\}$ is a LARCH process if it satisfies

$$w_t \;=\; \sigma_t\,\epsilon_t\,, \qquad \sigma_t \;=\; a_0 + \sum_{j=1}^{\infty} c_j w_{t-j}, \tag{24}$$

where $a_0, c_j \in \mathbb{R}$ with $a_0 \neq 0$, and ϵ_t is as in Definition 1.

In the LARCH process in equation (24), volatility is a linear combination of past values of the process. Contrast this with the standard GARCH process in equation (22), where *squared* volatility is a linear combination of past *squared* values of the process.

The LARCH process describes two stylized facts of asset return data that are not adequately reflected by the standard GARCH model (Fan and Yao 2003, p. 170): (i) long memory or long-range dependence, and (ii) the leverage effect. The former describes the empirical observation of profound autocorrelation in absolute and squared returns, along with very slow decay as lag increases, slower than the exponential decay implied by GARCH. The latter describes the higher increase in volatility that is experienced after a fall in the market, as compared with volatility increases after a rise in market values, and as contrasted with the symmetric behavior of volatility in GARCH.

Now, $\{L_t\}$ in equation (11) and the LARCH process $\{w_t\}$ in equation (24) differ only in the finiteness of the series. Proposition 3 appears to suggest that the conditions on parameters of the loss process become progressively more stringent as one transits from requiring strict stationarity, to weak stationarity, to the existence of moments of increasing order, until one reaches a sufficiently high-order moment that is infinite. Giraitis et al. (2000, 2004) employ a Volterra series expansion of the LARCH 'volatility' term σ_t in equation (24), and they use the combinatorial formalism of *diagrams* to find various moments of w_t and σ_t. They also find that the existence of higher-order moments requires progressively greater restriction on the LARCH parameters.

Now, a well-known moment inequality in probability theory suggests that infinite higher-order moments should lead to a heavy-tailed probability distribution. At a fundamental level, heavy tails can be explained by an application of renewal theory to stochastic difference equations, resulting in products of random matrices, as shown in section 4.3 (Kesten 1973). This is suggestive, therefore, of possible significant tail risk in the pension fund. If large losses occur with a higher probability than anticipated, then larger contributions will be required from plan sponsors. As argued by Franzoni and Marín (2006), this may be underestimated by investors who therefore overprice firms with significantly underfunded pension plans.

5 Stochastic Simulations

In this section, we use stochastic simulations of the model in section 3 to complement the analysis of section 4 and investigate tail risk in the pension fund.

Tail risk We investigate the distribution of both the unfunded liability $\overline{F}_t$ and the cash flow or supplementary contribution C_t. In the case of $\overline{F}_t$, we are considering the risk that the pension fund does not accumulate enough assets. Leibowitz and Henrikssan (1988) and Sharpe and Tint (1990) refer to this as surplus risk. In the case of C_t, we are considering the risk that capital must be unexpectedly diverted to the pension fund to make up for large losses or deficits. This can potentially result in bankruptcy of the corporate plan sponsor (credit risk) or closure of the plan (covenant risk). We are interested in the extremes of both $\overline{F}_t$ and C_t, and therefore in tail risk.

Simulation inputs Some representative parameter values and assumptions are required to simulate the pension fund. To draw meaningful, realistic conclusions, we allow for liability growth through inflation. We assume that real returns are i.i.d., consistent with Assumption 2. We use the OECD data of Tapia (2008) on real returns on US defined-benefit voluntary occupational pension funds between 1988 and 2005 and we fit a normal distribution, with mean 6.83% and standard deviation 8.91%, to the log-return. Note that this is return per annum and net of price inflation.

Initialization and normalization There is no evidence in the OECD data (Tapia 2008) that the surveyed pension funds are managed solely with a view to hedging liabilities, but we assume that this is the case, again consistent with Assumption 2. More precisely, we assume that pension assets hedge inflation in pension liabilities. Hence, we set $\delta = 6.83\%$ and $\varepsilon_t \sim$ i.i.d. lognormal with a zero mean and a standard deviation of 9.6%.[7] The pension liability Y is normalized at 100% and an initial normal contribution of 4.2% is calculated.

Averaging and amortization To operationalize the counter-cyclical and smoothing behavior of pension funds, we simulate the U.S. practice of averaging asset values and amortizing resultant losses. The averaging period is denoted by n and the amortization period by m, so that the order p of the loss process $\{L_t\}$ in equation (10) is $p = m + n - 1$. Details are provided in Appendix E.

Simulation outputs The loss L_t and unfunded liability $\overline{F}_t$ are expressed as a percentage of the liability, whereas the supplementary contribution C_t is expressed as a percentage of the normal contribution. Simulations are performed for different smoothing levels, that is, for different combinations of amortization period m and averaging period n.

Computation To capture the tail of the distributions, we carry out 100,000 simulations, with checks for convergence using 50,000 and 200,000 simulations. We use a fixed seed to generate reproducible pseudo-random numbers and minimize sampling error when comparing results. To investigate strict stationarity, we estimate the top Lyapunov exponent, as defined in equation (17), using (see Bougerol and Picard 1992a)

$$\gamma \stackrel{\text{a.s.}}{=} \lim_{t\to\infty} \frac{1}{t} \log \|\mathbf{A}_t \cdots \mathbf{A}_1\|. \tag{25}$$

[7] If $\log X \sim \mathrm{N}(\mu, \sigma^2)$, then $\mathrm{Var}X = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right)$. Substituting $\mu = 6.83\%$ and $\sigma = 8.91\%$ gives a standard deviation of 9.6%.

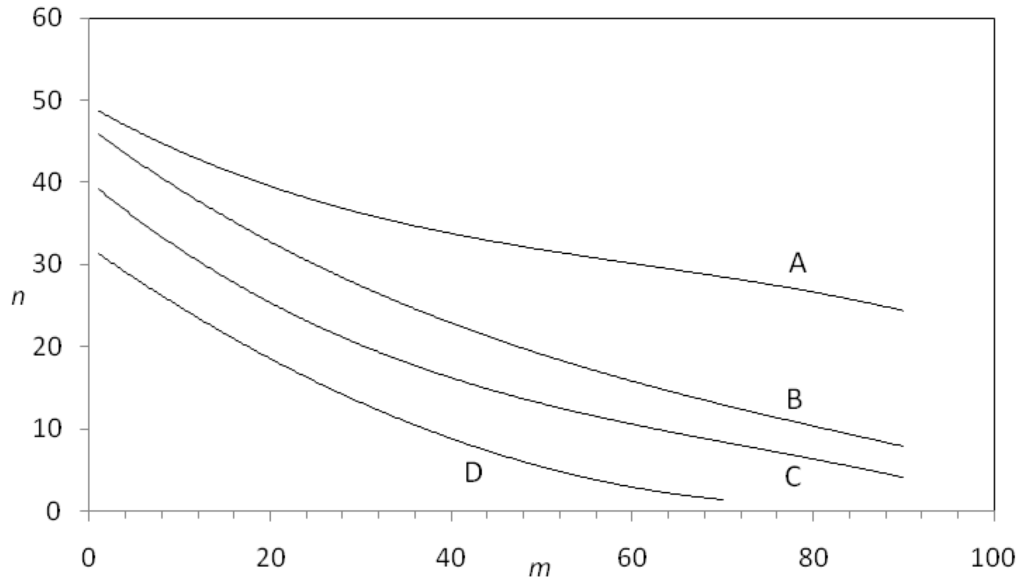

Fig. 1 Plots of smoothing parameters n versus m showing frontiers of strict stationarity (A), weak stationarity (B), existence of 4th moments (C) and existence of 6th moments (D). The region of stationarity or existence of moments is the region under and including the relevant curve.

For weak stationarity, we use the condition in part (ii) of Proposition 2 directly, so no simulation is required. To investigate the existence of moments, we could use the condition in part (iii) of Proposition 3 but this is a sufficient condition and could be unnecessarily restrictive, so we use stochastic simulation instead to test whether moments are finite.

Stationarity and existence of moments Our results are presented in Figure 1, which shows regions of stationarity or existence of moments, that is, the maximum values of n for different values of m for which stationarity or existence of moments holds.[8] As expected from the discussion in section 4.6, more stringent conditions, in the form of shorter averaging and amortization periods, apply as moments of a higher order are required to exist. This is indicative of a heavy-tailed distribution.

[8] Equivalently, the frontiers in Figure 1 give the maximum m for various n. Note that $m, n \in \mathbb{N}$, so a polynomial curve is fitted to non-integer values for readability. Similar depictions appear elsewhere: see Bollerslev (1986, Fig. 1) in the econometrics literature on GARCH models, and Tong (1990, Fig. 4.10, p. 171) in the non-linear time series literature on bilinear processes.

Probability plots Non-parametric estimates of the probability density functions (pdf) and the decumulative distribution functions (dcdf), as well as normal quantile-quantile (QQ) plots for the loss, the unfunded liability and the supplementary contribution are shown in Figures 2, 3 and 4 respectively. (The dcdf is the complementary cdf, i.e. $1 - \text{cdf}$.) Since we are concerned with tail risk, the behaviour of the right tail is of interest. The QQ-plots, in the bottom left panels of Figures 2, 3 and 4, show deviation from normality in the right tails. L_t, $\overline{F}_t$ and C_t all exhibit heavy-tailedness on the right.

Heavy-tailedness A measure of heavy-tailedness is the tail index introduced by Kesten (1973). This captures the fact that heavy-tailed distributions have tails that are Pareto-like and that decay like a power law. The tail index κ is an estimate of the rate of tail decay in the probability distribution of a random variable X (Fan and Yao 2003, p. 156):

$$\mathbb{P}(X > x) \ \sim \ Cx^{-\kappa}, \tag{26}$$

where C is some constant and $\sim$ means that the ratio of the l.h.s. and r.h.s. of equation (26) tends to 1 in the limit as $x \to \infty$. The log-log plots of the dcdf's of L_t, $\overline{F}_t$ and C_t, in the bottom right panels of Figures 2, 3 and 4 respectively, exhibit linearity. This gives a strong indication that their tails decay slowly according to a power law. In our numerical investigations, the power law decay was particularly evident for large smoothing parameter values (that is, large m and n). For comparison, we also show the tails of the normal distribution, with corresponding mean and standard deviation, which exhibit fast decay.

Tabulated statistics Various statistics for the loss, unfunded liability and supplementary contribution are shown in Tables 1, 2 and 3 respectively. There is some minor repetition in the Tables, e.g. statistics for $m = n = 5$, but this ensures legibility and aids with interpretation.

First moments First moment values are within $\pm 0.5\%$, consistent with zero expectation from part (iv) of Proposition 3, and are not reported in the Tables.

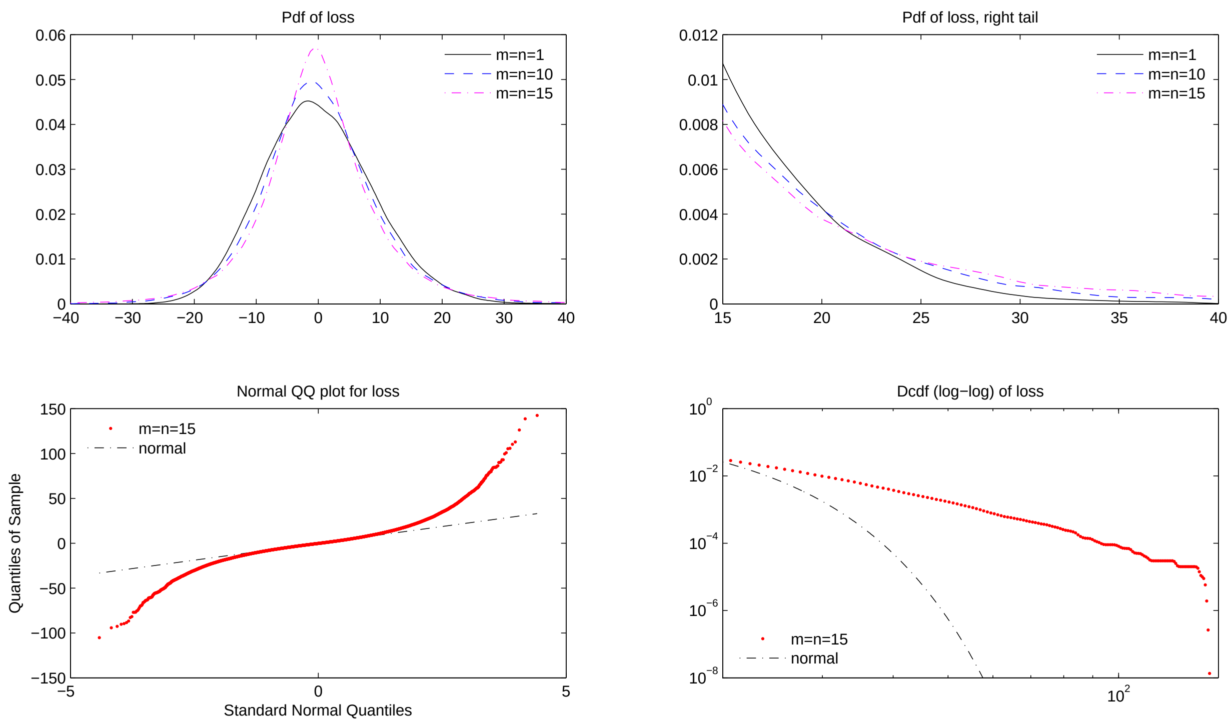

Fig. 2 Probability density estimates for the loss L_t (top left), with right tail enlarged (top right), for different smoothing parameters m and n. The normal QQ-plot (bottom left) shows deviation from normality in the tails. The log-log plot of the decumulative distribution function (bottom right) shows the slowly decaying Pareto-like right tail of the distribution of the loss, compared to the normal distribution.

Third and fourth moments We observe from Tables 1–3 that the loss, unfunded liability and supplementary contribution are all positively skewed and leptokurtic, with both skewness and kurtosis increasing as smoothing increases (i.e. as either m or n increases). It is noteworthy that the pdf curves of the loss in Figure 2 cross twice on either side of the mean, but this happens only once for the pdf curves of the unfunded liability in Figure 3.

Risk measures We investigate three risk measures: (i) The standard deviation is a classical measure of risk for symmetrically distributed random variables, and is used in Markowitz mean-variance portfolio theory, for example. (ii) The 95th percentile is commonly used, usually appearing in the form of the Value-at-Risk. (iii) We also consider the tail conditional expectation (TCE) at 95%. For

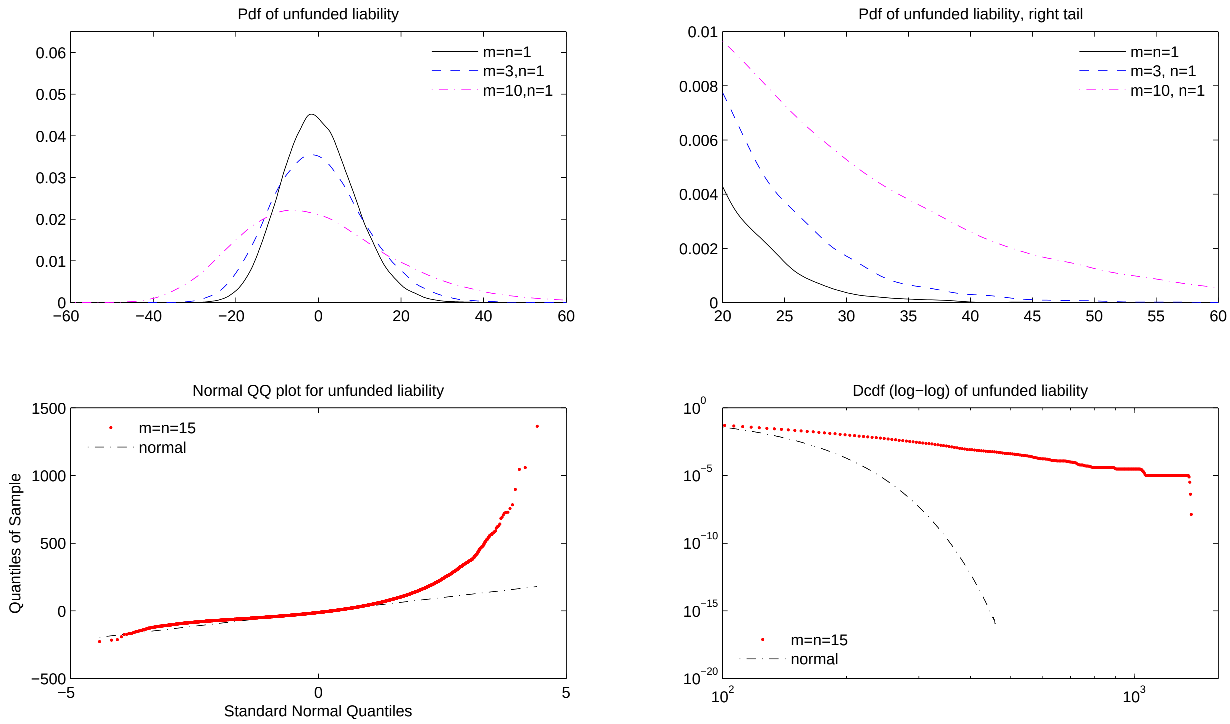

Fig. 3 Probability density estimates for the unfunded liability $\overline{F}_t$ (top left), with right tail enlarged (top right), for different smoothing parameters m and n. The normal QQ-plot (bottom left) shows deviation from normality in the right tail. The log-log plot of the decumulative distribution function (bottom right) shows the slowly decaying Pareto-like right tail of the distribution of the unfunded liability, compared to the normal distribution.

a continuously distributed random variable X with cdf $F_X(x)$, the 95th percentile is $F_X^{-1}(0.95)$ and the TCE is $\mathbb{E}\left[X \mid X > F_X^{-1}(0.95)\right]$.

Risk of underfunding ($\overline{F}_t$) In Tables 1 and 2, the loss and unfunded liability exhibit monotonically increasing standard deviations, 95th percentiles and tail conditional expectations as either m or n increases. The risk of underfunding therefore increases as smoothing increases. This is reasonable because more smoothing means that losses are paid off more slowly, so that losses tend to accumulate, leading to more volatile pension fund deficits.

Cash flow risk (C_t) Whereas funding risk increases with more smoothing, cash flow risk behaves in a less clear-cut fashion. Numerical work shows that there are roughly two situations:

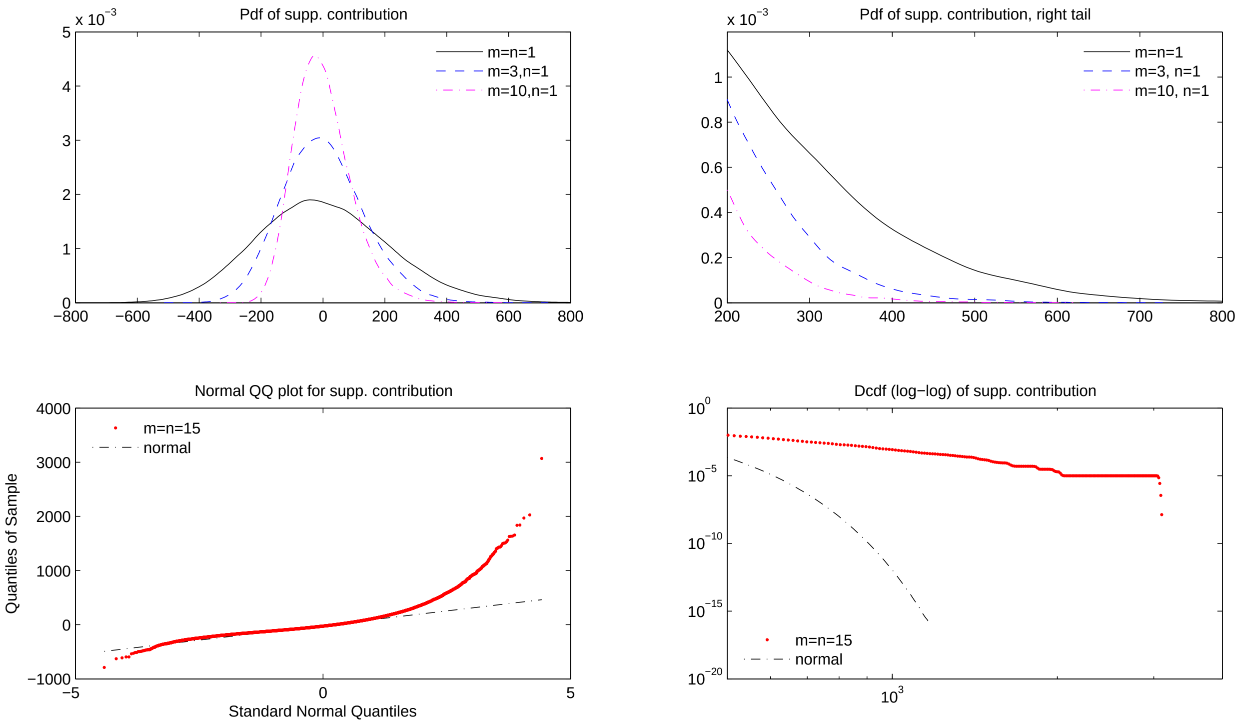

Fig. 4 Probability density estimates for the supplementary contribution C_t (top left), with right tail enlarged (top right), for different smoothing parameters m and n. The normal QQ-plot (bottom left) shows deviation from normality in the right tail. The log-log plot of the decumulative distribution function (bottom right) shows the slowly decaying Pareto-like right tail of the distribution of the supplementary contribution, compared to the normal distribution.

1. *Low smoothing levels.* Risk in C_t appears to have a minimum wrt. m and n, at least for low values of m and n. In Table 3, the standard deviation and the 95th percentile decrease and then increase as n increases when $m = 1, 5$ (and also as m increases when $n = 1, 5$). The behavior of the tail conditional expectation is somewhat more ambiguous but also appears to have a minimum. This feature of a minimum in contribution risk is also reproduced in Table 4 when $m = n$ and m increases.
2. *High smoothing levels.* For larger values of n (resp. m), risk in C_t behaves as in L_t and $\overline{F}_t$ in that it increases monotonically as m (resp. n) increases. We show only the case for $n = 20$ in Table 5 for brevity.

m	n	SD	Skewness	Kurtosis	95th perc.	TCE
1	1	8.9448	0.2709	3.1624	15.3640	19.8759
1	5	9.0055	0.2764	3.3121	15.4796	20.1974
1	10	9.1119	0.2894	3.6574	15.5676	20.7237
1	20	9.5933	0.3338	5.5416	15.9046	22.8381
1	30	11.3312	1.5157	51.5845	16.9074	27.9428
5	1	8.9982	0.2756	3.2924	15.4601	20.1599
5	5	9.0910	0.2873	3.5815	15.5404	20.6256
5	10	9.2470	0.3066	4.1541	15.6346	21.3688
5	20	9.9363	0.3827	7.4414	16.1100	24.0680
5	30	12.8186	4.8390	297.1861	17.5273	30.9406
1	1	8.9448	0.2709	3.1624	15.3640	19.8759
5	1	8.9982	0.2756	3.2924	15.4601	20.1599
10	1	9.0734	0.2847	3.5300	15.5041	20.5354
20	1	9.2813	0.3035	4.2465	15.7279	21.5301
30	1	9.5759	0.3342	5.4294	15.9127	22.7265
1	5	9.0055	0.2764	3.3121	15.4796	20.1974
5	5	9.0910	0.2873	3.5815	15.5404	20.6256
10	5	9.2070	0.3018	4.0025	15.6185	21.1817
20	5	9.5132	0.3268	5.2216	15.8628	22.5085
30	5	9.9500	0.4225	7.9966	16.0414	24.0746

Table 1 Standard deviation, skewness, kurtosis, 95th percentile and tail conditional expectation of the loss L_t for various combinations of m, n.

The above observations about contribution risk are consistent with the results of Owadally and Haberman (1999, 2004a) in the context of optimal pension funding. These authors use a simple pension fund model and look only at standard deviation rather than tail risk. Their explanation, in terms of a controlled process, is nevertheless applicable here. Smoothing, at low levels, means that the recognition of random losses is deferred so that contributions are dampened. But too much smoothing leads to feedback with long delays: losses are paid off too slowly and they accumulate leading to volatile deficits, and ultimately larger and more volatile contributions are required to the

m	n	SD	Skewness	Kurtosis	95th perc.	TCE
1	1	8.9448	0.2709	3.1624	15.3640	19.8759
1	5	14.3627	0.4995	3.4593	25.4955	33.7607
1	10	21.3724	0.7939	4.1891	38.9556	53.4606
1	20	39.6760	1.5662	8.6251	73.5585	110.1710
1	30	75.9424	3.4658	44.9189	129.4896	223.7662
5	1	13.8046	0.4784	3.4204	24.4489	32.2817
5	5	20.1227	0.7461	4.0338	36.5750	49.8702
5	10	27.7233	1.0638	5.1931	51.2338	72.7586
5	20	48.9183	2.0575	13.8148	89.5563	140.4511
5	30	97.8561	5.1100	108.4428	154.5248	286.5379
1	1	8.9448	0.2709	3.1624	15.3640	19.8759
5	1	13.8046	0.4784	3.4204	24.4489	32.2817
10	1	19.1937	0.7167	3.9599	34.8725	47.2743
20	1	29.0291	1.1599	5.6909	53.9020	77.1303
30	1	38.6314	1.6503	9.1320	71.6141	108.1702
1	5	14.3627	0.4995	3.4593	25.4955	33.7607
5	5	20.1227	0.7461	4.0338	36.5750	49.8702
10	5	25.9875	1.0047	4.9242	47.8631	67.5400
20	5	37.1329	1.5634	8.2782	69.4890	103.4388
30	5	48.7451	2.2625	15.6745	89.3260	142.1886

Table 2 Standard deviation, skewness, kurtosis, 95th percentile and tail conditional expectation of the unfunded liability $\overline{F}_t$ for various combinations of m, n.

pension fund. Thus, the application of counter-cyclical and smoothed funding rules, which aim to reduce contribution volatility for pension plan sponsors, may have a counter-productive effect if they are not properly calibrated.

6 Conclusion

Like other financial institutions, corporate pension funds suffered deep losses in the financial crisis of 2008. Defined benefit pension funds are unique in that they are governed by specialized pension

m	n	SD	Skewness	Kurtosis	95th perc.	TCE
1	1	212.9714	0.2709	3.1624	365.8107	473.2349
1	5	111.9033	0.4337	3.3821	196.7087	259.3782
1	10	99.3020	0.6546	3.9611	176.9619	241.7429
1	20	119.6733	1.1068	6.9659	212.4116	314.4828
1	30	195.8838	2.0625	26.7591	315.2703	535.9906
5	1	110.1281	0.4323	3.3636	193.8599	255.2223
5	5	106.5412	0.7012	3.9506	191.8650	261.2202
5	10	105.2168	0.9814	5.0788	190.9653	271.4525
5	20	133.2633	1.5779	10.3216	238.4419	368.5780
5	30	233.7507	3.2776	53.3562	357.2302	648.6941
1	1	212.9714	0.2709	3.1624	365.8107	473.2349
5	1	110.1281	0.4323	3.3636	193.8599	255.2223
10	1	92.7231	0.6465	3.8408	166.3623	225.2631
20	1	90.3605	1.0495	5.4235	165.8961	235.4938
30	1	98.5287	1.5051	8.4797	180.8426	271.4921
1	5	111.9033	0.4337	3.3821	196.7087	259.3782
5	5	106.5412	0.7012	3.9506	191.8650	261.2202
10	5	100.3773	0.9646	4.8758	183.9542	258.3171
20	5	103.2623	1.4590	7.7762	190.7173	283.5039
30	5	115.4426	2.0245	12.8424	208.7338	333.2008

Table 3 Standard deviation, skewness, kurtosis, 95th percentile and tail conditional expectation of the supplementary contribution C_t for various combinations of m, n.

m	n	SD	Skewness	Kurtosis	95th perc.	TCE
1	1	212.9714	0.2709	3.1624	365.8107	473.2349
5	5	106.5412	0.7012	3.9506	191.8650	261.2202
10	10	112.6476	1.3109	6.8758	206.8592	303.4156
15	15	142.3291	2.1914	15.5769	254.3503	412.1991
20	20	211.2505	4.4592	70.2110	337.4722	621.8401

Table 4 Standard deviation, skewness, kurtosis, 95th percentile and tail conditional expectation of the supplementary contribution C_t for various m, n when $m = n$.

m	n	SD	Skewness	Kurtosis	95th perc.	TCE
1	20	119.6733	1.1068	6.9659	212.4116	314.4828
5	20	133.2633	1.5779	10.3216	238.4419	368.5780
10	20	154.4189	2.1965	17.0755	270.4143	444.9742
20	20	211.2505	4.4592	70.2110	337.4722	621.8401
30	20	291.9566	10.4929	414.1530	390.6603	817.6284

Table 5 Standard deviation, skewness, kurtosis, 95th percentile and tail conditional expectation of the supplementary contribution C_t for various m when $n = 20$.

funding and accounting rules, which permit cash flows and costs to be spread over time. In this paper, we sought to establish whether the smoothing mechanisms entrenched in pension funding rules and practice could be related to large pension losses and deficits.

We reviewed the evidence in the accounting literature that spreading cash flows over time hinders investors' efficient valuation of firms. There is also significant evidence in the corporate finance literature that increases in mandatory pension contributions negatively impact firms' capital investment and are accompanied by depressed stock returns subsequently. Pension funding rules and practice also influence the value of firms because of the tax savings and operational leverage that flexibility in discretionary pension contributions affords, after allowing for market imperfections. The mechanism underlying funding for pension plans is therefore important to investors and management, as well as to the employee membership of these plans.

The evidence, from the U.S. and worldwide, that funding rules are implemented in a countercyclical fashion was also reviewed. This was particularly visible in the aftermath of the financial crisis in 2008. Proposed funding rules in the European Union are explicit in allowing for countercyclical measures. These are also implicit—through averaging in liability discount rates, asset values and choice of valuation parameters—in funding practices utilized in the U.S. and elsewhere. These methods are intended to limit the volatility of cash flows required from plan sponsors.

We built a stylized model, with the starting premise that pension liabilities are optimally but imperfectly hedged by assets held in the pension fund, and showed that pension losses follow a stochastic bilinear process. We were able to leverage results from time series analysis and Markov chain theory to derive a complete description of the stochastic dynamics of the pension fund. The range of permissible funding parameters appears to become more restricted if one demands a certain level of stability (more precisely, stochastic stationarity in the strict and weak senses) in the pension fund. The higher-order moments of the pension loss may be infinite. Closer inspection revealed that the loss process follows an ARCH-type model.

We postulated therefore that the loss process would exhibit bursts of volatility and that its distribution would be leptokurtic. Stochastic simulations were carried out using data on U.S. defined benefit pension funds and implementing the U.S. practice of asset value averaging and gain/loss amortization. They revealed that the distributions of the loss, unfunded liability and supplementary contributions required to the pension fund did indeed exhibit skewness, leptokurtosis and heavy tails with Pareto-like slow decay. These were more pronounced when the amount of allowable smoothing, through longer amortization and averaging periods, was itself more pronounced.

An important implication for policy-makers and regulators is that the amount of smoothing in pension funding rules should be limited because of the emergence of tail risk. In other words, these rules may work under normal circumstances but nonlinear dynamic effects may entail unpredictably large but rare losses either in individual pension funds or in a systemic way. Policy should also have regard for all in-built sources of counter-cyclicality and for the combined dynamic effects caused not just by funding rules but also by funding practices and by discretionary employer contributions being managed for tax or operational reasons. In particular, the Solvency II capital adequacy standard, as proposed by the European Union for pension funds, undergoes various Quantitative Impact Studies and these may need re-calibration to identify the presence of tail risk. Finally, our findings have a valuation implication which is of key relevance to investors in firms which engage in counter-cyclical deferral and smoothing of pension cash flows. Large pension losses, should they occur, are likely to affect capital investment in these firms and lead to depressed stock returns.

A Proof of Approximation in Equation (5)

Recall that Y is the liability when it is calculated at the term-independent discount rate δ. From equation (2),

$$Y \;=\; \sum_{x=E}^{R} m_x \ell_x \widetilde{\alpha}(x), \tag{27}$$

where $\widetilde{\alpha}(x)$ is the present value, at rate δ and at time t, of a deferred annuity paying \$1 yearly in advance from retirement age R till death, for an individual aged $x \leq R$ at time t:

$$\widetilde{\alpha}(x) \;=\; \sum_{j=0}^{\infty} e^{-(R-x+j)\,\delta}\, \ell_{R+j}/\ell_x. \tag{28}$$

The normal cash flow Z_t is $Z_t = Z_t^{(1)} - Z_t^{(2)} + Z_t^{(3)}$ where each component is defined in section 3.2.

$$Z_t \;=\; m_R \ell_R \alpha(t,R) \;-\; \sum_{x=E}^{R-1} (m_{x+1} - m_x) \ell_x \alpha(t,x) \;+\; e^{-\delta}(\delta_t - \delta) Y_t. \tag{29}$$

By analogy to Y in equation (27), we may define Z to be the normal cash flow Z_t when it is calculated using the term-independent discount rate δ.

$$Z \;=\; m_R \ell_R \widetilde{\alpha}(R) \;-\; \sum_{x=E}^{R-1} (m_{x+1} - m_x) \ell_x \widetilde{\alpha}(x). \tag{30}$$

Noting that $m_E = 0$ by definition (no benefit right endowed at entry into the plan), the above may be rewritten as

$$Z \;=\; \sum_{x=E}^{R} m_x \left[\ell_x \widetilde{\alpha}(x) - \ell_{x-1} \widetilde{\alpha}(x-1) \right]. \tag{31}$$

The term in square brackets above may be further simplified using equation (28):

$$\ell_x \widetilde{\alpha}(x) - \ell_{x-1}\widetilde{\alpha}(x-1) \;=\; \sum_{j=0}^{\infty} e^{-(R-x+j)\,\delta}\, \ell_{R+j} (1 - e^{-\delta}) \;=\; \ell_x\, \widetilde{\alpha}(x)\, (1 - e^{-\delta}). \tag{32}$$

Finally, substituting equation (32) in equation (31) and using Y as defined in equation (27) gives

$$Z \;=\; \sum_{x=E}^{R} m_x\, \ell_x\, \widetilde{\alpha}(x)\, (1 - e^{-\delta}) \;=\; Y(1 - e^{-\delta}). \tag{33}$$

As in Assumption 1, where the liability Y_t is linearized around δ, we may linearize Z_t in equation (29):

$$Z_t \;=\; Z e^{-D_Z(\delta_t - \delta)} \;+\; e^{-\delta}(\delta_t - \delta) Y_t, \tag{34}$$

where $D_Z = -Z^{-1} \partial Z / \partial \delta$. Using equation (3) from Assumption 1 as well as equation (33), we obtain

$$\frac{Z_t}{Y_t} \;=\; \frac{Z}{Y} \frac{Y}{Y_t} e^{-D_Z(\delta_t - \delta)} \;+\; e^{-\delta}(\delta_t - \delta) \;=\; (1 - e^{-\delta}) e^{-(D_Z - D_Y)(\delta_t - \delta)} \;+\; e^{-\delta}(\delta_t - \delta). \tag{35}$$

The difference between the modified durations of Z and Y, which appears in equation (35), is easily derived from equation (33). We note that $\partial Z/\partial\delta = (\partial Y/\partial \delta)(1 - e^{-\delta}) + Y e^{-\delta}$. Upon dividing by both sides of equation (33), we find that $Z^{-1}(\partial Z/\partial \delta) = Y^{-1}(\partial Y/\partial\delta) + (e^{\delta} - 1)^{-1}$. Hence, $D_Z - D_Y = -(e^{\delta} - 1)^{-1}$. Substituting this into equation (35) results in

$$\frac{Z_t}{Y_t} \;=\; (1 - e^{-\delta}) \exp\left[-\frac{\delta_t - \delta}{e^{\delta} - 1} \right] \;+\; e^{-\delta}(\delta_t - \delta) \;\approx\; (1 - e^{-\delta}) \left[1 - \frac{\delta_t - \delta}{e^{\delta} - 1} \right] \;+\; e^{-\delta}(\delta_t - \delta) \;=\; 1 - e^{-\delta}. \tag{36}$$

This concludes the proof of the approximation in equation (5). $\square$

B Proof of Proposition 1

It immediately follows from equation (7) and Assumption 2 that

$$\mathbb{E}\left[F_t \mid \mathcal{F}_{t-1}\right] \;=\; \left(F_{t-1}+C_{t-1}-Z\right)\mathbb{E}\left[\exp\left(r_t^A-r_t^L+\delta\right)\right] \;=\; e^{\delta}\left(F_{t-1}+C_{t-1}-Z\right). \tag{37}$$

From the definition of the loss L_t in equation (8), we therefore obtain

$$L_t = e^{\delta}(F_{t-1}+C_{t-1}-Z)-F_t. \tag{38}$$

It is convenient to work in terms of the unfunded liability, $\overline{F}_t = 1-F_t$, and replace Z using equation (5), yielding $\overline{F}_t - e^{\delta}\overline{F}_{t-1} = L_t - e^{\delta}C_{t-1}$. Upon substituting the supplementary contribution from equation (9), we get $\overline{F}_t - e^{\delta}\overline{F}_{t-1} = L_t - e^{\delta}\boldsymbol{\pi}'\,\mathbf{L}_{t-1}$. This is a linear difference equation in $\overline{F}_t$ which may be solved using standard techniques. It is easily verified that its solution is

$$\overline{F}_t \;=\; \boldsymbol{\lambda}'\,\mathbf{L}_t, \tag{39}$$

where $\boldsymbol{\lambda} = (\lambda_1,\ldots,\lambda_p)'$ and $\lambda_j = e^{(j-1)\delta}+\pi_j-\sum_{k=1}^{j} e^{(j-k)\delta}\pi_k$ for $j\in[1,p]$. Equation (39) shows that the unfunded liability is the accumulated value of unpaid losses.

Now, the main recurrence relation may be obtained by substituting F_t from the budget constraint in equation (7) directly into the r.h.s. of equation (38). Simplifying with the help of equation (5) and using the unfunded liability $\overline{F}_t = 1-F_t$ results in

$$L_t \;=\; \left[\exp(r_t^A-r_t^L)-1\right]\left[e^{\delta}(\overline{F}_{t-1}-C_{t-1})-1\right]. \tag{40}$$

Replacing C_t from equation (9) and $\overline{F}_t$ from equation (39) then yields the recurrence relation in equation (10) in Proposition 1. We note that $\boldsymbol{\beta} = e^{\delta}(\boldsymbol{\lambda}-\boldsymbol{\pi})$ and that $\beta_p = 0$. □

C Proof of Proposition 2

Part (i). $\mathbb{E}L_t = \mathbb{E}\left[\varepsilon_t\right]\mathbb{E}\left[v_t\right]=0$, by independence of ε_t and v_t. $\mathbb{E}L_t$ exists $\Leftrightarrow \mathbb{E}|L_t|<\infty$ (Feller 1971, p. 136). $\{L_t\}$ is evidently adapted to $\{\mathcal{F}_t\}$. For $\tau<t$, ε_t is independent of $\mathcal{F}_\tau$ and $\mathbb{E}\left[L_t \mid \mathcal{F}_\tau\right]=\mathbb{E}\left[\varepsilon_t\right]\mathbb{E}\left[v_t \mid \mathcal{F}_\tau\right]=0$. Hence, $\{L_t\}$ is a martingale difference sequence (m.d.s.) (Davidson 1994, p. 230). The covariance result is a property of m.d.s.: see e.g. Davidson (1994, Corollary 15.4). It is easily verified: $\mathbb{E}\left[L_tL_\tau\right]=\mathbb{E}\left[\varepsilon_t v_t L_\tau\right]=\mathbb{E}\left[\varepsilon_t\right]\mathbb{E}\left[v_tL_\tau\right]=0$ when $t>\tau$.

Part (ii). Given the results in part (i), we need only consider $\mathrm{Var}L_t$ to establish covariance-stationarity of $\{L_t\}$. Since $\mathbb{E}\varepsilon_t=0$ and $\mathbb{E}\left[L_tL_\tau\right]=0$ for $\tau\neq t$, it follows that $\mathbb{E}v_t^2=\sum\beta_j^2\mathbb{E}L_{t-j}^2+1$. Now, $\mathrm{Var}L_t=\mathbb{E}L_t^2=\mathbb{E}\left[\varepsilon_t^2\right]\mathbb{E}\left[v_t^2\right]$, giving a pth-order linear difference equation: $\mathrm{Var}L_t \;=\; \sum_{j=1}^{p}\sigma^2\beta_j^2\mathrm{Var}L_{t-j} \;+\; \sigma^2$. It is well-known that stability depends on the zeros of the characteristic polynomial $z^{p-1}-\sigma^2\beta_1^2z^{p-2}-\cdots-\sigma^2\beta_p^2$ satisfying $|z|<1$. Since the coefficients $\sigma^2\beta_j^2$, $j=1,\ldots,p$ are non-negative, a necessary and sufficient condition for stability is that $\sum\sigma^2\beta_j^2 =$

$\sigma^2\boldsymbol{\beta}'\boldsymbol{\beta} < 1$. The initial values $L_l, \ldots, L_{l-p+1}$ are known at some time $l < 0$ and, letting $l \to -\infty$, all second moment terms containing the initial values $L_l, \ldots, L_{l-p+1}$ vanish geometrically iff $\sigma^2\mathbf{b}'\mathbf{b} < 1$. Finally, from the pth-order linear difference equation in $\mathrm{Var}L_t$ above, $\mathrm{Var}L_t$ converges to $\sigma^2/\left(1-\sigma^2\sum\beta_j^2\right)$. □

D Proof of Proposition 3

First, note that $L_t = \varepsilon_t v_t$ (equation (11a)) with v_t independent of ε_t and $\{\varepsilon_t\}$ being i.i.d. (Assumption 2) and hence trivially stationary and ergodic, so that ergodicity and stationarity carry across from $\{v_t\}$ to $\{L_t\}$ and vice-versa.

Part (i). To apply Theorem 1, we first obtain the conditions under which $\{\mathbf{L}_t\}$ is irreducible.

$\{L_t\}$ in equation (10) is not a bilinear process, according to Definition 1 because $k \neq 0$ in equation (18). However, using equations (11a) and (11b), we can write

$$v_t \;=\; \sum_{j=1}^{p} b_j v_{t-j}\epsilon_{t-j} \;-\; 1, \tag{41}$$

$$v_t + 1 \;=\; \sum_{j=1}^{p} b_j (v_{t-j}+1)\epsilon_{t-j} \;-\; \sum_{j=1}^{p} b_j\epsilon_{t-j}. \tag{42}$$

This conforms to the class of bilinear processes in equation (20) and Theorem 4 is therefore applicable.

The state-space representation of $\{v_t\}$ consists of the state equation (13a), with $\mathbf{A}_t$ and $\mathbf{B}_t$ defined in equations (14) and (15) respectively, together with the observation equation $v_t = \boldsymbol{\beta}'\mathbf{L}_t - 1$. In accordance with Theorem 4, the irreducibility of the Markov chain $\{\mathbf{L}_t\}$ in equation (13a) may be guaranteed by insisting on the condition on the density of ε_t, which does not violate Assumption 2, specifically $\mathbb{E}\varepsilon_t = 0$. Controllability of $\{v_t\}$ follows by comparing equations (41) and (20) and noting that there is no "autoregressive" element and no possibility of coincident roots in the polynomials of equation (21).

To conclude the application of Theorem 1, we proceed to show that the technical conditions $\mathbb{E}\log^+\|\mathbf{A}_t\| < \infty$ and $\mathbb{E}\log^+\|\mathbf{B}_t\| < \infty$ do hold. Using the maximum row sum matrix norm induced by the max vector norm (Horn and Johnson 1985, p. 295), $\|\mathbf{A}_t\| = \max(1, |\varepsilon_t|\sum|\beta_j|)$ and $\|\mathbf{B}_t\| = |-\varepsilon_t| = |\varepsilon_t|$. Hence, $\log^+\|\mathbf{A}_t\| = \log\left[\max(1, \max(1, |\varepsilon_t|\sum|\beta_j|))\right] = \log^+\left(|\varepsilon_t|\sum|\beta_j|\right)$.

Now, $0 \leq \log^+(k|\varepsilon_t|) \leq k|\varepsilon_t|$, for any $k \geq 0$, and hence $\mathbb{E}\log^+(k|\varepsilon_t|) \leq \mathbb{E}\left[k|\varepsilon_t|\right]$. By assumption, $\mathbb{E}\varepsilon_t = 0$. Therefore, $\mathbb{E}\left[k\varepsilon_t\right]$ exists and is finite, which is equivalent to $\mathbb{E}|k\varepsilon_t| < \infty$ (Feller 1971, p. 136). Hence, $\mathbb{E}\log^+(k|\varepsilon_t|) < \infty$. Replacing k by $\sum|\beta_j|$ yields $\mathbb{E}\log^+\|\mathbf{A}_t\| < \infty$. Replacing k by 1 yields $\mathbb{E}\log^+\|\mathbf{B}_t\| < \infty$. This may also be confirmed by a direct application of Theorem 5 of Kristensen (2009).

This concludes the proof of part (i).

Part (ii). This is an application of Theorem 2. In equation (41), $\{v_t\}$ conforms to the bilinear process in Definition 1 as required in Theorem 2. Theorem 2 is therefore applicable to $\{v_t\}$ with $\widetilde{\mathbf{A}}_t = \mathbf{A}_t$ and $\widetilde{\mathbf{B}}_t = \mathbf{B}_t$. A straightforward inspection of the proof of Theorem 2 (Pham 1985, Theorem 4.1) shows that, because $\mathbf{B}_t$ in equation (15) does not contain a quadratic in ε_t (as would be the case for a general bilinear process under Theorem 2), we need not insist that $\mathbb{E}\varepsilon_t^4 < \infty$ and merely require that $\mathbb{E}\varepsilon_t^2 < \infty$.

We claim that $\mathbf{Q} = Q\mathbf{I}_p$, where Q is the stationary variance of $\{L_t\}$, given on the r.h.s. of equation (12), is a positive solution of the matrix equation (19), thereby satisfying the condition of Theorem 2. This may be seen by noting that $\mathbb{E}\mathbf{L}_t = \mathbf{0}_p$ from part (i) of Proposition 2 and using the state equation (13a) to obtain $\mathrm{Var}\mathbf{L}_{t+1} = \mathbb{E}\left[\mathrm{Var}\left[\mathbf{L}_{t+1} \mid \varepsilon_{t+1}\right]\right] + \mathrm{Var}\left[\mathbb{E}\left[\mathbf{L}_{t+1} \mid \varepsilon_{t+1}\right]\right] = \mathbb{E}\left[\mathbf{A}_{t+1}(\mathrm{Var}\mathbf{L}_t)\mathbf{A}'_{t+1}\right] + \mathrm{Var}\mathbf{B}_{t+1}$. In the covariance-stationary state, this equation becomes $\mathrm{Var}\mathbf{L}_t = \mathbb{E}\left[\mathbf{A}_t(\mathrm{Var}\mathbf{L}_t)\mathbf{A}'_t\right] + \mathrm{Var}\mathbf{B}_t$ which corresponds to equation (19), and is satisfied by $\mathrm{Var}\mathbf{L}_t = Q\mathbf{I}_p$ in accordance with Proposition 2.

This concludes the proof of part (ii).

Part (iii). This is an application of Theorem 3. In equation (41), $\{v_t\}$ conforms to the bilinear process in Definition 1 as required in Theorem 3, which is therefore applicable to $\{v_t\}$ with $\widetilde{\mathbf{A}}_t = \mathbf{A}_t$ and $\widetilde{\mathbf{B}}_t = \mathbf{B}_t$. Note, from equation (11a), that $\mathbb{E}\left[L_t^{2m}\right] = \mathbb{E}\left[\varepsilon_t^{2m}\right]\mathbb{E}\left[v^{2m}\right]$.

Part (iv). From Proposition 2, $\{L_t\}$ is a martingale difference sequence (m.d.s.). Therefore, $\{C_t\}$ and $\{\overline{F}_t\}$ are weighted sums of m.d.s. from equations (9) and (39). Moments of C_t, $\overline{F}_t$ and L_t follow from equations (9), (39), (10) respectively. $\{L_t\}$ is serially uncorrelated but $\{C_t\}$ and $\{\overline{F}_t\}$ have zero autocorrelation at lags p and greater. Parts (iv) and (v) are consistent for lag $\tau = 0$ since, e.g., $\sum_{j=1}^{p}(\boldsymbol{\pi}\boldsymbol{\pi}')_{(j,j)} = \mathrm{tr}(\boldsymbol{\pi}\boldsymbol{\pi}') = \boldsymbol{\pi}'\boldsymbol{\pi}$. □

E Averaging and amortization

Actuarial standards of practice discuss asset value averaging in detail: see for example ASB (2009). A taxonomy of practical methods is presented by Committee on Retirement Systems Research (2001) and described mathematically by Owadally and Haberman (2004b). This results in averaging losses, say over $n \in \mathbb{N}$ years, with allowance for interest:

$$L_t^A = \frac{1}{n}\sum_{j=0}^{n-1} e^{j\delta} L_{t-j}. \tag{43}$$

Smoothed losses are then paid off, usually in equal tranches, over a number of years (say $m \in \mathbb{N}$) through the device of amortization (ASB 2007a; McGill et al 2004). A supplementary contribution is paid at time t by amortizing every previous "averaged" loss that has not yet been paid in full over, say, m years:

$$C_t = \frac{1}{\alpha_m}\sum_{j=0}^{m-1} L_{t-j}^A, \tag{44}$$

where $\alpha_m = 1 + e^{-\delta} + \cdots + e^{-(m-1)\delta}$ is the present value of an annuity over m years.

Working through the double summation in equations (43) and (44), one may express the supplementary contribution in terms of the filter specified in equation (9), with $p = m + n - 1$. Letting $\overline{p} = \max(m, n)$ and $\underline{p} = \min(m, n)$

and using some algebra returns the following filter weights π_j.

$$\pi_j \;=\; \begin{cases} (n\alpha_m)^{-1}\, e^{(j-1)\delta}\, \alpha_j & \text{if } j \in [1, \underline{p}] \\ (n\alpha_m)^{-1}\, e^{[\min(j,n)-1]\delta}\, \alpha_{\underline{p}} & \text{if } j \in [\underline{p}, \overline{p}] \\ (n\alpha_m)^{-1}\, e^{(n-1)\delta}\, \alpha_{p-j+1} & \text{if } j \in [\overline{p}, p] \\ 0 & \text{otherwise} \end{cases} \tag{45}$$

The weights $\{\pi_j\}$ above do satisfy the requirement that $\sum_{j=1}^{p} e^{-(j-1)\delta}\pi_j = 1$ as set out under equation (9). Equation (9) shows that the supplementary contribution includes a proportion π_j of the loss that occurred j years ago, $j \in [1, p]$.

References

Adams J, Smith DJ (2009) Mind the gap: using derivatives overlays to hedge pension duration. Financ Anal J 65(4):60–67

ASB (2007a) Actuarial Standard of Practice No. 4. Measuring Pension Obligations and Determining Pension Plan Costs or Contributions. Actuarial Standards Board, Washington, DC

ASB (2007b) Actuarial Standard of Practice No. 27. Selection of Economic Assumptions for Measuring Pension Obligations. Actuarial Standards Board, Washington, DC

ASB (2009) Actuarial Standard of Practice No. 44. Selection and Use of Asset Valuation Methods for Pension Valuations. Actuarial Standards Board, Washington, DC

Bergstresser D, Desai MA, Rauh JD (2006) Earnings manipulation, pension assumptions, and managerial investment decisions. Q J Econ 121:157–195

Berkelaar A, Kouwenberg R (2003) Retirement saving with contribution payments and labor income as a benchmark for investments. J Econ Dyn Control 27:1069–1097

Blake, D (2006a) Pension Economics. John Wiley, Chichester, UK

Blake, D (2006b) Pension Finance. John Wiley, Chichester, UK

Bollerslev T (1986) Generalized autoregressive conditional heteroscedasticity. J Econometrics 31:307–327.

Bougerol P, Picard N (1992a) Strict stationarity of generalized autoregressive processes. Ann Probab 20(4):1714–1730

Bougerol P, Picard N (1992b) Stationarity of GARCH processes and of some nonnegative time series. J Econometrics 52:115–127

Broeders D, Chen A (2010) Pension regulation and the market value of pension liabilities: a contingent claims analysis using Parisian options. J Bank Financ 34(6):1201–1214

Campbell JL, Dhaliwal DS, Schwartz WC (2010) Equity valuation effects of the Pension Protection Act of 2006. Contemp Account Res 27(2):469–536

Campbell JL, Dhaliwal DS, Schwartz WC Jr (2012) Financing constraints and the cost of capital: evidence from the funding of corporate pension plans. Rev Financ Stud 25(3):868–912

Citigroup (2010). Citigroup Pension Liability Index—revised methodology. http://www.soa.org/Files/Sections/pen-cpli-methodology.pdf. Accessed 10 July 2012

Committee on Retirement Systems Research (2001) Survey of asset valuation methods for defined benefit pension plans. Pension Forum, 13(1):1–49. Society of Actuaries, Schaumburg, Illinois

Davidson, J (1994) Stochastic limit theory. Oxford University Press

Dickson DCM, Hardy MR, Waters HR (2009) Actuarial mathematics for life contingent risks. Cambridge University Press

Duffie D, Richardson HR (1991) Mean-variance hedging in continuous time. Ann Appl Probab 1:1–15

EIOPA (2012). Solvency II. European Insurance and Occupational Pensions Authority (EIOPA), Frankfurt am Main, Germany. https://eiopa.europa.eu/activities/insurance/solvency-ii/. Accessed 12 July 2012

Engle RF (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50:987–1007

Fan J, Yao Q (2003) Nonlinear time series: nonparametric and parametric methods. Springer

FASB (1985) Statement of Financial Accounting Standards no. 87: employers accounting for pensions. Financial Accounting Standards Board (FASB), Stamford, CT

FASB (2006) Statement of Financial Accounting Standards no. 158: employers accounting for defined benefit pensions and other postretirement plans. Financial Accounting Standards Board (FASB), Stamford, CT

Feller W (1971) An introduction to probability theory and its applications, vol. 2, 2nd ed. John Wiley & Sons

Franzoni F (2009) Underinvestment vs. overinvestment: evidence from price reactions to pension contributions. J Financ Econ 92:491–518

Franzoni F, Marín JM (2006) Pension plan funding and stock market efficiency. J Financ 61(2):921–956

Giraitis L, Leipus R, Robinson PM, Surgailis D (2004) LARCH, leverage and long memory. J Financ Economet 2(2):177–210

Giraitis L, Robinson PM, Surgailis D (2000b) A model for long memory conditional heteroscedasticity. Ann Appl Probab 10(3):1002–1024

Girola JA (2011) The HQM yield curve: basic concepts. U.S. Department of the Treasury. http://www.treasury.gov/resource-center/economic-policy/corp-bond-yield/Documents/ycp_oct2011.pdf. Accessed 10 July 2012

Grant CT, Grant GH, Ortega WR (2007) FASB's quick fix for pension accounting is only first step. Financ Anal J 63:21–35

Hann RN, Heflin F, Subramanayam KR (2007) Fair-value pension accounting. J Account Econ 44(3):328–358

Horn RA, Johnson CR, (1985) Matrix Analysis. Cambridge University Press

Jiang X (2011) The smoothing of pension expenses: a panel analysis. Rev Quant Finance Account 37:451–476

Josa-Fombellida R, Rincón-Zapatero JP (2006) Optimal investment decisions with a liability: the case of defined benefit pension plans. Insur Math Econ 39:81–98

Kesten H (1973) Random difference equations and renewal theory for products of random matrices. Acta Math 131:207–248

Kristensen D (2009) On stationarity and ergodicity of the bilinear model with applications to GARCH models. J Time Ser Anal 30(1):125–144

Leibowitz ML, Henrikssan RD (1988) Portfolio optimization within a surplus framework. Financ Anal J 44(2):43–47

Love DA, Smith PA, Wilcox DW (2011) The effect of regulation on optimal corporate pension risk. J Financ Econ 101:18–35

McGill DM, Brown KN, Haley JJ, Schieber SJ (2004) Fundamentals of private pensions, 8th ed. Oxford University Press

Mercer (2012) Mercer pension discount yield curve and index rates in the US. http://www.mercer.com/articles/1213490. Accessed on 10 July 2012

Mitra S, Hossain M (2009) Value-relevance of pension transition adjustments and other comprehensive income components in the adoption year of SFAS No. 158. Rev Quant Finance Account 33:279–301

Novy-Marx R, Rauh J (2011) Public pension promises: how big are they and what are they worth? J Financ 66(4):1211–1249

OECD (2007) Protecting pensions. Policy analysis and examples from OECD countries. Private pensions series no. 8, Organization for Economic Cooperation and Development. OECD Publishing, Paris

Owadally MI, Haberman S (1999) Pension fund dynamics and gains/losses due to random rates of investment return. N Amer Actuarial J 3(3):105–117

Owadally MI, Haberman S (2004a) Efficient gain and loss amortization and optimal funding in pension plans. N Amer Actuarial J 8(1):21–36

Owadally MI, Haberman S (2004b) The treatment of assets in pension funding. ASTIN Bull 34(2):425–433

Petersen MA (1994) Cash flow variability and firm's pension choice. A role for operating leverage. J Financ Econ 36:361–383

Pham DT (1985) Bilinear Markovian representation and bilinear models. Stoch Proc Appl 20:295–306

Pham DT (1986) The mixing property of bilinear and generalised random coefficient autoregressive models. Stoch Proc Appl 23:291–300

Pham DT (1993) Bilinear time series models. In: Tong H (ed) Dimension estimation and models. World Scientific, Singapore, pp 191–223

Picconi M (2006) The perils of pensions: does pension accounting lead investors and analysts astray? Account Rev 81:925–955

Rauh JD (2006) Investment and financing constraints: evidence from the funding of corporate pension plans. J Financ 61:33–71

Rauh JD (2009) Risk shifting versus risk management: investment policy in corporate pension plans. Rev Financ Stud 22(7):2687–2733

Rouge R, El Karoui N (2000) Pricing via utility maximization and entropy. Math Financ 10:259–276

Rudolf M, Ziemba WT (2004) Intertemporal surplus management. J Econ Dyn Control 28(5):975-990

Sharpe WF, Tint LG (1990) Liabilities—a new approach. J Portfolio Manage 16:5–10

Shivdasani A, Stefanescu I (2010) How do pensions affect corporate capital structure decisions? Rev Financ Stud 23(3):1287–1323

Sundaresan S, Zapatero F (1997) Valuation, optimal asset allocation and retirement incentives of pension plans. Rev Financ Stud 10:631–660

Tapia W (2008) Comparing aggregate investment returns in privately managed pension funds: an initial assessment. OECD working papers on insurance and private pensions no. 21, Organization for Economic Cooperation and Development. OECD Publishing, Paris

Tong H (1990) Non-linear time series. A dynamical system approach. Clarendon Press, Oxford

Towers Watson (2011) Global pension asset study. Towers Watson, New York

Yermo J, Severinson C (2010) The impact of the financial crisis on defined benefit plans and the need for counter-cyclical funding regulations. OECD working papers on finance, insurance and private pensions no. 3, Organization for Economic Cooperation and Development. OECD Publishing, Paris