Institute
and Faculty
of Actuaries

Pensions, today and tomorrow: Every risk in its right place?

The Frank Redington Prize

by Craig Turnbull

October 2022

Pensions, today and tomorrow: Every risk in its right place?

1. Introduction

What would be a sustainable and effective UK pension system for the people? This seemingly straightforward question has featured prominently in UK politics since Lloyd George first introduced a basic State pension (for over 70s only) in 1908. Politicians' ambitions and society's expectations for younger retirement ages and bigger retirement incomes increased notably after the Second World War, as reflected in William Beveridge's pension legislation of 1948. Forms of private sector pension provision in the UK pre-date the introduction of the State pension by over one hundred years. Since Beveridge, the overall affordability, efficiency and equity of pension policy, including the meshing of individual, employer and State pension elements, has never been far from the top of a Chancellor's list of political challenges.

Despite the huge increases in real wealth and incomes that have been generated by the UK economy over the last seventy or so years, a seemingly disparate set of unexpected external shocks and long-term societal trends mean that the challenges in formulating and implementing a sustainable and satisfactory pension policy are perhaps greater today than ever. Changes in population demographics and dependency ratios, a macroeconomic environment characterised by exceptionally low long-term real interest rates, society's increasing aspirations for their work / retirement lifecycle, and the increasing costs of providing medical and social care for longer in late life all have profound implications for what pension systems may be preferable, practical or even possible.

1.1 On the role of the actuary and the objectives of this essay

Actuaries have played an important role in UK pensions throughout its 200-year history and should continue to do so. However, for the actuarial contribution to be most effective, it is important to recognise the boundaries of our professional expertise. We should have a disciplined focus on the areas where we can make unique and important professional contributions. And we should recognise the domains of pension policy and practice where we should either defer to other professionals or recognise that these are not questions for professional advisors but for society and its elected representatives.

Many of the most fundamental questions around the nature of the future UK pension system must be ones where the actuarial profession remains professionally agnostic. It is not for actuaries to recommend who should bear what portion of the pension burden – whether it be the rich, the poor, this generation or the next or the one after that; nor is it for us to recommend what living standards the retired should be entitled to or aspire to. Whether State pension contributions and benefits should be universal flat-rate or means-tested; whether pension contributions should be compulsory or otherwise; whether the State pension should be advance funded or pay-as-you-go; whether the investment of any advance funding of state pensions should be controlled by the State; these are all political choices and the authority to make such choices must ultimately lie with society's elected representatives.

The actuary's professional role is a more technocratic one. Our role is to provide insightful analysis and clear advice on the financial implications that would likely follow these alternative policy choices. Many of these implications are complex and long-term, and their analysis require substantial analytical skill and judgment. The profession should seek to deliver this analysis rigorously and zealously but should also take care not to advocate specific policy choices that are ultimately merely political preferences.

The approach taken in this essay is aligned to this perspective. This essay therefore does not seek to propose or endorse radical pension policy choices that should be found in political party manifestos. Nor does it directly discuss State pension provision. It is largely focused on the analysis of voluntary funded pension provision by individuals in conjunction with their employer, and on the application of actuarial reasoning to identify where it may be made to work better. A recurring theme in this essay is the question of whether our current funded pensions systems result in the risks that arise in any form of advance funding of pensions – especially financial market risk and longevity risk – being held in the places that are most conducive to the sustainable, equitable and efficient delivery of those pensions.

1.2 The Shifting Goalposts

The UK's Gross Domestic Product has more than doubled in real terms since 1980[1]. But several factors have conspired to complicate the provision of future pensions, despite this substantial increase in the overall wealth of society. Remarkable rates of improvement in longevity over the last several decades mean longer expected lifetimes and longer expected retirements. Since 1980, life expectancy at birth has increased by some 9 years for females and 7 years for males[2]. Some of this elongation of life will not only be spent being necessarily economically unproductive but may also entail incurring increasingly substantial costs relating to the provision of social care and medical treatment[3].

The pension challenges that follow the success of technological developments in medicine and societal improvements in lifestyle are first order. From an actuarial perspective, however, there is arguably an even more profound change in the financial environment that completely alters the economic arithmetic of long-term pension provision: between 1981 and 2021, long-term interest rates in the UK fell from 13% to less than 2%.

The adequacy of long-term pension provision must be measured in real terms, and so it is the long-term real interest rate that is of more fundamental importance to this discussion. When UK index-linked gilts were first issued in 1981, they offered a long-term real yield in excess of 3%. Whilst it was briefly as low as 2% at some points in the 1980s and early '90s, by 1995 it was again above 3%. However, since 1997, notwithstanding a brief hiccup during the 2007/8 global financial crisis, UK long-term real rates have fallen relentlessly. This is a phenomenon that has occurred across much of the globe, but that is of little consolation to British savers or actuaries. Since the global financial crisis, the UK long-term real interest rate has behaved in a way that would have been regarded as quite extraordinary by most actuaries and economists twenty years or so ago. Indeed, the standard actuarial model used for stochastic analysis of Defined Benefit pension funds at the start of the 21st century assumed it was not possible for long-term real yields to go below zero[4]. In the UK, the long-term real rate has been below zero since 2010, and at the time of writing it was around -2%.

Negative real interest rates are not an unprecedented phenomenon. For example, short-term real interest rates were negative in the mid-1970s whilst inflation was out of control. Long-term real

---

[1] Gross Domestic Product: chained volume measures: Seasonally adjusted £m - Office for National Statistics (ons.gov.uk)

[2] National life tables – life expectancy in the UK - Office for National Statistics (ons.gov.uk)

[3] See Chapter 4 of Goodhart & Pradhan (2020) for an excellent discussion of the economic implications of the increasing costs of care in retirement.

[4] See Section 9 of Wilkie (1995). Some other models in use at the time did better - for example, in the late '90s the Smith model implied a probability of over 5% that the long-term real rate would be less than -2% in 2021 (see Smith (2021)); Hibbert, Mowbray and Turnbull (2001) implied this probability was around 1%. Nonetheless, the point remains that the behaviour of long-term real rates over the last decade has been a tail event relative to our understanding of real rates at the start of the century.

interest rates were not observable at this time, as the first index-linked gilts were not issued until a few years later. But the level of long-term government bond yields at this time suggests they reflected long-term inflation expectations in such a way as to offer a positive expected real return (and perhaps an inflation risk premium too).

The situation today is different. The UK long-term real interest rate, as observed in index-linked gilt prices, has been negative for more than a decade. There are many potential contributors to this situation. Perhaps expectations for productivity growth have significantly fallen; perhaps the long-term real rate has been affected by globalisation and a long-term savings glut in Asia; perhaps central banks' unprecedented quantitative easing programs have significantly elevated asset prices. The relative importance of factors such as these and many others is a subject of ongoing professional debate amongst economists. From the actuarial perspective, and in the context of the sustainable and effective provision of pensions, speculation about what might have caused these unprecedented falls in the long-term real interest rate to arise is of secondary concern (unless they were to provide a very strong indication of the likely permanence or otherwise of their effect). The more pressing actuarial point is that a sustained shift in future expected long-term real interest rates from a previous norm of +3.5% to its current level of c. -2% has a quite transformative impact on the cost of funding long-term pension provision. Now, the celebrated power of long-term compound interest is working in reverse[5].

Little is more important to the actuarial analysis of pension provision than the long-term real interest rate, and in recent years it has moved to levels that are literally off the charts. Whilst we may hope to find ourselves once again in a +3% real interest rate world, we must continue to adapt to this new environment and the consequences that will follow if it persists. There are unlikely to be easy fixes. The cost of funding a pension will be greater, more risks will be taken in doing so, and there will therefore be a need for more risk management, especially by individuals. The pension system is in a state of flux, a transitional phase as it re-orientates itself to this changed environment. We must address the legacy problem of under-funded benefits that were accrued in a different era, whilst also improving the system of pension funding that is designed for today's economic reality. These two subject areas are considered in turn in Sections 2 and 3 below.

2. The legacy of Defined Benefit pension fund liabilities

The Defined Benefit (DB) pension scheme, where benefits are advance funded by a mixture of employer and employee contributions based on a percentage of salary, paid over the period of service of the employee and invested in a trust fund, with benefits defined with reference to the salary history and years of service of the employee, has a long history in the UK. It emerged as an important component of worker compensation in the second half of the nineteenth century. By the second half of the twentieth century, a significant proportion of British workers were accruing pension benefits through participation in a DB pension scheme.

This advance funding approach, however, has not immunised pension funds' financial position from the falls in the long-term real interest rate that have occurred over the last twenty-five years. Broadly speaking, there are three reasons for this: the contributions were generally not invested in matching assets (and whilst the contributions were invested in risky assets that have performed well, their returns have not kept pace with the extraordinary returns generated by long-term index-linked bonds as a result of the precipitous falls in their yield); as real interest rates fell from the late '90s onwards, the amount of contributions paid to fund these increasingly costly accruals did not respond in accordance with the increase in the market cost of the accrual; and finally, increases in longevity

---

[5] A simple illustration: a worker pays x% of their income every year from age 25 into a pension fund that earns a real return equal to the long-term real interest rate. The worker wishes to purchase an annuity at age 65 that will pay a real income of half their pre-retirement income for 20 years (we assume zero real income growth throughout the worker's life). With a 3.5% real interest rate, x is less than 9%. When the assumed real interest rate is changed from +3.5% to -2%, x increases from less than 9% to around 50%.

expectations further contributed to the cost of ongoing accrual and the cost of already accrued benefits.

Most remaining UK DB pension schemes are now closed to new members and around half are also closed to future accrual by existing members. As of 2021, almost 10 million UK workers or retirees remain members of a DB pension scheme, and almost 1 million are still accruing new benefits (albeit in most cases at an accrual rate that has been reduced to better reflect the economic cost of the benefits in today's economic environment)[6].

2.1 Deficits

The result of these historical investment mismatches, contribution underpayments and longevity surprises is that the aggregate position of UK DB pension funds is a buy-out deficit of considerable scale[7]. The total assets of the DB schemes in the Pension Protection Fund's (PPF's) eligible universe in 2021 was £1,721bn. The aggregate buy-out cost of their liabilities was £2,336bn, implying a deficit of £615bn[8].

This aggregate deficit is a notably large number. It may be argued, however, that buy-out deficits are of little importance, and that DB pension funds' contribution plans and investment strategies were not designed to deliver buy-out solvency – instead, they were designed to produce asset proceeds that would be sufficient to meet the liability cashflows as they fell due. It is true that there has never been unambiguous clarity within the British actuarial profession on how to determine the amount required to adequately advance fund accrued DB liabilities. This has been the subject of professional actuarial debate since the nineteenth century. But it should also be noted that professional actuarial orthodoxy has not always chosen the convenient perspective that buy-out valuations simply do not matter very much. As recently as the 1990s, leading pension actuaries were arguing that DB pension funds should have funding levels that are not only sufficient to meet the cost of buy-out but that also include a significant additional buffer to allow for the risk that buy-out costs may increase in the future[9]. The argument for this level of advance funding is simple: it is the level that ensures accrued pension benefits can be secured without further reliance on the (credit-risky) sponsor.

Those sceptical of the relevance of the buy-out cost are certainly right to assert that a pension fund with a current buy-out deficit need not necessarily conclude with insufficient asset proceeds to meet liability outgo as it falls due. The investment mismatches that resulted in the market values of pension funds' assets (partly) missing the extraordinary increase in value of long-term index-linked gilts during the fall in real rates could have the opposite effect if real rates revert to their historical norms (though DB pension funds today are more closely matched to bond yields than they were ten or twenty years ago[10]). Moreover, ongoing deficit-funding contributions by pension fund sponsors could plug the gap (though these contributions totalled £12bn in 2021 and their annual size is forecast to fall to less than £1bn by 2031[11], so could plausibly only cover a fraction of the £615bn buy-out deficit). Another source of surplus on a buy-out basis arises when scheme members elect to receive pension transfer values below their benefits' buy-out cost.

---

[6] Pension Protection Fund (2021), Chapter 3.

[7] 'Buy-out' is the transfer of the pension liabilities to a third-party, and a liability valuation on a buy-out basis is an estimate of the market cost of this transfer. (A 'buy-in' is where the pension fund purchases an insurance policy that covers a specified set of the pension liabilities.)

[8] Pension Protection Fund (2021), Chapter 4.

[9] Thornton and Wilson (1992). Paul Thornton went on to be President of the Institute of Actuaries from 1998-2000. See p. 254-272 of Turnbull (2017) for further discussion.

[10] Pension Protection Fund (2021), Figure 7.2.

[11] Pension Protection Fund (2021), p. 5.

Some combination of these factors could result in the entirety of the £615bn buy-out deficit being recovered, and the legacy of DB pension funds could thereby be put to rest without further loss. But it is inevitable that some of these pension funds' sponsors will fail whilst their pension fund's assets are insufficient to meet the buy-out cost, resulting in the scheme's assets and liabilities being transferred to the PPF. For these pension funds, the cost of the deficit will be borne in part by the pension fund members, many of whom will suffer some reduction in their benefits relative to what was promised to them. To the extent these benefit reductions are still not sufficient to fund the shortfall, the PPF levy on remaining firms is intended to plug the gap. Presumably, if that proves insufficient, the government may step in to fund the remainder, or there will be some reduction in the future level of benefits provided by the PPF.

2.2 Endgames

A DB pension scheme that is closed to new members and future accruals may nonetheless have a schedule of future pension payments that extends several decades into the future. It is natural that economies of scale may incentivise closed pension funds to outsource elements of the management of their liability run-off. Moreover, with the pension scheme no longer playing an ongoing strategic role in the delivery of future employee compensation, the sponsoring employer may be attracted to transactional options that allow their ongoing liability for accrued pension benefits to be crystallised and closed out.

The demand for insurance buy-out solutions

Currently, the main route for a pension fund seeking to transfer at least part of its DB pension fund liability is through an insurance buy-out (or buy-in, which is usually a steppingstone to buy-out). Some £200bn of DB pension liabilities have been subject to buy-out or buy-in between 2007 and 2021[12]. This is a relatively small fraction of the total universe of UK DB pension liabilities, and many DB pension schemes are not currently in the financial position to afford a buy-out transaction. But buy-out (or some similar form of transfer of liabilities to a third party) is now the exit strategy of choice for many private sector DB pension schemes.

The juxtaposition of the growth in pension buy-out business alongside the decline in individual annuity business is an empirical curiosity. Today, only a small proportion of retiring Defined Contribution (DC) pension fund members opt to invest some of their pension fund in a life annuity. And even fewer DC pension fund members in their accumulation phase opt to invest in a deferred annuity. There could be many reasonable grounds for this apparent lack of DC members' interest in annuity products: inheritance planning, investment and / or longevity risk appetite, current market pricing and a view that these products are not good value for money, etc.

But when the deferred members of a DB pension fund are given the choice between receiving a cash transfer value that can be invested in a DC plan or transferring their accrued deferred annuity benefit to a buy-out provider, it is perhaps surprising that so many of them opt for the latter when it is a kind of benefit that DC members appear to find so unappealing. The costs and complications involved in transferring from DB to DC schemes (including the costs of necessary financial advice) may help to explain this apparent anomaly, and the lack of economic equivalence (at current interest rates) in the tax treatment of DB and DC pensions, particularly in respect of the Lifetime Allowance, may be another important disincentive for DB members considering transferring to DC. Nonetheless, the consequent difference in retirement solution that results from the free choices of these two groups is quite striking. It is not obvious why these two groups of pension savers should have the distinctly

---

[12] Pension Protection Fund (2021), p. 30.

different risk preferences that this implies. The alternative explanation is that one of these groups of pension savers is inadvertently far away from a solution that is aligned to their preferences.

This situation is, of course, different for DB pension fund members whose pension is in payment. As any actuarial student knows, a transfer value option generally cannot be offered to a life annuitant, as it creates a strong longevity anti-selection effect (all pensioners would take their transfer value on their deathbed). And DB pension schemes using buy-out have often focused their initial buy-out on the in-payment segment of their membership.

Whilst retired DB pension fund members may be necessarily attached to the longevity risk insurance that comes with their in-payment life annuity income, it is less clear why they must remain attached to a retirement income solution that has zero investment risk. As a thought experiment, suppose buy-out firms provided their in-payment policyholders with the option of taking a unit-linked version of their annuity, with a similar investment mix to that which the buy-out firms use to back those annuities. The investment risk appetite implied by DC pension preferences suggests such a choice could be of interest to many buy-out policyholders.

The supply of insurance buy-out solutions

The increase in pension scheme demand for buy-out solutions has become a major source of new business for the UK life industry. Indeed, with the gradual but seemingly inevitable demise of with-profits business since the start of the century and the end of compulsory annuitisation in 2015, the buy-out of legacy DB pension funds has become the most significant source of new long-term liabilities for the UK life sector.

It may be interesting to reflect on how different a buy-out-focused UK life business is to the industry’s historical norms[13]. Very briefly, let’s start at the historical beginning of the UK life industry. Around 1800, UK life offices wrote a mixture of with-profit and non-profit business. The non-profit business was priced to make an expected profit for with-profit policyholders but came with attendant risk – if the non-profit business made a loss, it was the with-profit policyholders who bore this loss, and as members of a mutual company with unlimited liability, their exposure was meaningful. The non-profit business was backed with high quality assets, and the profits were expected to accrue primarily from mortality / longevity margins rather than investment risk-taking.

Moving on to the mid-twentieth century, assets backing with-profit business were increasingly invested in risky assets such as equities and property. The basic idea was that those risky assets were not backing guaranteed liabilities, but the variable element of with-profit pay-outs (in particular, the ‘terminal bonus’ element that was introduced in the 1950s). By the late twentieth century, with-profit funds’ risky asset allocations tended to exceed the amount that was set aside for variable benefits and so included some of the assets that backed guaranteed benefits. This resulted in increasingly volatile balance sheets. By the early years of the twenty first century, equity market volatility, together with the low interest rate-induced losses that resulted from guaranteed annuity options, helped to put the entire 200-year-old UK with-profit system into run-off.

In today’s buy-out sector, there are no with-profit policyholders to bear the losses in the long-term non-profit business. And these fixed liabilities tend to be backed by asset portfolios that are riskier than the historical norm for these types of long-term fixed liabilities. Business innovation is, of course, a vital element of any healthy and vibrant industry. But it is perhaps under-appreciated how much of a departure this buy-out-focused business strategy is from the historical norms of the UK life sector.

---

[13] For more on the early historical development of UK life business and the actuarial thinking behind its financial management, see Turnbull (2017), Chapter 2.

In the absence of the buffers provided by the variable benefits of with-profit policyholders, the risk of losses on this business is entirely supported by loss-absorbing capital, and the minimum amount of the capital that must be held is set by regulatory capital requirements. Assessing the appropriate level of capital required to provide a high level of security for long-term guarantees backed by risky assets is a complex and necessarily judgment-laden process. The non-stationarity and inherent uncertainty of our economic world makes this inevitable.

This leads to a rather costly game between the insurance firms and their regulator. The shareholders' equity position in the buy-out balance sheet is, in economic terms, a call option on the risky asset portfolio, with a strike price equal to the benefits promised to the pension fund members. The cost of the call option to the shareholders is the amount of capital they need to provide to support the business. The prospective value of the call option increases with the riskiness of the asset portfolio. Firms have a natural economic incentive to increase the value of this option by maximising risk-taking relative to capital requirements, and it is the basic purpose of prudential solvency regulation to keep capital requirements aligned to risk-taking and such that policyholder protection is not diminished by the firm's investment choices. A considerable pool of costly expertise is consumed in this contest. Whilst this certainly generates substantial demand for actuarial services (including my own!), it is not self-evident that it is a necessary component of the most efficient form of pension provision, particularly when DC investment choices suggest there is a preference for simple, transparent solutions with direct exposures to risky asset returns.

It may be argued that buy-out firms' investment risk-taking is in the annuitants' interests, as part of the expected asset return associated with these strategies will be passed through to the policyholder (or the DB pension fund sponsor) in the pricing of the product. And it may also be argued that the long-term risks associated with these asset strategies is really very low, and very little additional capital is required to ensure it is extremely likely that the proceeds of these asset strategies will be sufficient to pay the members' benefits as they fall due. Whilst this is indeed a strong argument for backing guaranteed liabilities with such assets, it is perhaps an even stronger argument for offering pension fund members a form of retirement income product that has a direct participation in these asset returns.

As in the above discussion of the demand for buy-out solutions, this discussion of their supply again points to the potential usefulness of some form of investment-linked annuity as a simpler, cheaper solution that provides the former DB pension fund member with a level of direct investment risk participation that is more commensurate with the risk appetite preferences implied by DC members' investment choices.

Superfunds

This discussion of the resolution of legacy DB pension liabilities would not be complete without mention of superfunds, which are emerging as a new alternative exit route for DB pension funds. Like buy-out insurers, a superfund would take over the full obligation for meeting a specified block of the outstanding liabilities of the pension scheme. A superfund is not an insurance company and therefore is not subject to the prudential solvency capital requirements of the insurance sector. However, superfunds are regulated by the UK's Pension Regulator (TPR), and this regulator published guidance in June 2020 that states superfunds must hold a risk-based capital requirement and 'the overarching objective of our capital requirements is for there to be a very high probability of members' benefits being paid in full'[14]

Despite this very-high-probability ambition, TPR expects pension funds to use an insurance buy-out rather than a superfund if the pension scheme can afford the buy-out or is likely to be able to afford it

---

[14] Guidance for DB superfunds | The Pensions Regulator

soon. So, a superfund will be a cheaper route to exit for the pension scheme sponsor, and it would seem counter-intuitive to suppose it does not also deliver a less secure form of benefit provision than that delivered by an insurance buy-out. This doesn't mean the superfund is delivering an inadequate or inappropriate solution: to assess that we would need to consider whether the superfund solution is likely to offer a more secure route to paying members' benefits than would be the case if the benefits were left in the DB pension scheme with whatever ongoing sponsor support was available. The answer to that question will be highly specific to a scheme's circumstances.

From the members' perspective, the presence of the superfund route creates greater complexity. Now different routes to securing guaranteed benefits necessarily come with different levels of security. It would seem fundamentally important that the advice that supports members' decisions on whether or not to retain a guaranteed form of benefit that will be delivered by a superfund should explicitly highlight that this particular form of solution will mean their benefits are less secure than would be the case in a buy-out - and similarly less secure than the (deferred) life annuity that could be purchased with their transfer value (albeit that annuity income would be lower than their current promised benefit).

3. Risk management in Defined Contribution plans

As noted above, an economic environment with a long-term real interest rate some 500 basis points lower than its historical norm has a profound impact on the saving power of advance funding. Of course, one important lesson that can be taken from the experience of the last thirty years is that the prospects for future long-term real interest rates are highly unpredictable. We should not rule out the possibility that whatever forces have brought real rates to their current extraordinary low levels over the last couple of decades will prove transitory. But robust actuarial analysis cannot rely on this hope.

A shorter retirement period, delivered through a later retirement age, is an inevitable economic consequence of prolonged negative long-term real interest rates. Greater flexibility in today's approaches to working life may create more options for how older workers choose to participate in the workforce in later years. For many, particularly those in non-manual occupations, this may mean the transition from working life to retirement is not a binary step, but a gradual change in the volume and nature of their work over many years.

The future transition to retirement is therefore likely to be later and longer than has been the norm in recent decades. There is also likely to be more uncertainty around the timing of the start of the retirement transition. This uncertainty is a natural consequence of the investment risk appetite that DC pension plan members tend to exhibit in their accumulation phase. A 40-year-old worker with the objective of retiring when their DC pension pot accumulates to a value that can generate a sustainable real income in retirement of, say, 40%, of their full-time working income, may plausibly find themselves able to retire in 25 years or 35 years, depending on the paths of future investment returns and real interest rates.

For the reasons above, the notion of a 'normal retirement age' is likely to become an increasingly outdated concept. Instead, we will have a blurred transition from pension accumulation to income. Individuals' management of their pension risk, in both the accumulation and income phases, is now of fundamental importance to pension outcomes. This section focuses on a couple of key aspects of this risk management that are notable for both the scale of their importance and their plausible potential for significant improvement relative to current practices: the management of longevity risk (section 3.1); and the management of investment risk (section 3.2).

3.1 Managing longevity risk in the retirement income / asset decumulation phase

The relaxation in 2015 of the requirement for a significant portion of DC pension pots to be invested in a life annuity at the time of retirement grants today's DC retirees a great deal of freedom (and responsibility) to determine how much longevity risk they will retain during retirement. As noted in the Introduction, an individual's longevity risk appetite may be a function of a range of factors. Whilst there is a natural incentive to avoid the significant costs that may be associated with transferring their longevity risk to a third party, the prospect of the pension pot being exhausted at a time when many years of life remain ahead should be deeply unattractive to most retirees. It seems highly counter-intuitive that individuals should prefer to bear a risk that is so consequential and that is also largely diversifiable.

DC retirees can, of course, fully transfer their longevity risk to an insurer with the purchase of a life annuity. Individuals may choose to invest at least a part of their retirement wealth in this way, though relatively few currently do. A slightly more sophisticated longevity risk management strategy could involve the use of a deferred annuity that is purchased at retirement and that would start to pay at, say, age 85 or even 90. This asset would act as an out-the-money longevity put option, allowing new retirees to cost-effectively manage their longevity 'tail' risk. However, it is less clear that life assurers are willing to write such products. Deferred annuities have greater longevity risk content than immediate life annuities (relative to premium), and this can make them more capital-intensive.

Taking a step back, there might appear to be something counter-intuitive about this situation – why does the pooling of largely diversifiable longevity risk result in substantial capital requirements for insurers? The answer, of course, is that only a part of longevity risk is diversifiable, and conventional annuities transfer both the diversifiable and non-diversifiable elements of longevity risk from the individual to the life assurer.

What do we mean by diversifiable and non-diversifiable longevity risk? Suppose we have a pool of 1000 newly retired individuals all age 65. Let's suppose these are homogenous lives and the current expectation is that they will live for another 20 years. Developments in medicine, changes in social care provision, unexpected environmental developments, changes in social attitudes to food, alcohol and exercise, the appearance of new pandemics, wars, etc. may mean that the aggregate mortality experience for large pools of lives such as these 1000 people is ultimately not what was expected when they were age 65. The ultimate average years lived by this pool of lives from their retirement at age 65 may turn out to be 17 years or 23 years instead of the 20 years that was expected at the time of their retirement. This is a non-diversifiable risk for the insurer in the sense that the risk does not diversify away as more of it is pooled. This is the form of risk that life assurers hold longevity risk capital to support.

There is a second form of longevity risk: suppose societal trends turn out, miraculously, to be exactly as expected, and the average years lived by the 1000 lives is indeed 20 years. Amongst this pool, some may have died after 1 year because of any number of misfortunes; equally, some may have lived for 40 years to the ripe old age of 105. This risk is diversifiable – the law of large numbers ensures that pooling this risk will lead to a more stable aggregate average outcome.

A key point here is that, for the individual, the diversifiable element of longevity risk is much bigger than the non-diversifiable element. The diversifiable part is therefore the part of longevity risk that individuals ought to be most keen to transfer. And this is the element of longevity risk that insurers can efficiently diversify. Indeed, this is the very point of insurance – to facilitate the pooling of diversifiable risk. Whenever insurers stray from this principle and seek to place intrinsically non-diversifiable risk on their balance sheets, unstable outcomes will inevitably follow, and these will tend to require lots of up-front capital.

This points to a form of longevity risk solution for DC retirees where the pension income varies with the longevity experience of the pool. This variability will be only a fraction of the individual's own longevity risk exposure. Such a solution would not place any non-diversifiable risk on the insurer's balance sheet, and it would therefore be a very capital-light, and hence cheaper, product. Such a product could take many forms, and it is not the aim of this essay to identify or advocate a specific product design. Longevity risk product innovation has been surprisingly uneventful in the years that have followed the removal of compulsory annuitisation. However, there is a growing actuarial academic literature[15] that examines the feasibility of some forms of longevity risk sharing that are aligned to the above observation that a solution that pools diversifiable risk whilst returning non-diversifiable risk to the retiree is the likely route to more cost-effective longevity risk solutions.

The discussion in Section 2 developed the argument that a form of investment-linked annuity could be a more capital and cost-efficient retirement solution than an existing life annuity, and that this could also be better aligned to retirees' investment risk preferences. This section has developed this argument a stage further. Not only is it efficient to allow investment risk content into the retirement income solution, but it is also more efficient to include some longevity risk in the retirement solution - specifically, the element of longevity risk that is non-diversifiable.

3.2 Managing investment risk: can intergenerational investment risk pooling deliver?

DC asset allocations imply that most individuals have substantial appetite for investment risk during the advance funding period, and perhaps also in the retirement income / asset decumulation phase. This investment risk appetite is quite intuitive given the size of asset risk premia that have been empirically observed in long-term investment in equity and property assets. It is also likely that investment risk appetites have been amplified by the falls in long-term real interest rates, as individuals seek to maintain expected real investments returns at something closer to their historical norms.

The presence of this considerable investment risk appetite means the risk management of individuals' DC portfolios is crucially important to obtaining satisfactory pension outcomes. This risk management challenge is likely to gain an increasing profile in future years as DC funds continue to accumulate and risky portfolio values experience the bouts of severe volatility that periodically visit the financial markets. In this setting, a tempting but contentious idea has emerged: is it plausible that an investment vehicle can 'smooth out' risky investment returns over time, thereby delivering all the risk premia to the investor whilst exposing them to only a fraction of the risk?

This idea is often referred to as intergenerational risk sharing and, in the UK DC world, early implementations of it are being piloted in Collective Defined Contribution (CDC) schemes. To some, intergenerational risk sharing is a dubious and unscientific idea akin to alchemy[16]. Others have argued that the idea is little more than good common sense and foresee a major role for it in long-term DC pension accumulation[17]. Such debates can easily become polarised and result in more heat than light. The following discussion does not argue for or against the idea per se: rather, we seek to clarify what technical conditions are logically necessary for intergenerational risk pooling to be equitable and sustainable.

Before we proceed with this analysis, we may note that the ambition to deliver smoothed risky investment returns is not a new one for long-term UK savings institutions. Throughout the second half

---

[15] For example, see the recent seminar by Professor Catherine Donnelly: ARC Webinar series 2022: The next generation of CDC pensions? - YouTube; and Bernhardt & Donnelly (2021).

[16] See Ralfe (2021)

[17] See IFoA (2019) and IFoA (2021).

of the twentieth century, UK life assurers' with-profit funds provided participation in long-term investment risk premia (primarily in equities and property), along with some forms of underlying guarantee, together with some element of 'smoothing' of investment risk through intergenerational risk pooling. The structure of some CDC plans, where a prudently calibrated initial level of pension is increased over time as investment returns are earned, is reminiscent of a with-profit product design.

UK with-profit funds are now in near universal terminal decline and managed run-off. For DC investors, some new asset management solutions have emerged to support individuals' financial risk management of their DC pension pots: lifestyling asset strategies and target-date funds are popular examples. These automated asset allocation strategies reduce investment risk over time as the individual approaches their intended retirement date. Such approaches have some intuitive appeal – a reduction in investment risk as the intended retirement date draws closer can provide the investor with a more stable expectation for their ultimate retirement pot, helping to avoid unexpected last-minute bumps on their planned path to retirement.

With these asset management solutions, the reduction in asset risk is necessarily accompanied by a reduction in the participation in asset risk premia. The appeal of the intergenerational risk pooling concept is that it may achieve a similar reduction in short-term return volatility without the need for a reduction in participation in risk premia. This risk pooling idea is alluring but it might sound too good to be true. Can the risk in asset returns really be reduced without reducing the risk premium? After all, the textbooks say financial market risk is inherently non-diversifiable – so what good can risk pooling do? More specifically: what characteristics of financial market risk are necessary for this risk pooling to deliver in a sustainable way?

Pooling investment risk over time: the characteristics of market risk necessary for success

The essential characteristic of investment risk that is necessary for viable intergenerational risk transfer lies in the 'time structure' of risky excess asset returns[18]. For intergenerational risk transfer to be sustainable, equitable and meaningful, some element of short-term market returns must be transient or temporary. Put another way, *expected excess returns on risky assets must materially change over time in a discernible way.*

In the absence of such a characteristic, intergenerational investment risk pooling cannot accomplish the objective of providing a stable participation in asset risk premia (expected return) with less risk. Suppose the expected market return on risky assets is constant over time. Then intergenerational risk pooling would merely imply robbing Peter to pay Paul. If cohort A experiences a poor market return, and their return is then 'topped up' through a cross-subsidy that must be funded from a future cohort B, then cohort B's expected return is necessarily reduced below the (constant) expected market return. B's future participation in the asset risk premium has therefore been reduced before they have even invested their money.

It is difficult to reconcile this effect with the principle of intergenerational fairness – why does A deserve this extra return more than B? It might be argued it is fair because this is what cohort B has signed up to do in these circumstances. But cohort B is a new cohort investing new money. They may not have signed up for anything at the point at which cohort A's investments have been exposed to the poor market returns. It might also be argued that that the nature of long-term pension saving - investing a series of regular contributions – means that there is a large overlap between cohort A and cohort B, and so the scale of the transfers between A and B may be immaterial. But, in the absence of any external third-party, any smoothing benefit to cohort A must result in an offsetting cost to other cohorts. The 'over-lapping' argument is only effective if it means there is no material smoothing

---

[18] Throughout the remainder of this essay, we use 'risky asst return' to refer to a risky asset's excess return, i.e. the risky asset return in excess of the risk-free rate.

impact for any cohort. And if this is the case, surely it would be preferable to avoid the cost and complexity associated with implementing it.

In a world where expected market returns are stable over time, the smoothing mechanism will result in poor experienced returns being passed around cohorts of investors like a hot potato, but they will never find a natural home. Any reduction in risk that is created by the smoothing of returns must necessarily result in a new source of volatility: the smoothing creates a new variation in expected returns for the new money that will be invested in the scheme in the future. We have simply swapped one form of volatility for another, and it is not obvious what has been gained other than complexity.

However, if the poor market return that has been experienced by cohort A in the above example implies an increase in expected market return for cohort B, then such a cross-subsidy would seem much more reasonable and fair - B's good fortune at buying 'cheap' assets provides the rationale for them subsidising the poor outcome that arises for A when they must sell them. In this case we can foresee a system of pooling that removes the market's intertemporal variation in short-term expected returns and offers an expected return to all participants at any point in time that is equal to the long-term expected market return. Such a system would arguably be delivering equitable (ex ante though not ex post) outcomes for all generations – all investors have the same stable expected participation in the (now volatile) asset risk premium, whilst risk has been reduced for all by the scheme's approach of 'looking through' excess short-term market price volatility.

Evidence for intertemporal variation in expected risky asset returns

This essay will not attempt to offer a definitive assessment of the empirical evidence in support of the presence of material intertemporal variation in expected risky asset returns. But the good news for advocates of intergenerational risk pooling of market risk is that there is a strong prima facie case for it. Substantial research on the empirical behaviour of long-term risky asset returns and the extent to which they have exhibited long-term mean-reversion has been produced by academic economists since the early 1980s. Mean-reversion in realised returns implies intertemporal variation in expected returns – if an asset price is pulled back towards some 'correct' level whenever it departs from it, then the expected return on the asset is inevitably altered by this correction mechanism[19]. Several classic papers were published on the long-term empirical behaviour of risky asset returns in the 1980s by economists such as Shiller[20], Fama & French[21] and Roll[22]. Fama and Shiller were jointly awarded (together with Hansen) the 2013 Nobel Prize in economics for their work in this field.

Today, there is some consensus that risky asset classes such as equities have exhibited significant mean-reversion in excess returns when analysed over periods of years. Basically, market prices appear to be more volatile than can be explained by the volatility of their underlying income streams. Changes in prices appear to be greater than those required to reflect rational changes in expected future income. This means that falls in market prices tend to predict higher future expected returns and vice versa. As the leading US economics professor John Cochrane put it: "Historically we find that virtually all variation in price / dividend ratios has reflected varying excess expected returns."[23]

Naturally, when dealing with the long-term and empirical, caveats must be attached. The inherent non-stationarity of our economic system may mean that the past is simply not a reliable guide to the future. Moreover, there is quite limited historical data from which to base these inferences about long-term behaviour. Quoting Cochrane again, he notes: "The slow movement of the price / dividend ratio

---

[19] Excess volatility is exactly the same thing as return predictability.", Cochrane (2005), p. 396.

[20] Shiller (1981)

[21] Fama and French (1988a) and Fama and French (1988b)

[22] Roll (1988)

[23] Cochrane (2005), p. 397.

means that on a purely statistical basis, return forecastability is an open question. What we really know is that low prices relative to dividends and earnings in the 1950s preceded the boom market of the early 1960s; that the high price / dividend ratios of the mid-1960s preceded the poor returns of the 1970s; that the low prices of the mid-1970s preceded the current boom. We really have three post-war data points: a once-per-generation change in expected returns."[24]

In today's world of big data analysis, three data points may sound like a rather modest foundation for reliable inference. But it's essentially all we have. The following analysis will proceed from the premise that risky assets do indeed exhibit intertemporal variation in expected returns, but we must always bear in mind that we cannot 'prove' that future returns will indeed exhibit this characteristic.

Modelling intertemporal variation in expected risky asset returns

Intertemporal variation in expected risky asset returns can be intuitively conceptualised by a model that supposes the asset's market price can be de-composed into two components: a 'permanent' component, which reflects changes in the expected stream of future cashflows generated by the asset, and could be supposed to follow a random walk; and a 'temporary' component, which can be characterised in various ways (rational changes in risk premia; inexplicable excess short-term market price volatility), and which mean-reverts back to zero over time.

This type of model and its empirical calibration has been extensively analysed by academic economists[25]. It also fits with some 'traditional' actuarial conceptions of how to measure and manage market price risk. For example:

- The Wilkie model[26], the most influential stochastic asset-liability for UK actuarial use in the 1980s and '90s, produced material intertemporal variation in the expected returns of equities and property that could be characterised as a model of permanent and temporary components of asset prices;
- • A traditional actuarial perspective on the equitable distribution of surplus generated by equity and property returns in with-profit funds would view the component of price return arising from increases in asset income as more fit for immediate distribution than the component of return arising from a change in income yield[27];
- • The historical treatment of asset valuations in DB pension fund valuations, where expected future equity income was discounted using a stable off-market actuarial assumption, could be viewed as an attempt to include only the permanent component of the asset value in the pension fund valuation[28]

From the perspective of intergenerational risk pooling, the key point is that it may be possible to use an intergenerational risk pooling mechanism to 'correct' the temporary component of market price changes immediately rather than waiting for the market to eventually do it at some point in the future. Put another way, the system would 'look through' excess short-term return volatility by only allocating the investor the return on the permanent component of the asset price. In this way, the investor would participate fully in the asset risk premia, whilst avoiding exposure to the short-term volatility generated

---

[24] Cochrane (2005), p. 395.

[25] Poterba and Summers (1988)

[26] Wilkie (1986) and Wilkie (1995).

[27] Turnbull (2017), p. 191-2.

[28] See Turnbull (2017), p. 240-247.

by changes in the temporary component of the asset value. These ideas are quantitatively explored in the technical modelling appendix.

Delivering investment risk pooling over time in the presence of intertemporal variation in expected returns: some immediate questions

The intergenerational risk pooling idea is alluring and could play a promising role in reducing short-term investment risk associated with DC investing. But it is crucial to recognise that its delivery is not some form of actuarial magic trick. It is only viable (i.e. sustainable and equitable) if the risks that we seek to pool – long-term risky investment returns – have some quite specific statistical characteristics. Many questions immediately follow that are important to explore in depth in an assessment of the fundamental viability of an intergenerational risk pooling system. Below we provide a brief outline of some of these questions and where the answers might be found.

How confident can we be that risky asset classes such as equities will exhibit intertemporal variation in expected returns in the future?

As noted above, it is widely accepted that there is good empirical support for the hypothesis that risky assets generate significant intertemporal variation in expected excess returns over the long term[29].But this empirical support is necessarily derived from a small sample of data. And, even with plenty of historical data, we should not forget the fundamental epistemic issue that the non-stationarity and inherent uncertainty in our economic future necessarily renders any inference about the future from the past a fallible one. Any intergenerational risk pooling system must therefore recognise the possibility that historical price time series will not turn out to be a reliable basis for inference about future price behaviour.

Following a period of poor returns that implies markets now have elevated expected returns, surely any investor would find it preferable to invest new funds in an 'un-smoothed' strategy rather than in a strategy that subsidises the investment losses of others' previous investments?

How does an intergenerational risk pooling scheme attract new funds in the aftermath of a period of very poor returns? If such returns have arisen from a negative experience in the temporary component of risky asset prices, then the logic of mean-reversion dictates that the expected returns on the market will, at this point, be high. Why would a rational investor opt to invest their new money such that this prospective excellent return will be partly used to shore up the losses of others' previous investments? This question applies even in the presence of return mean-reversion – whilst the smoothing mechanism can stabilise investor's participation in the market's time-varying asset risk premia, this may be unenticing during times when market prices imply risk premia are elevated.

It is not easy to find a reason why someone would choose to invest in the smoothed scheme in these circumstances, other than that they are compelled to do so by some form of mandatory system. Ultimately, the case for compulsory investment in a specific investment vehicle can only be decided by government. As actuaries we must highlight that the viability of intergenerational risk pooling requires future generations to be there when earlier generations need their support. Uncertainty in the participation of future generations could therefore constitute a significant vulnerability for such schemes.

If mean-reversion means long-term risky returns are more stable and predictable than they otherwise would be, why do we even need intergenerational risk pooling?

---

[29] There is less consensus amongst economists on why this asset price behaviour occurs, i.e. does the empirical price behaviour reflect irrational 'animal spirits' or rational changes in the risk premia required for holding assets through risky periods? For our actuarial purposes, a reliable causal explanation for the empirical behaviour is arguably less important.

If returns do have a temporary component, then in what sense does a long-term investor benefit from the intertemporal risk pooling mechanism? After all, they can time diversify away the temporary component by simply buying and holding for long enough. The benefit of the intertemporal risk pooling idea is a reduction in the *short-term* volatility of the return, but it does little to the long-term volatility of the return. So, if we have a 1-year investment horizon, the risk pooling could be very useful. But this essay is focused on the long-term investment horizons that arise in DC pension accumulation. The most compelling benefit in this setting is that it reduces the need for a form of lifestyling strategy towards the end of the accumulation period. Even long-term investors eventually have short-term investment horizons, and it is in these circumstances where the intertemporal risk pooling mechanism could usefully permit participation in asset risk premia for longer. Put another way, the presence of short-term excess volatility is a key driver of the need for lifestyling, and intergenerational risk pooling can provide another way of mitigating this type of volatility.

The above discussion argued intergenerational risk pooling could work by distinguishing between changes in the permanent and temporary components of assets prices. How do we identify, at any point in time, which component of a market price is permanent, and which is temporary?

This is a very real and practical challenge in the context of intergenerational risk pooling schemes. We will not find the permanent and temporary component values of the FTSE 100 in the Financial Times. A considerable amount of judgement will inevitably be required in determining how much intergenerational cross-subsidy is appropriate in any given scenario – or in determining a formula that will make this determination. This same challenge arose in the equitable distribution of equity gains / losses across different generations of with-profit policyholders.

An intuitive approach to differentiating between permanent and temporary components of risky asset prices could assume there is a long-term value that income yields revert to – and so the permanent component could be determined as the value implied by this yield, and the temporary component as the difference between this value and the market value. Of course, the challenge that immediately follows is that this long-term yield value is not a directly observable property, and we have no way of knowing if it will be stable in the future. Changes in long-term income growth expectations could rationally drive a change in the ‘neutral’ income yield level. But, if we put such concerns to one side, then this approach would identify the permanent component of the return as that arising from the experienced change in income, and the temporary component as that arising from the change in yield.

The above discussion is entirely qualitative. How big does the volatility of the temporary component of returns need to be relative to the permanent component such that the risk pooling is quantitatively meaningful?

This question is explored in a technical appendix which presents some stochastic modelling analysis.

Issues for intergenerational risk pooling beyond intertemporal variation in expected risky asset returns

The above discussion has only considered the behaviour of the excess risky asset return (i.e. the return in excess of the risk-free rate). The smoothing mechanism must also adapt to the inevitable changes in long-term interest rates and long-term inflation expectations that will arise over its time of operation, such that an equitable treatment is delivered across the different generations of participants.

It is notable that much of the second half of the twentieth century's British actuarial literature was concerned with a very similar challenge in the context of the equitable treatment of different generations of with-profit policyholders[30]. This literature is full of deep and insightful discussion of these types of issues. Much of the actuarial profession's experience and expertise in this area is now historical. A perusal of this literature does not give the impression that these challenges have very straightforward solutions, and we should not under-estimate the challenges that are created for equitable intergenerational risk pooling by the very unpredictable nature of our long-term economic future.

4. Closing Thoughts

If there is a single theme that runs through this essay's discussion of solutions for both legacy DB benefits and the ongoing provision of DC benefits, it is that sustainable, equitable and cost-effective solutions will tend to be ones where savings institutions only bear risks that they can diversify. This principle was present in the discussion of legacy DB, where it was argued that, to the extent that the member / policyholder has appetite for market risk in the funding of their pension benefits, it is likely to be more efficiently borne directly by them rather than on the balance sheet of an insurer. The principle was again present in the DC discussion, albeit with a couple of twists: longevity risk has both diversifiable and non-diversifiable elements, and whilst the diversifiable element can be efficiently pooled by an insurer, it is likely to be more efficient and cost-effective for the non-diversifiable (and smaller) element to be retained by the individual; and, finally, if some component of market risk is temporary in nature (as defined in section 3.2), then there is scope for this component of risk to be pooled, thereby reducing individual members' investment risk over short-term horizons.

The practical implications of the above arguments for pension product design, the provision of financial advice and government policy in areas such as taxation and regulation are naturally best further explored by those with specific expertise in those fields. But this essay closes with a few general assertions that naturally follow from the above.

There should be a focus on the removal of unnecessary obstacles to the transfer of accrued DB benefits into DC funds (whilst retaining robust safeguards that ensure members receive appropriate financial advice and are offered a fair transfer value), thereby allowing individuals to directly participate in the long-term asset risk premia generated by their pension assets. Inadvertent tax disincentives to transfer from DB to DC such as the lack of economic equivalence in the treatment of DB and DC benefits in the Lifetime Allowance tax calculation should be addressed. If accompanied by a reduction in the barriers to efficiently investing in long-term illiquid credit assets in DC decumulation strategies, such a focus could align well with government ambitions for more pension assets to be invested directly in the UK real economy.

DC product innovation since the removal of compulsory annuitisation in 2015 has been disappointingly uneventful. But there are some developments in both academic literature and practical

---

[30] See, for example, Kennedy et al (1976) and Redington (1981), and Turnbull (2017), pp. 188-192 for some discussion of this topic.

DC provision that align to the above arguments about where different forms of longevity risk (diversifiable and non-diversifiable) and market risk (permanent and temporary) should naturally reside. Intergenerational pooling of investment risk can find an intellectual basis in the empirical evidence for mean-reversion in the excess returns of risky assets. Conceptually, this offers a less costly way of mitigating short-term excess volatility in market returns than 'conventional' approaches such as lifestyling. But the meaningful practical challenges associated with intergenerational risk pooling should not be understated. Some form of compulsion in participation is likely to be necessary to ensure new policyholders arrive when they are needed and when it is not necessarily in their own best interest to do so. And we should not lose sight of the fact that the long-term properties of the future behaviour of financial markets that provide the technical basis for this form of risk pooling is inherently uncertain, no matter what the statistical analysis of historical returns might suggest. Any intergenerational risk pooling scheme therefore needs to be built with 'safety valves' that can be used if long-term market behaviour is of a kind that is not deemed plausible by probabilistic models calibrated to available historical data.

5. Technical Appendix: Some illustrative quantitative analysis of intergenerational risk pooling

An illustrative stochastic process for a risky asset with intertemporal variation in expected excess returns

Suppose the market value of a risky asset at time t, M(t), is:

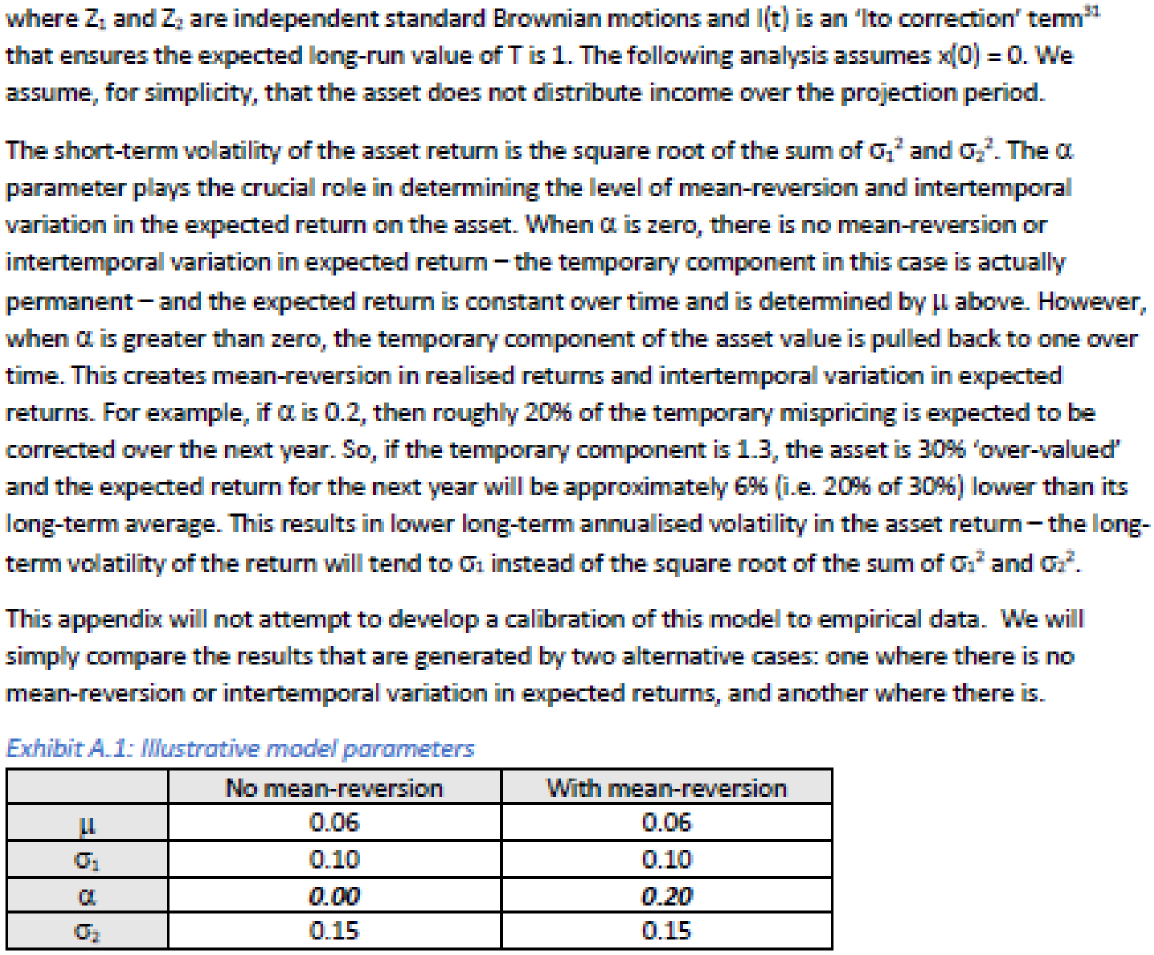
$M(t) = P(t)T(t)$

P(t) can be regarded as the permanent component of the asset price, and it is assumed to follow a geometric brownian motion (i.e. a form of random walk). T(t) can be regarded as the temporary component, and it is a stochastic process that mean-reverts to a value of 1. Specifically, we assume its logarithm follows an Ornstein-Uhlenbeck process (i.e. it is mean-reverting). The stochastic processes driving M(t) are therefore:

$$dP(t) = \mu P(t)dt + \sigma_1 P(t)dZ_1(t)$$

$$T(t) = I(t)\exp(x(t))$$

$$dx(t) = -\alpha x(t)dt + \sigma_2 dZ_2(t)$$



where Z_1 and Z_2 are independent standard Brownian motions and $I(t)$ is an 'Ito correction' term[31] that ensures the expected long-run value of T is 1. The following analysis assumes $x(0) = 0$. We assume, for simplicity, that the asset does not distribute income over the projection period.

The short-term volatility of the asset return is the square root of the sum of σ_1^2 and σ_2^2. The α parameter plays the crucial role in determining the level of mean-reversion and intertemporal variation in the expected return on the asset. When α is zero, there is no mean-reversion or intertemporal variation in expected return – the temporary component in this case is actually permanent – and the expected return is constant over time and is determined by μ above. However, when α is greater than zero, the temporary component of the asset value is pulled back to one over time. This creates mean-reversion in realised returns and intertemporal variation in expected returns. For example, if α is 0.2, then roughly 20% of the temporary mispricing is expected to be corrected over the next year. So, if the temporary component is 1.3, the asset is 30% 'over-valued' and the expected return for the next year will be approximately 6% (i.e. 20% of 30%) lower than its long-term average. This results in lower long-term annualised volatility in the asset return – the long-term volatility of the return will tend to σ_1 instead of the square root of the sum of σ_1^2 and σ_2^2.

This appendix will not attempt to develop a calibration of this model to empirical data. We will simply compare the results that are generated by two alternative cases: one where there is no mean-reversion or intertemporal variation in expected returns, and another where there is.

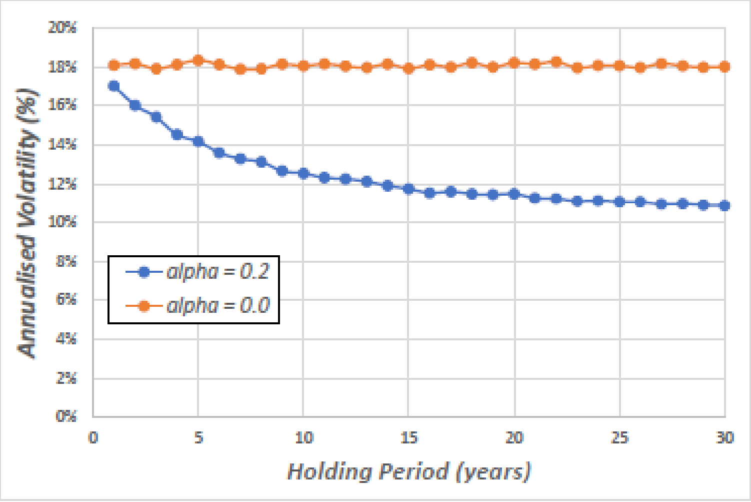
Exhibit A.1: Illustrative model parameters

	No mean-reversion	With mean-reversion
μ	0.06	0.06
σ_1	0.10	0.10
α	*0.00*	*0.20*
σ_2	0.15	0.15

31

The chart below shows how the annualised volatility[32] of the asset return behaves as a function of holding period. In the analysis that follows, results have been produced using 10,000 stochastic simulations, run with a monthly time-step over 30 years.

Exhibit A.2: Annualised volatility term structures produced by the two parameterisations

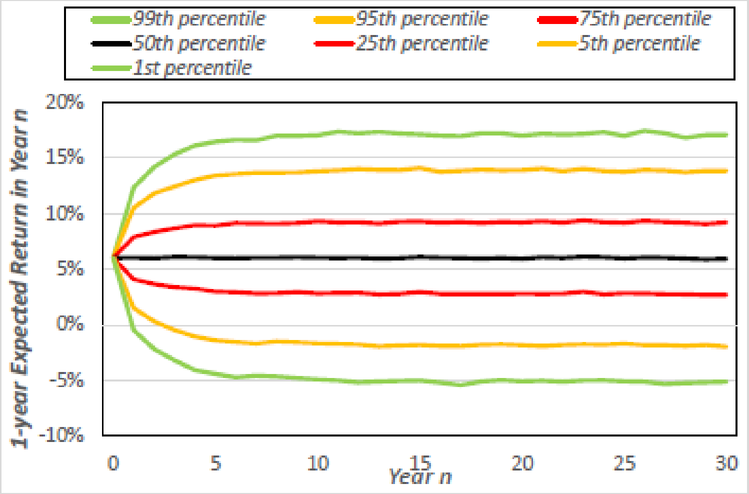


The above results provide an intuitive description of the effects of mean-reversion. In the case where a is zero, the asset price simply follows a geometric Brownian motion with a constant volatility of

18%[33]. However, with a of 0.20, the mean-reversion in the model produces a temporary component in the asset price which mean-reverts over time, reducing the long-term annualised volatility of the asset return. In this case, the annual volatility in any given year is again approximately 18%. But the negative serial correlation in the returns generated over time means that the annualised volatility measure reduces with the length of the holding period. Specifically, the annualised volatility tends to o-1 (assumed to be 10% in this illustration) as the holding period increases. The annualised volatility term structure produced with a of 0.2 will bear some qualitative comparison with the annualised volatility term structure produced for equities by the Wilkie model.

As noted earlier, mean-reversion in realised returns and intertemporal variation in expected returns are two sides of the same coin. With a of 0, the expected return on the asset price is 6% at all times in all scenarios. But with a of 0.2, the short-term expected return of the asset changes over time – with a positive a, the short-term expected return is reduced when x(t), the stochastic variable that represents the temporary component of the market price, is positive, and the expected return is increased when x(t) is negative. This results in the short-term expected asset return having a probability distribution. The distribution produced for the above a = 0.2 case is shown below in exhibit A.3.

Exhibit A.3: The distribution of the short-term expected assets return; $\alpha = 0.2$ ($x(0) = 0$)



The above chart highlights that our example parameterisation of this model generates significant variation in short-term expected market returns. Indeed, the model produces many scenarios where the short-term expected return on the asset is negative. This may seem quite hard to square with a notion of market efficiency. Whilst this calibration is entirely illustrative, material mean-reversion in excess realised returns, for which there is arguably significant empirical evidence, can only arise in the presence of material intertemporal variation in excess expected returns.

We should also note that this postulated asset price behaviour implies simple tactical asset allocation switching strategies could generate high expected returns, and such strategies could be more effective for long-term pension savers than intergenerational risk pooling. Nonetheless, the crucial point in our current context is that, if we subscribe to the view that risky asset prices behave in this way, then intergenerational risk pooling can pool and hence mitigate excess short-term price volatility (and this is also equivalent to providing all participants with a constant expected asset return over all time horizons).

---

[31] [31] $I(t) = \exp[-t\sigma_2^2(1-e^{-2\alpha t})/4\alpha]$. Note $I(t) = \exp(-t\sigma_2^2/2)$ when $\alpha = 0$.

[32] We define the annualised volatility of M(t) as [standard deviation of ln(M(t)/M(0))]/sqrt(t).

[33] Note the square root of (10%2 + 15%2) is 18.03%.

A simple model of an intergenerational risk pooling scheme

Next, we consider what the above model implies for the behaviour of an intergenerational risk pooling scheme. We consider a very simple case that may serve to demonstrate the ideas discussed in section 3.2 above.

Suppose a CDC scheme writes a 1-year lump sum investment product where the premium is invested in the above asset. Further suppose that the product pay-out is determined by the return earned by the permanent component of the asset return rather than the market value return. That is, if 1 is invested in the product at time t, it pays out P(t+1)/P(t) at time t+1.

The premium is invested in the asset at time t, and so the scheme makes a profit / loss at time t+1 that reflects the difference between the change in the market value of the asset and the product pay-out. So, an investment of 1 at time t generates a profit (loss) at time t+1 for the scheme of M(t+1)/M(t) - P(t+1)/P(t).

The following two charts show the cumulative profit / loss that is generated for the scheme by writing 1 unit of this product every year for 30 years with the two model parameterisations in exhibit A.1.

Exhibit A.4: Cumulative P/L, $\alpha = 0$ (i.e. no asset price mean-reversion)

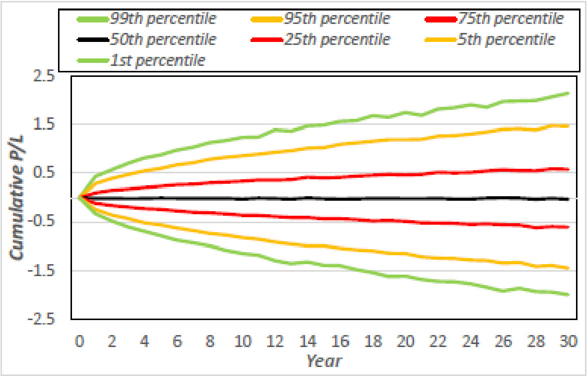
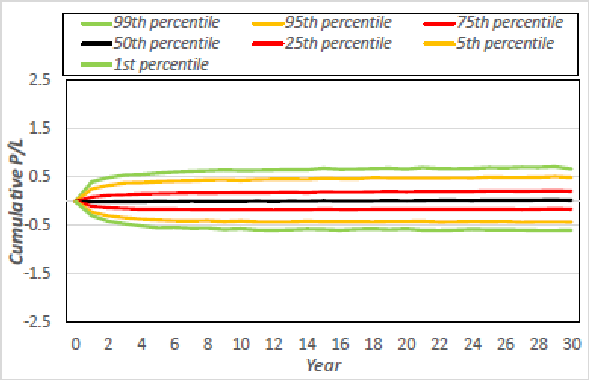


Exhibit A.5: Cumulative P/L, $\alpha = 0.2$ (i.e. with asset price mean-reversion)



These two charts highlight important quantitative and qualitative differences between the two cases. When asset returns are mean-reverting (exhibit A.5), the temporary component of market price changes that the scheme is absorbing largely offsets itself over time, and the long-term outlook for the scheme's risk pooling strategy is stable. There would be no need in this case for an additional mechanism to distribute the past smoothing profits / loss – in this case the market's own correction mechanism is already doing this job. In this setting, members are receiving an asset return with an annual volatility of 10% instead of 18%, and with an expected return that is being maintained at 6% rather than changing materially from year to year.

However, a very different picture emerges in the absence of any mean-reversion (exhibit A.4). Here, the temporary component of market price changes that the scheme is attempting to pool and diversify is actually permanent, and this risk therefore accumulates over time. The uncertainty of the profit / loss of the scheme expands with time horizon, and some form of mechanism will be required to redistribute the accumulating profit / loss back into future generations' returns, thereby materially altering their expected returns.

Let us further develop our quantitative understanding of this through consideration of a specific example scenario. Consider where the market return in year 1 is -34%, and this is comprised of a permanent component of return of 6% and a temporary component of -40%. So, as per the defined product pay-out, the year-1 cohort receives a 6% return (i.e. the permanent component of the return). The -40% loss that has resulted from this pay-out will have to be borne by future cohorts.

Suppose we are in the no-mean-reversion (□= 0) world, and we recognise the need for a mechanism to redistribute the smoothing profit / loss to future cohorts. We decide that 1/5 of the smoothing profit/loss account will be distributed to the next cohort of annual members every year. This smoothing mechanism therefore means that the expected return for the cohorts in years 2, 3, 4, 5 and 6 will have an expected return that is approximately 8% lower than it would be in the absence of smoothing. By definition, the market's expected return in the no-mean-reversion case is unchanged by the year-1 market fall. So, the expected market asset return is still 6%. After the market fall, the product is therefore offering an expected return of around -2% (i.e. 6% - 8%) to new money invested in years 2, 3, 4, 5 and 6. At this point, we might expect fairly searching questions to be asked about the reasonableness of compelling anyone to invest on these terms.

In the case where markets do mean-revert, the product will again offer a lower expected return than the market's expected return following the year-1 market fall of this example scenario. But in that case the market's expected return is highly elevated above its 6% long-term average level. New money invested in the product in year 2 and beyond will have an expected return of 6%, just as it does every other year. Whilst the 6% expected return may appear unenticing at this time when market expected returns are elevated, this is arguably a fundamentally more equitable situation than the one arising in the no-mean-reversion case described above.

We have argued in this essay that, if mean-reversion is present in risky asset returns, intergenerational risk pooling can play the useful function of offsetting the intertemporal variation in expected returns that is generated by the mean-reverting market return process. Subject to the (many and significant) caveats discussed earlier in this essay, the above quantitative analysis supports the proposition that the presence of intertemporal variation in expected returns allows intergenerational risk pooling to be a sustainable and effective system of financial market risk management for DC pension members.

6. Acknowledgments and References

Acknowledgements

The ideas, analysis and arguments developed in this essay are the result of my good fortune of working with many better minds than my own over many years. The preparation of this essay particularly benefited from feedback and discussions with Andrew Barrie, Professor John Hibbert and John Taylor. Any errors and all views are mine alone.

References

Bernhardt and Donnelly (2021), "Quantifying the trade-off between income stability and the number of members in a pooled annuity fund", ASTIN Bulletin: The Journal of the IAA, 51(1).

Cochrane (2005), Asset Pricing (Revised Edition), Princeton University Press.

Fama and French (1988a), “Permanent and temporary components of stock prices”, Journal of Political Economy 96, pp. 246-273.

Fama and French (1988b), “Dividend Yields and Expected Stock Returns”, Journal of Financial Economics 22, pp. 3–25.

Goodhart and Pradhan (2020), The Great Demographic Reversal, Palgrave Macmillan.

Hibbert, Mowbray and Turnbull (2001), “A Stochastic Asset Model and Calibration for Long-Term Financial Planning”, Barrie & Hibbert.

Institute and Faculty of Actuaries (2019), Delivering Collective Defined Contribution Pension Schemes: IFoA response to the Department for Work and Pensions.

Institute and Faculty of Actuaries (2021), The Great Risk Transfer: Campaign Recommendations.

Kennedy et al (1976), “Bonus Distribution with High Equity Backing”, Journal of the Institute of Actuaries, Vol. 103, pp. 11-58.

Poterba and Summers (1988), “Mean reversion in stock prices: Evidence and Implications”, Journal of Financial Economics, Vol. 22, pp. 27-59.

Pension Protection Fund (2021), The Purple Book 2021.

Redington, F.M. (1981). “The Flock and the Sheep and Other Essays”. Journal of the Institute of Actuaries, Vol. 108, pp. 361-404.

Ralfe (2021), “Can Collective Defined Contribution turn base metal into pensions gold?”, Lecture at Imperial Business School.

Roll (1988), “R2”, Journal of Finance 43, No. 3 (July 1988), pp. 541-566.

Shiller (1981), “Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?” American Economic Review 71, No. 3 (June 1981), pp. 421-436.

Smith (2021), “Economic Scenario Generators and Public Policy”, Presentation to the IFoA Finance and Investment Conference, 10 May 2021.

Thornton and Wilson (1992), “A Realistic Approach to Pension Funding”, Journal of the Institute of Actuaries, Vol. 119, No. 2, pp. 229-312.

Turnbull (2017), A History of British Actuarial Thought, Palgrave Macmillan.

Wilkie (1986), “A Stochastic Investment Model for Actuarial Use’, Transactions of the Faculty of Actuaries, Vol. 39, pp 341-403.

Wilkie (1995), “More on a Stochastic Asset Model for Actuarial Use”, British Actuarial Journal, Vol. 1, p. 777-964.

Institute
and Faculty
of Actuaries

Beijing

14F China World Office 1 · 1 Jianwai Avenue · Beijing · China 100004
Tel: +86 (10) 6535 0248

Edinburgh

Level 2 · Exchange Crescent · 7 Conference Square · Edinburgh · EH3 8RA
Tel: +44 (0) 131 240 1300

Hong Kong

1803 Tower One · Lippo Centre · 89 Queensway · Hong Kong
Tel: +852 2147 9418

London (registered office)

7th Floor · Holborn Gate · 326-330 High Holborn · London · WC1V 7PP
Tel: +44 (0) 20 7632 2100

Oxford

1st Floor · Park Central · 40/41 Park End Street · Oxford · OX1 1JD
Tel: +44 (0) 1865 268 200

Singapore

5 Shenton Way · UIC Building · #10-01 · Singapore 068808
Tel: +65 8778 1784

www.actuaries.org.uk