Network for Studies on Pensions, Aging and Retirement

Netspar PANEL PAPERS

Peter Broer, Thijs Knaap and Ed Westerhout

Risk Factors in Pension Returns

Peter Broer, Thijs Knaap and Ed Westerhout

Risk Factors in Pension Returns

PANEL PAPER 21

Netspar Network for Studies on Pensions,Aging and Retirement

Colophon
Panel Papers is a publication of Netspar
December 2010

Design
B-more Design
Bladvulling, Tilburg

Printing
Printing Office Tilburg University

Editorial address
Netspar, Tilburg University
PO Box 90153, 5000 LE Tilburg
info@netspar.nl

CONTENTS

PREFACE

Netspar stimulates debate and fundamental research in the field of pensions, aging and retirement. The aging of the population is front-page news, as many baby boomers are now moving into retirement. More generally, people live longer and in better health while at the same time families choose to have fewer children. Although the aging of the population often gets negative attention, with bleak pictures painted of the doubling of the ratio of the number of people aged 65 and older to the number of the working population during the next decades, it must, at the same time, be a boon to society that so many people are living longer and healthier lives. Can the falling number of working young afford to pay the pensions for a growing number of pensioners? Do people have to work a longer working week and postpone retirement? Or should the pensions be cut or the premiums paid by the working population be raised to afford social security for a growing group of pensioners? Should people be encouraged to take more responsibility for their own pension? What is the changing role of employers associations and trade unions in the organization of pensions? Can and are people prepared to undertake investment for their own pension, or are they happy to leave this to the pension funds? Who takes responsibility for the pension funds? How can a transparent and level playing field for pension funds and insurance companies be ensured? How should an acceptable trade-off be struck between social goals such as solidarity between young and old, or rich and poor, and individual freedom? But most important of all: how can the benefits of living longer and healthier be harnessed for a happier and more prosperous society?

The Netspar Panel Papers aim to meet the demand for understanding the ever-expanding academic literature on the consequences of aging populations. They also aim to help give a better scientific underpinning of policy advice. They attempt to provide a survey of the latest and most relevant research, try to explain this in a non-technical manner and outline the implications for policy questions faced by Netspar's partners. Let there be no mistake. In many ways, formulating such a position paper is a tougher task than writing an academic paper or an op-ed piece. The authors have benefitted from the comments of the Editorial Board on various drafts and also from the discussions during the presentation of their paper at a Netspar Panel Meeting.

I hope the result helps reaching Netspar's aim to stimulate social innovation in addressing the challenges and opportunities raised by aging in an efficient and equitable manner and in an international setting.

Henk Don
Chairman of the Netspar Editorial Board

Affiliations

At the time of writing this panel paper, Peter Broer, Thijs Knaap and Ed Westerhout worked at CPB Netherlands Bureau for Economic Policy Analysis, The Hague.

RISK FACTORS IN PENSION RETURNS

Abstract

Pension schemes are vulnerable to a number of macroeconomic risks. This paper assesses these risks. It reviews the literature on risks in demography, productivity and financial markets, paying attention to their development over time and possible interactions. Using this review, we construct a VAR model for The Netherlands to derive the implications of risks for the returns on four prototype pension schemes, namely defined-benefit (DB) and defined-contribution (DC) schemes that are either funded or PAYG-financed. We find that the returns on funded pension schemes are substantially more risky than those on PAYG-financed pensions. This holds in particular for funded schemes that offer DB pensions. In the latter case, uncertainty increases over time, due to the decrease in the ratio of workers to retirees. Hybrid pension schemes that combine elements of these prototypes offer more stable rates of return, but the diversification gains seem rather modest.

1. Non-technical summary

The economic crisis that currently challenges economies all over the world has reminded us of the fact that risks are part of life. Moreover, the crisis demonstrates that risks need not be confined to typical markets. Indeed, the current crisis started as a problem in part of US mortgage markets, but spread out over other financial markets, in both the US and other countries, and eventually turned into a global economic crisis.

This paper reviews major sources of aggregate risk. In particular, we distinguish between labor productivity risk, demographic risk and financial risks. The latter include equity price risk, interest rate risk and inflation risk. Demographic risk is subdivided in mortality risk, fertility risk and migration risk. Moreover, we account for rare events: low-probability events that have potentially large effects upon equity returns, interest rate and output. The current crisis and previous economic crises are examples, but also wars match this definition. We do not account for environmental risks (such as climate change) and epidemiological risks (pandemics, for example), as these are clearly beyond our expertise. Risks are defined (as is common) as realizations that deviate from earlier expectations. This may look trivial, but is relevant in case of variables that exhibit structural change, like labor productivity or longevity. We describe the different risks in two ways: in the text of the literature review, we discuss both the size of the risk as well as the uncertainty surrounding the estimate of this size. The latter may come from statistical errors that surround estimation, as well as methodological uncertainty. After the review, we encapsulate the main results in a numerical model. From this point on, we ignore the second kind of uncertainty.

Common wisdom holds that equity returns are very risky – and the same holds true, although to a somewhat lesser extent, for inflation and longevity. This paper asks whether this common wisdom is true: are there other factors that are equally or even more risky? Moreover, the paper explores to what extent different risk factors are related. Is a simultaneous decline in output and asset markets (as seen in the current crisis) more the rule or the exception? Are demographic risks related to productivity risks and financial risks or are they orthogonal to each other?

With regard to equity markets, the paper confirms that equity returns are quite risky. The literature has shown that mean reversion of equity returns reduces their risky nature over longer horizons. Campbell and Viceira (2002) argue that over a horizon of 50 years, bonds may be as risky as equity returns. We instead find that the quantitative significance of the mean reversion phenomenon is quite limited. Even over long horizons, equity offers much more risky returns than bonds and bills. Moreover, times in which equity returns are high are, on average, also times in which the volatility of equity returns is above average. This suggests that high expected equity returns may not provide sufficient reason for investors to shift their portfolios towards more equity.

Inflation is a risk factor as well, as historical evidence several decades back has shown. The current crisis is held to have increased inflation risk: the crisis itself bears the risk of deflation, whereas the relaxation of monetary and fiscal policies that followed upon the credit crisis may fuel inflation. Still, we find that the real return on short-term nominal bonds is substantially less risky than the return on equity. The same holds true for the return on long-term nominal bonds.

Another major risk in the economy is labor productivity growth. Labor productivity growth rates vary substantially across countries and over time. Different from shocks in equity returns, shocks in productivity have a long life. The unit-root puzzle in the literature illustrates this nicely: it asks whether productivity shocks are permanent or tend to vanish slowly over time; the possibility that labor productivity shocks are temporary events is no more than a theoretical argument. This long-lasting nature of labor productivity shocks explains our finding that labor productivity shocks may be as persistent as shocks in equity returns. Although the US seems to feature some mean reversion in productivity growth, this mean reversion seems to be lacking in several European countries, among them the Netherlands.

As to demographic risks, we find fertility risk to be huge: in the coming decades, the total fertility rate could fall to a value as low as 1.5 — although the rate could also increase to a value as high as 2.5. In terms of their impact upon the dependency ratio, longevity risks, on the other hand, are much more moderate. This is contrary to common wisdom, which holds that longevity risks are substantial. The difference is explained by the fact that large additions to longevity can be expected in the coming decades on statistical grounds. Only the shocks around this increasing trend are accounted as risks, however, and these risks are an order of magnitude smaller than fertility risks.

The paper has also explored the relationships between different risk factors. Three such relationships stand out. The first is the relationship between productivity and equity returns. The constancy of labor income as a share of GDP suggests that labor income and capital income move together. Benzoni et al. (2007) find that wage income and dividend income are cointegrated. A second important relationship is that between the interest rate and population ageing. The view that population ageing will depress interest rates is widely held; the effect could be as large as a full percentage point. Thirdly, the ageing of the population may depress labor productivity on account of a compositional effect. This effect is relatively weak and, compared to the other effects, much less certain.

What are the implications of this assessment of aggregate risks for the riskiness of returns to various pension schemes? Providing an answer to this question is a tricky exercise, as the uncertainties are huge. However, we can still continue if we bear in mind that the calculations cannot be more than illustrative only. The returns to PAYG schemes depend on productivity growth and changes in the dependency ratio, whereas the returns to funded DC plans depend on capital market returns. The returns to funded DB schemes depend on a mixture of these factors. This implies that the riskiness of the returns to different pension schemes depends on different risk factors: demographic risks and productivity risks mainly drive the returns to PAYG schemes, whereas capital market risks mainly drive the returns to funded schemes.

We do not try here to assess the efficiency of pension schemes. Efficiency of a pension scheme involves both the variability of its rate of return as its mean rate of return. The mean rate of return reflects compensating transfers that have been made during the introduction of the scheme. For example, the low mean rate of return on a PAYG scheme as compared to that on a funded scheme in a dynamically efficient world merely reflects the transfers made to old generations when the PAYG scheme was introduced (Sinn, 2000). Neglecting transitional effects (as in our present analysis) biases the mean rate of return and any measures based on it. One way to correct for this deficiency would be to adjust welfare calculations for these transitional effects, as in Storesletten et al. (1999). Alternatively, one can refrain from calculating integral welfare effects, as done in Matsen and Thøgersen (2004) and in this paper.

We compare the contributions and benefits that correspond to four prototype pension schemes: PAYG DB, PAYG DC, funded DB and funded DC. This merely reflects two well-known themes. First, the variability of contribution rates and replacement rates depends to a large extent upon the risk nature of the scheme: DB or DC. Second, funded schemes imply much lower (higher) and more variable contribution rates (replacement rates) than the corresponding PAYG schemes. Rather than looking at contribution and replacement rates, it is more interesting to look at the internal rates of return of different pension schemes, however. This internal rate of return is defined as the rate of return that equates the net present value of a person's contributions to a pension scheme and benefits from this pension scheme to zero. It is a (much) more comprehensive measure than the contribution rate or the replacement rate. A comparison between the four pension schemes in terms of this internal rate of return reveals that the two PAYG schemes are rather similar, that the rate of return of the funded DC scheme is much more variable than that of the two PAYG schemes, and that the funded DB scheme features the highest degree of rate-of-return variability.

A decomposition with respect to the four sources of risk that we distinguish shows how the four pension systems are susceptible to different kinds of risk. Unsurprisingly, financial risks dominate the variation in returns to funded systems, but play no role in the returns to unfunded systems. Quite unexpectedly, demographic risks turn out to be relatively unimportant for all kinds of pension schemes. Indeed, the main drivers of variation in the returns to PAYG systems turn out to be fluctuations in productivity and, especially, rare disasters.

We also assess the gains from diversification by looking at the variability of the net present value of mixtures of pension schemes relative to the variability that would result if there were no diversification gains to be reaped (perfectly positive correlation). We translate our results into implicit correlation coefficients. These are easy to interpret, and can be compared with some earlier analyses in the literature. Our correlation coefficients are actually quite large: between 0.5 and 1. This suggests that the risk reduction that can be achieved by mixing different pension schemes should not be exaggerated.

2. Introduction

What is the most important risk that is facing pension schemes in the coming century? Is it the continued ageing of the population, or the risk of the current crisis leading to a prolonged slump in economic growth? Does the volatility of financial returns represent the largest uncertainty, or is it the occasional disaster that wipes out years of earnings? For all interested parties, the answers to these strategic questions are key to planning for the future of retirement provision.

This paper aims to give insight into the long-term development of the determinants of pension returns. What exactly determines these returns depends on the financing mode (pay-as-you-go (PAYG) versus funding), the mechanism for shock absorption (defined benefit (DB) versus defined contribution (DC)) and the degree of actuarial neutrality (Lindbeck and Persson, 2003). The returns to funded schemes are often higher than those on PAYG schemes, for example, if the capital market rate of return exceeds the rate of economic growth.

But the choice of pension plan does not just affect average returns; the uncertainties involved in various types of plans are of a different character and likely of different magnitude. They depend on productivity growth and changes in the dependency ratio for a PAYG scheme, on capital market returns for funded DC plans, and on a combination of these determinants for funded DB schemes. All of these determinants are themselves intrinsically uncertain, and the uncertainties often increase with the time horizon.

We review the literature on four types of macroeconomic risk, each of which may have an important influence on pension returns. These are productivity risk, demographic risk, financial asset return risk and the risk of rare, but high-impact, disasters. We review the literature in two different ways. The first is a discussion of the main contributions to the literature. Here, we bring together many different strands in the literature. There is a large literature on macroeconomic risks (Nelson and Plosser, 1982; Cochrane, 1988; Barro and Ursúa, 2008), on financial risks (Campbell, 2003; Lettau and Ludvigson, 2009), and on demographic risks (Lee and Carter, 1992; Lee and Tuljapurkar, 1994; Keilman et al., 2008). There is also a large literature that explores interactions between different risk factors, like the relation between the interest rate and the age structure of the population, or the relation between asset returns and productivity growth. We review these literatures and present an integrated assessment of the importance of different risk sources for a small open economy. Our paper is the first (to our knowledge) that brings together all these fields and integrates the analysis of demographic, financial and macroeconomic risks, including the development of these risks over time and the interactions that exist between them.

The second way we review the literature is by constructing a vector-autoregressive (VAR) model that brings together the insights from the literature. The model represents a quantitative formulation of the different risks and their influence upon each other. Actually, the model is based on three sources: i) the Campbell and Viceira (2002) model that specifies capital market risk, ii) empirical estimates for other parts of the model, as can be found in the literature, such as the link between demography and the interest rate and the correlation between capital income risk and labor income risk, and iii) empirical estimates for some parts of the model that we did not find in the literature and that we thus had to produce ourselves, like the development of Dutch fertility rates and the relation between US labor productivity and labor productivity in our own country.

Compared with the first part of the literature review, the model is (much) more concrete. Model builders are forced to adopt a specific modeling, even when opposing views exist in the literature. Thus, part of the uncertainty in the literature is swept under the rug — and as a consequence, risks may be underestimated. This approach also bears the risk that at some points the wrong choice is made (so that model calculations will be biased). An attractive alternative would have been to estimate all the relations in the VAR model ourselves, but this would obviously have been far beyond the scope of a panel paper.

We use the VAR model to simulate developments in the Dutch economy over several decades and to determine the expectation and variability of growth, demography, financial returns and rare disasters. With these simulations, we assess the variation of pension benefits and contributions under various schemes, as well as the behavior of more comprehensive measures of the return to a pension scheme. Using stylized examples, we distinguish between pension schemes that differ in financing structure (PAYG, funding) and in the allocation of risk (DB, DC).

We also explore whether a mixture of pension schemes reduces risks by allowing for more diversification. A related paper that explores comparable ideas is Matsen and Thøgersen (2004). Their theory is that if the returns to different pension schemes are driven by imperfectly correlated risk factors, a combination of pension schemes may yield more stable returns. Their empirical application suffers from some drawbacks, however. The risks in their article are stylized forms based on past realizations. The current analysis offers a more detailed representation of future risks, which allows us to tackle the issue in a more realistic setting. Moreover, the two-period model they employ does not allow them to account for the autocorrelation patterns that can be found in the data and that are reflected in our analysis. Furthermore, the Matsen and Thøgersen paper does not allow for rare disasters, an omission that would be difficult to motivate in a 2010 paper. Wagener (2000) is another related paper that shows how PAYG pensions can add an element of diversification to insure against macroeconomic risks. Using a model with two periods and a limited number of risks, he shows that this argument works in favor of defined-benefit (DB) pension schemes.

The structure of the paper is as follows. Section 3 reviews the literature on the four major risk factors, including the literature on the relations between them. Results from a VAR model that reflects the main findings of this review are the subject of section 4. Section 5 specifies how macroeconomic risks affect premiums and benefits of four prototype pension schemes that are either PAYG or funded and DB or DC. Section 6 uses stochastic simulation to compare the four pension schemes. Finally, section 7 concludes.

3. Risk Factors

We consider the following risk factors in an economy: the fertility, mortality and migration rates of different cohorts over time; the development of total factor productivity; and financial returns to bonds, stocks, money (inflation) and safe assets. This section first reviews productivity risk, and then discusses demographic and financial risks. We end with a discussion of systemic risks, which have regained interest recently. The discussion is based on the survey of the literature in Broer (2010).

3.1 Productivity Risk

In Western countries, economic growth is largely the result of increases in productivity. Total factor productivity (TFP) can be usefully split into a deterministic- and a stochastic component (see Box 1). Deterministic increases in productivity can be the result of planned new investment, such as infrastructure; the stochastic component captures technological shocks and various macroeconomic shocks.

Measurements of the data-generating process behind the stochastic component of technology shocks have shown that there is a large degree of autocorrelation. This makes it difficult to distinguish statistically between transitory shocks and persistent shocks. An autocorrelated transitory shock can affect the growth rate for a long period of time, but eventually washes out of the process; persistent shocks lead to a permanent change in the level of TFP. The long-run implications of these two possibilities are very different: with an autocorrelation coefficient of 0.98, for example, after 40 years more than half a shock in the level of TFP has vanished.

Box 1: Productivity

The aggregate production function can be written in terms of a function *F* that combines the input from capital and labor and total factor productivity, *S*:

$$y_t = S_t F\left[K_t, L_t\right] = \zeta_t X_t F\left[K_t, L_t\right] \qquad (1)$$

$$\ln\zeta_t = \rho\ln\zeta_{t+1} + \gamma t + \varepsilon_t \qquad (2)$$

The rate of growth of total factor productivity, $\dot{S}/S$, also known as the "Solow residual", can be decomposed in a stochastic component, ζ, and a deterministic component, *X*. This model has a unit root in TFP growth when $\rho = 1$. Empirical research for the U.S. shows that technology shocks are highly correlated over time, with $\rho \approx 0.98$ and $\sigma_\varepsilon \approx 0.0072$ on a quarterly basis (see King and Rebelo, 1999). On an annual basis, the standard deviation of output shocks is 0.028. The implausibly large size of these shocks, which are difficult to associate with observable events, is one of the main criticisms of Real Business Cycle (RBC) models. Indeed, the high correlation of shocks triggers intertemporal substitution effects and is one of the main driving mechanisms of RBC models (King and Rebelo, 1999). Filtering out ζ from *S* should lead to an unbiased estimate of both components of productivity. There are, however, substantial problems with the correct measurement of inputs (e.g. with respect to quality adjustments and spillovers from R&D). Any systematic measurement error in the Solow residual will also affect the TFP measure. See Carlaw and Lipsey (2003).

Economic agents also respond differently to the two types of shocks. With persistent shocks, the income effect of a shock is much larger, whereas the effect in terms of intertemporal substitution is smaller. It is therefore important to be able to discriminate between these possibilities, and a large body of empirical research deals with the question of whether shocks are persistent. This question is usually put in terms of whether there is a unit root in gross domestic product (GDP) per capita (GDP is less prone to measurement error than TFP). Having a unit root in the process means that there is no tendency to revert to earlier levels after a shock: the best predictor for future values of TFP is its value today.

Any insights into future fluctuations of economic growth must involve understanding this aspect of the data-generating process. We therefore spend some time on the unit-root question.

3.1.1 A Unit Root in GDP?

The issue whether there is a unit root in GDP was put on the agenda by Nelson and Plosser (1982), who concluded that many macroeconomic variables appear to be difference stationary over a 60-year span,[1] implying that an unexpected shock has a permanent effect on the level of the variable. In particular, a 1% shock in gross national product (GNP) should lead to a revision of the level of future GNP of *more* than 1%, due to the lags in the data-generating process. Their results suggest that real per capita GNP may possess a unit root. Nelson and Plosser's result was confirmed by Campbell and Mankiw (1987), among others, who

1 The sample period in the Nelson-Plosser data set ranges from 1860-1990 to 1909-1970.

used a more elaborate testing framework. During the eighties, a consensus emerged about the presence of a unit root in GNP.

The unit-root consensus was broken in the nineties. Rappoport and Reichlin (1989) and Perron (1989) show that if structural breaks in the regression coefficients are allowed, GNP is in fact stationary. Zivot and Andrews (1992) identify the most likely candidates for structural breaks: the Great Depression and the oil price shock of 1974. The relevance of the Nelson-Plosser findings for the post-war period has also been questioned. Christiano and Eichenbaum (1990) and Rudebusch (1993) argue that if uncertainty with respect to lag length is taken into account, unit-root tests based on post-war US data have insufficient power to reliably distinguish between both hypotheses. The choice of sample period appeared to be vital to the conclusions, in any case. Using a longer sample period, starting in 1870, Diebold and Senhadji (1996) were able to reject the null hypothesis of a unit root in favor of stationarity – even in the absence of structural breaks.[2] One side effect of using long sample periods is that the homogeneity of the data can be violated more easily, as argued by Murray and Nelson (2000). In Murray and Nelson (2002), the authors use a Markov switching model to show that the Great Depression was characterized by more volatile shocks. Taking this conditional heteroskedasticity into account, they conclude that GDP does have a unit root.

The debate has continued unabated. Lima and de Jesus Filho (2008) and Cook (2008) argue that a nonlinear alternative to the

2 Sen (2004) and Vougas (2007) take a similar approach as Perron, and Rappoport and Reichlin, using an extension of the original Nelson-Plosser data set. They both conclude that the data are trend-stationary, provided that one allows for breaks or nonlinearities in the constant term or the time trend.

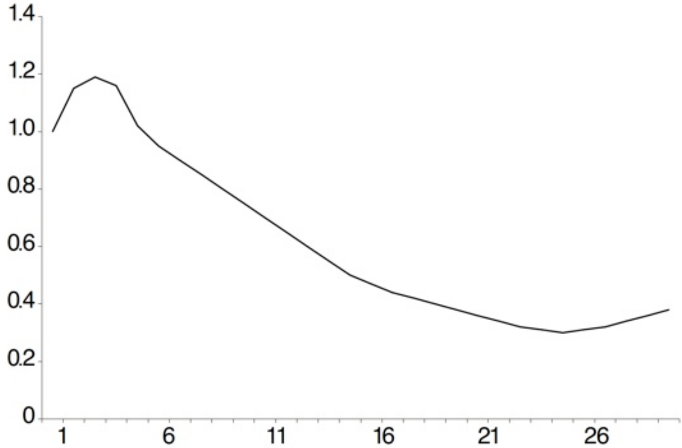

Figure 1: Annualized variance ratio of log GNP for the U.S. (source: Cochrane (1988))

standard ADF test leads to the conclusion that US GDP is trend stationary. On the other hand, Benati (2007) concludes that labor productivity is a unit-root process, which virtually implies that GDP per capita is also nonstationary.

Most research in this area focuses on the United States. Cochrane (1988) analyses the behavior of the annualized variance ratio of US GNP, $\frac{1}{k}\operatorname{var}\left(\ln\frac{\mathrm{GNP}_t}{\mathrm{GNP}_{t-k}}\right)$, over time. If GNP is a random walk, this annualized variance ratio is a constant. But if GNP is trend stationary instead, the annualized variance ratio falls to zero. Cochrane (1988) finds that the annualized variance ratio does decline over time – but not all the way to zero (see Figure 1): This indicates that GNP may indeed be non-stationary, but with a variance that grows less rapidly than that of a random walk.

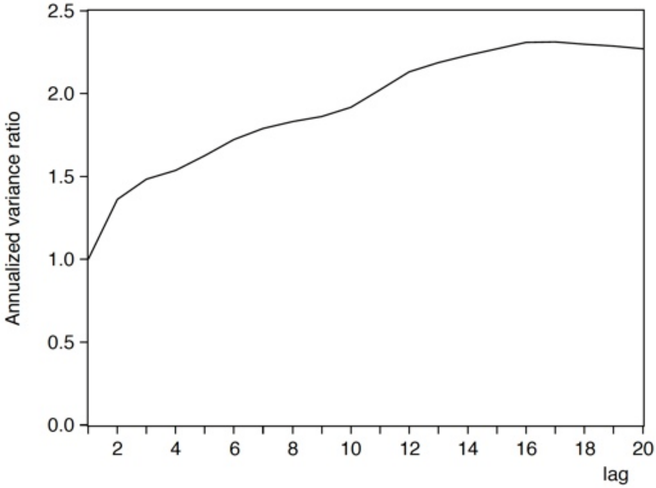

Figure 2: Annualized variance ratio of log GDP for the Netherlands

Cogley (1990) applied the method of Cochrane (1988) to a number of OECD countries. It appears that the mean reversion found for the United States is not present to the same degree in other countries. Several countries (e.g. Italy, France) display increasing annualized variances. Figure 2 shows that for the Netherlands, too, the annualized variance ratio increases for periods up to twenty years. Indeed, the variance ratio grows more strongly than that of any country included in Table 2 of Cogley, with a 15–20 year ratio of about 2.3. The case for a unit-root specification seems therefore to be stronger for the Netherlands than for the US.[3]

3 For the Netherlands, the augmented Dickey-Fuller (D-F) statistic for postwar log GDP per capita is -1.66; for the growth rate of GDP per capita it is -5.26. This suggests a unit root in the level of GDP. For log labor productivity per hour, the D-F statistic is -1.07, and for productivity growth it is -2.8. It follows that a random walk in productivity *growth* cannot be rejected on statistical grounds at the 5% level.

One problem with the whole approach is that it denies possible spillovers between countries. This position is difficult to defend in an era in which economies become increasingly integrated. This may have implications for the findings as well. As the US was the technology leader over the sample period, the apparent nonstationarity of GNP for the other countries might be due to catching up. In addition, to the extent that convergence between countries occurs (Barro, 1996; Temple, 1999), GNP processes must be cointegrated, which precludes a fundamental (i.e., asymptotic) difference in variance profiles.

The current dominant view in the unit-root discussion is that quantity series may generally be characterized as trend stationary, *provided that* breaks or nonlinearities in the level or the trend parameter are included. The problem with this point of view is that the possibility of future breaks in the parameters of the data-generating process (DGP) is left open. In a sense, conditioning the stationarity of the data on exogenous structural breaks begs the unit-root question. If the breaks are not one-time events, they must be part of the DGP.[4] As such, the *unconditional* development of GNP, for instance, may be nonstationary.

In addition, there is evidence that the stochastic process for GNP in Anglo-Saxon countries is best described as an intermediary case between stationarity and non-stationarity, in the sense that forecast uncertainty increases with the horizon — but not as fast as in the case of a random walk. However, the available evidence for continental Europe firmly points in the direction of a unit root for GNP.

4 In Barro (2006), Barro and Ursúa (2008) and Barro et al. (2009), the breaks are treated as rare disasters, see Section 2.4.

3.1.2 Spillovers from demographic risk to productivity risk

Apart from the essentially stochastic developments in technology, it is likely that productivity is related to demographic factors. Indeed, with age-dependent productivity, the age composition of the labor force directly affects aggregate productivity. In addition, the ability to handle product innovations or process innovations may deteriorate with age. We discuss both channels in turn.

Direct measurement of the relation between age and productivity is difficult. Implicit contracts and tenure effects obscure the effect of productivity on wages (see the survey of de Hek and van Vuuren, 2008). A few studies attempt to measure the effect of age on productivity directly. OECD (1998) surveys the psychological literature to find that cognitive skills deteriorate only to a small extent between the ages of 40 and 65. Kotlikoff and Gokhale (1992) use a panel data set on age-earnings profiles within a single large US firm. They find that productivity does indeed fall with age for older workers for all job types considered. They tentatively conclude that productivity at age 65 is only one-third of peak productivity (at age 45). However, Hellerstein et al. (1999), in a study that uses data from different plants, did not find that productivity of older workers is lower than that of the reference group. Dostie (2006) accounts for unobserved heterogeneity between workers and plants, and also concludes that productivity of older workers does not fall.[5] Prskawetz and Lindh (2006) reach a different conclusion. They present a survey of piece-rate studies from which they conclude that individual productivity has an inverted U-shaped profile. They attribute the productivity decline at later age to a reduction in

5 However, older workers with at least a college degree are paid more than their productivity.

cognitive abilities. A limitation of their survey is that it contains mainly studies on the performance of artists and scientists, which are rather special groups.

There is some literature that connects output to health and longevity. Bloom et al. (2004) estimate that a one-year increase in life expectancy raises productivity by 4%. Weil (2007) also finds a positive effect from health on productivity. However, Acemoglu and Johnson (2007) fail to find a casual relationship from improvements in life expectancy to economic growth.

Boersch-Supan (2003) calculated the expected effect of ageing on German productivity growth, assuming that the age-productivity profiles found by Kotlikoff and Wise (1989) may be applied. He finds that the expected drop in productivity will probably not exceed 3% over the next two decades (i.e. at most 0.15% per year, or 0.1% per percent increase in the dependency ratio). This suggests that the effect of demographic shifts on productivity growth is minor. From the estimates in Nahuis et al. (2000), the effect of the dependency ratio on log labor productivity is about −0.09, in line with the calculations of Boersch-Supan. Based on these findings, we conclude that there is an effect of the dependency ratio on log labor productivity, equal to about −0.1.

3.2 Demographic Risk Factors

Population growth rates are difficult to predict over long periods. Keilman et al. (2008), Table 2.1, presents official forecasts of old-age dependency ratios in 2050 that shift by as much as 10%-points over the period 1994-2004. This volatility is an indication of non-stationarity of the underlying demographic process. Part of the problem is that there are three demographic risk factors, of which each may follow a different stochastic process: fertility, mortality and migration. We discuss the three separately.

3.2.1 Mortality

Life expectancy has risen steadily over the twentieth century and may well continue to rise for the next few decades. Estimates for the expected remaining life of the current population are essential information for anyone running a pension scheme. Yet revisions of this estimate are a frequent occurrence, indicating that large shifts in expected mortality happen quite regularly. Several methods are available to forecast future mortality (e.g. life tables, hazard models; see Sickles and Taubman (1999) for a survey). De Waegenaere et al. (2010) offer a general discussion of longevity risk and a survey of current statistical mortality models. A simple stochastic model that adequately describes the uncertainty in mortality rate changes is that of Lee and Carter (1992; see box 2), in which the mortality for different age groups depends upon an unobservable variable, an indicator for general developments in health and medicine. The Lee-Carter (LC) model has thus become the workhorse of demographics. The forecasting performance of the LC model was tested by Lee and Miller (2001), who concluded that it ``produces surprisingly accurate forecasts over rather long periods.'' Moreover, the LC model appears to generally outperform expert-opinion-based official projections of mortality in the past century. However, the nonstationarity of the dynamic state variable may lead to implausible long-run projections. De Waegenaere et al. discuss some alternatives.

Hári et al. (2008) estimated the Lee-Carter model on Dutch data using a Kalman filter approach. They showed that the hidden state variable can be characterized as a random walk with a trend. The

Box 2: Mortality

The standard model for mortality is Lee and Carter (1992), which is of the form

$$\ln \lambda_{t+1,\tau} = \alpha_{\lambda,\tau} + \beta_{\lambda,\tau}\mu_{\lambda,t} + \varepsilon_{\lambda,t+1,\tau} \qquad (3)$$

$$\mu_{\lambda,t+1} = \delta_{\lambda,1}\mu_{\lambda,t} + \delta_{\lambda,0} + \eta_{\lambda,t+1} \qquad (4)$$

Here, $\lambda_{T,\tau}$ is the mortality rate in year *t* for age τ and μ_λ is a hidden state variable. The $\varepsilon_{\lambda,T,\tau}$ are i.i.d. random variables with mean zero. The important feature of equation (3) is that mortality rates for different ages are all driven by the same state variable μ_λ . The hidden state develops according to a stochastic difference equation. The development of the hidden state variable appears to be described well by a random walk with drift ($\delta_{\lambda,1} = 1$), which renders the process for mortality rates non-stationary. Estimates for the U.S. show that $\delta_{\lambda,0} \approx -0.365$ and $\mathrm{E}\left[\eta_\lambda^2\right] \approx 0.65$. For persons of age τ<60, $\beta_{\lambda,\tau} \approx 0.03$, so that the rate of mortality of the elderly falls on average by about 1% per year.

trend, however, is rather uncertain if one allows it to be time-varying. This implies that predictions of future mortality rates do indeed, in a similar way as TFP, show increasing variation when the time period of the prediction is increased.

3.2.2 Fertility

Lee and Tuljapurkar (1994) used an ARIMA (1,0,1) model with a restriction on mean fertility of 2.1 child per woman to estimate fertility rates. Their estimates imply a 95% long-term confidence interval for fertility from 1 to 3 children per woman. Obviously, this implies huge uncertainty about population growth. In the "Uncertain Population of Europe" (UPE) project (Keilman et al., 2008), a similar approach is chosen, but Northern and Western European countries are assumed to have a long-run average fertility rate of 1.8, and Southern European countries an average fertility rate of 1.4. As before, these estimates imply a large degree of uncertainty in population growth rates.

The development of fertility of Dutch women can be explained reasonably well using a modified version of the Lee-Carter model. The modification involves fertility being modeled per cohort, instead of age, and uses two unobservable states, one describing the total fertility of a cohort, and the other the shift in the age distribution of fertility for that cohort. Modeling fertility shifts by cohort should be more compatible with changes in fundamental determinants of fertility than modeling fertility shifts by age. However, compared to Lee and Tuljapurkar (1994), the uncertainty about future fertility is just as large – basically, because the current value of the fertility state vector is imperfectly observed. The estimated fertility model is discussed in detail in Broer (2010). The estimates imply that the mean of the total fertility rate will converge to a level of 1.8 children per woman.

3.2.3 Migration

A third source of demographic uncertainty is migration. The UPE forecasts referred to earlier assume that the change in migration is normally distributed, $\Delta N_{t,\tau} = S_{t,\tau}\left(\eta_\tau + \delta_{t,\tau}\right)$, where the scales $S_{t,\tau}$ are country-specific and η and δ are random variables. For the Netherlands $S_{t,\tau} \approx 2.0$, and $\eta \infty N(0{,}03)$ and $\delta \infty N(0{,}07)$.

There is evidence that labor market conditions affect migration. In line with the Harris and Todaro (1970) model, wage and unemployment levels affect the inflow of immigrants (see e.g. Karemera et al., 2000; Clark et al., 2002; and Mayda, 2005). Restricting ourselves to the four risk factors that are the subject of this paper, we can capture labor market effects through the productivity risk factor. The main determinants of total net migration into the Netherlands are current and lagged productivity and population size. The long-run elasticity of productivity is rather small, however: only 0.1.

3.2.4 Spillovers from productivity risk to demographic risk

There is a considerable literature on economic determinants of fertility. Barro and Becker (1989) regard fertility as a choice made by forward-looking generations of households, linked by altruism. In their model a higher rate of technical progress leads to a lower fertility rate. Galor and Weil (2000) and Galor (2005) take this argument further to develop a complete model of the demographic transition that features an inverted U-shaped relation between productivity and fertility. Lehr (2009) empirically corroborates the predictions of the Galor-Weil model. Doepke (2004), Alders and Broer (2005) and Manuelli and Seshadri (2009) consider the effect of government policies on human capital accumulation and fertility.

Mortality is also affected by economic conditions. Fogel (1997) explains the secular decline in mortality from improved nutrition, due to improvements in food production and food distribution. In turn, increases in life expectancy may affect fertility, as argued by Kalemli-Oczan (2002) and Soares (2005). Birchenall (2007) considers the effect of health investment on child survival and life expectancy.

More research is needed, however, to ascertain whether these results are robust.

3.3 Financial Risks

Campbell (2003) and Lettau and Ludvigson (2009) provided an overview of some stylized facts on asset market returns. For the purposes of this paper, the following observations are particularly relevant:

1. For the US, the average (geometric) real return on stocks over the postwar period was 8%, while the real return on (risk-free) Treasury bills was 0.9%. This compares with a return on stock in the Netherlands of 14%, and on short-term government bonds of 3.4%. The EU average returns are about 9% for stock and 2% for government bonds. Hence, the post-war excess return of equity over bonds is about 7%, both in the US and the EU.[6]
2. Stock return has an annual standard deviation of 15% in US data, whereas the return on T-bills has a 1.7% standard deviation.
3. The US excess return of equity over bonds is, in substantial measure,predictable over longer horizons.
4. The Sharpe ratio (the "market price of risk") for the US stock market is countercyclical and highly volatile.
5. The volatility in the stock market is persistent, with a mean lag of about a year.

From this list, we see that there is substantial variation in the volatility of returns over time. In addition, the market price of risk (the Sharpe ratio) also fluctuates over the business cycle. These

6 Campbell's data for Europe cover the period 1970-1999. Taking the stock market crashes of 2002 and 2008 into account lowers the excess return by about 2%-points. For the US, the crash caused the excess return over the period 1891-2008 to fall by about 0.5%.

stylized facts are discussed below in terms of two related issues: the size of the equity risk premium, and the predictability of equity returns.

3.3.1 The Equity Premium

he first stylized fact above states that, historically, stock has earned an excess return over risk-free assets of about 7%. If this historical excess return is viewed as the equity premium, it is difficult to explain by standard consumption-based asset pricing in complete markets: the well-known equity premium puzzle. Recent surveys of this puzzle are given by Salomons (2008), Donaldson and Mehra (2008), and Mehra and Prescott (2008). For practical purposes, the important question is what can be said about future (expected) excess returns on equity. Do excess returns vary systematically with economic conditions? Are excess returns to some extent predictable?

A preliminary question, which must be addressed first, is whether the observed high excess returns to equity of the past are an indication of high expected excess returns over that same period. Did investors ex ante have any reason to expect these high returns? If the observed high excess returns in the US and other long-lasting stock markets are a consequence of survivor bias, as argued by Jorion and Goetzmann (1999), then the expected excess return (i.e. the equity premium) was probably much lower. However, Dimson et al. (2008), using a database of seventeen countries over a 106-year interval, estimated the historical world equity premium, relative to bonds, at 4.1%, a bit below the US figure of 4.5%.[7] After

7 Geometric means (i.e. annualized returns). The higher figure of stylized fact in Section 2.3 above refers to the post-war period.

adjustment for survivorship bias, this figure drops to 4.0%, a negligible correction.

A related point is that conditions in the past century were not representative for current conditions, so that *future* returns may turn out to be much lower. Fama and French (2002) argue that the high average return over the period 1950–2000 is caused by a decline in discount rates.[8] Using dividend/price ratios and dividend growth, Fama and French estimate an equity premium from 2.5% to 4.3% over that period, depending on whether one uses dividend or earnings forecasts. Dimson et al. (2008) use a similar decomposition into fundamentals to argue that the current equity premium should be in the range of 3%–3.5%. A "reasonable" lower bound for the equity premium, based on fundamentals, is in the order of 3%. An upper bound may be set at the historically observed excess return of 7%.

3.3.2 Asset Return Predictability

Estimating the current equity premium is a special case of the more general problem of efficient prediction of asset prices from current information. Under the random walk hypothesis of asset markets (see Fama, 1965), such an effort would be futile. However, research in the seventies showed that excess asset returns are to some extent predictable; see e.g. Bodie (1976) for the real return to equity and Fama and Schwert (1977) for the real return on short-term bonds. Predictability has also been established with respect to medium-term excess returns of equity; see e.g. Fama and French (1988).

8 This argument finds support in the results of Lettau and van Nieuwerburgh (2008) (section 2.3.2).

It appears from the survey in Lettau and Ludvigson (2009) that a whole range of variables is able to predict excess stock returns, including price-dividend ratios, term spreads, and proxies for the consumption-wealth ratio. Generally, the predictability of stock returns appears to improve with the horizon. A widely used variable is the price-dividend ratio (see Campbell and Viceira, 1999; Campbell and Shiller, 2001), which is able to explain 10% of the variation in excess returns in the US stock market over a ten-year horizon. However, this finding does not hold for continental Europe. For the Netherlands, in particular, Campbell and Shiller find no relation between price-dividend ratios and stock prices growth.

The dividend-smoothing model implies that price-dividend ratios contain information about discount rates of investors; see Cochrane (1994) and Broer (2010). If discount rates are stationary, so are price-dividend ratios – and any deviation from the unconditional mean must signal a future corrective price adjustment. A fall in the price-dividend ratio of 1% therefore generates a predicted excess return of 1%. The price-dividend ratio does not forecast well over short horizons, however. According to the survey in Table 2 of Lettau and Ludvigson (2009), it is outperformed by the consumption-wealth ratio proposed by Lettau and Ludvigson (2001). Moreover, Lettau et al. (2008) and Lettau and van Nieuwerburgh (2008) showed that the ability of dividend yields to forecast returns declined after 1995. There appears to be a structural break in dividend yields from 1995 on, with substantially lower dividend yields after this date. Lettau and van Nieuwerburgh show that conditional on this break, the predictability of returns from dividend yields remains intact. In line with the dividend-smoothing hypothesis, they argue that the break in dividend yields signals a break in discount rates, with substantially lower risk premiums being applied after 1994.

Business cycle information also helps predicting stock returns (see Fama and French, 1989; Campbell and Diebold, 2009; and Lettau and Ludvigson, 2001). The results from studies of return predictability suggest that predictability improves with the length of the horizon over which the return is measured. However, Valkanov (2003) corrects for the inconsistency of estimation because of overlapping data, and finds that it seriously weakens the case for long horizon predictability.

Another approach to long-run predictability is to estimate a VAR model of one-period returns and derive the multi-period returns from the VAR model. Campbell et al. (2003) and Campbell and Viceira (2005) investigate excess return predictability of a number of financial assets in the context of such a model.[9] The estimated model implies declining standard deviations of the real return to equity and five-year bonds, as a result of mean reversion in the returns. In contrast, the standard deviation of the real return on short-term bonds quadruples, due to inflation risk, which makes short-term bonds just as risky as equity in the long run.

Campbell and Viceira do not present standard errors of multiperiod return forecast errors, so that it is difficult to judge the reliability of the model for long-run forecasts. Potentially, parameter uncertainty could render their long-term results statistically insignificant.[10]

---

9 The structure of the Campbell and Viceira (2005) model is discussed in Broer (2010).

10 Moreover, the variance profile estimates in Figure 1 of Campbell and Viceira (2005) do not follow from the published VAR coefficients and error covariances in their Table 2. On the basis of these parameter values, the variance profiles are almost completely flat.

Hoevenaars et al. (2008) use an extended version of the Campbell and Viceira model, adding more assets, the credit spread, and inflation-linked liabilities. The additional assets are assumed not to generate dynamic feedback on the primary assets, which separates the model into a "core" VAR model and a supplementary VAR. Their results for the core VAR confirm those of Campbell and Viceira. Mean reversion in stock returns is attributed both to dividend yield behavior and to the credit spread.

The finding that excess returns are to a certain extent predictable was challenged by Goyal and Welch (2003, 2008), who used out-of-sample forecasting to test the predictive ability of dividend yields and other factors. They found unstable parameter estimates and poor forecasting performance, compared to forecasts based on the sample mean. Cochrane (2007, 2008) and Lettau and Ludvigson (2009) argued that the results in Goyal and Welch primarily show the effect of finite-sample bias on forecasting errors, and do not deny the existence of a structural relation between dividend yield and future excess returns. However, the results of Goyal and Welch (2008) do cast doubt on the possibility of successfully using dividend yield information in a portfolio investment strategy, as advocated in Campbell and Viceira (2002).

This review of the literature leads to the following conclusions:

1. There are strong indications for mean reversion in equity returns.
2. Mean reversion in asset returns does not necessarily indicate reduced risk exposure for long-term investors

The conclusion that mean reversion exists is based on several arguments. First, there is direct empirical evidence for the proposition, since asset returns show systematic dependence on *stationary* economic variables, notably the price-dividend ratio over medium-term horizons, and the consumption-wealth ratio over short- to medium-term horizons (Section 3.3.2). Second, there is substantial indirect evidence in terms of autocorrelation in volatility and the price of risk. Volatility is clearly correlated with the onset of a recession, and has a half time of about one year. ARCH models and stochastic volatility models show that there is a volatility-in-mean effect, implying that the equity premium goes up at the start of a recession (i.e. at the same time that realized excess returns fall). Third, the mean reversion result is reinforced by the finding that the price of risk varies countercyclically, and usually rises during a recession.

The existence of mean reversion in asset returns does not mean that it can be usefully exploited in terms of an investment strategy, however. First, out-of-sample prediction of asset returns is much more difficult than in-sample prediction, due to parameter uncertainty. Second, high expected excess returns coincide with, and are triggered by, high and persistent volatility. The above-normal expected return must be balanced against the above-normal return risk.[11] Third, the price of risk rises during recessions, which signals that the risk-absorbing capacity of investors is limited. Private investors start worrying about their consumption possibilities, along the lines sketched by Campbell and Cochrane (1999), and start to sell. Institutional long-term investors may be less risk-averse, but they need to meet specific solvency conditions – and with falling stock prices, these conditions become more stringent as well. Indeed, in a standard Value-at-Risk framework,

11 Time-varying volatility is missing from the VAR analyses of Campbell and Viceira (2002) and Campbell and Viceira (2005) on the argument that it is short-lived. However, a half life of one year of volatility corresponds very well with the average duration of a recession.

the increase in volatility may boost the price of risk of institutional investors just as much as that of private investors.

In the remaining part of this paper, we adopt the results from Campbell and Viceira (2005) to model financial risks. We must stress, however, that this model has several weaknesses, of which two stand out. First, asset return volatility is kept constant, whereas in the data it shows substantial correlation over time with mean lags of about one year. Volatility is also strongly countercyclical (Lettau and Ludvigson, 2009). Second, empirical studies show that the price of risk, which is also a constant in Campbell and Viceira (2005), is countercyclical and rises sharply during a recession (Campbell and Cochrane, 1999). In combination with countercyclical volatility, this implies strongly countercyclical expected returns.

Returning to our theme of the four main risk factors, it is quite likely that developments in the economy will have an effect on financial market returns. Demography will have an impact through the changing supply of capital that results from an ageing population. Productivity itself has an immediate effect on the return to capital, and must ultimately determine the performance of financial assets that are linked to production. We now discuss these two connections in turn.

3.3.3 Spillovers from Demographic Risk on Real Rates of Return

Worldwide ageing affects both supply and demand on capital markets. According to the standard life-cycle hypothesis, consumption smoothing leads to wealth accumulation during the working life, and dissaving in retirement. This implies that, *ceteris paribus*, population ageing will lead to a relative abundance of capital, rising real wages and falling rates of return on capital. However, this prediction depends strongly on the assumptions with respect to public spending and social security. Kotlikoff et al. (2001) generate a baseline scenario with a 3%-point increase in the real interest rate, due to the crowding out of capital. This contrasts with a fall in the interest rate of 1%-point that is predicted to occur if the social security system were to be gradually privatized.[12]

The pace of ageing is spread unevenly over countries, with OECD countries, Eastern Europe and China generally ageing faster than most non-OECD countries. This difference in age composition creates an opportunity for capital flows to the slow-ageing countries that may level most of the effect of ageing on factor prices. The effect of ageing on capital flows and world interest rates has been investigated in a number of papers. Brooks (2003b) obtains a fall of the world interest rate of 1%-point in a model without social security. This omission may not be crucial on a global scale. Domeij and Flodén (2006) did include a PAYG system in their model, and predicted a fall in interest rates of 1.5%-points over the next century and a half. However, Fehr et al. (2004), which includes capital flows among the US, the EU and Japan, reached essentially the same conclusion as Kotlikoff et al. (2001).

These differences in predictions depend substantially on the assumptions on saving behavior in developing countries. Fehr et al. (2005) extended the analysis of Fehr et al. (2004) by taking into account the impact of China on international capital flows. This reverses the interest rate scenarios of their previous paper. As China is ageing rapidly, and does not currently support a large social security system, Chinese savings are sufficiently large to compensate for the crowding-out effects that dominate factor price changes in

12 This scenario is indeed an important element of the 2006 Pension Protection Act, which provides for a massive shift to 401(k) plans.

the OECD. Without the effect of China on international capital flows, interest rates would increase by 1%,[13] while the inclusion of China changes the baseline scenario to a predicted *fall* of 1%-point. In addition, the fall in interest rates is predicted to be considerably larger if the OECD countries change their pension system – either by cutting replacement rates or by privatizing pensions.

Boersch-Supan et al. (2006) studied the impact of ageing on international capital flows, using a number of different capital mobility scenarios in which the rate of return on capital falls by about 1%-point. The fall in the rate of return is larger in the scenario where capital is mobile worldwide.[14]

The studies cited above do not consider macroeconomic risk. Hence, they cannot distinguish between bond returns and equity returns. However, the fundamental reason for a fall in interest rates due to population ageing can be expressed in terms of the ratio between accumulated net wealth and labor supply (i.e. the capital-labor ratio). The basic prediction is therefore that the rate of return on productive capital will fall. Although all other rates of return are linked with the marginal product of capital, they may be affected differently by ageing in a risky environment. Old households have a different portfolio composition than young households because their human capital hedge is much lower. As a result, population ageing *ceteris paribus* should result in lower demand for equity and a higher demand for low-risk assets. This shift in demand raises the

13 Another change with respect to the previous paper is that government investment is taken into account, which raises the saving ratio.

14 Krueger and Ludwig (2007) and Ludwig et al. (2007) extended the model of Boersch-Supan et al. with idiosyncratic productivity risk and mortality risk. This does not fundamentally affect saving behavior, and the interest rate predictions are very similar to those in Boersch-Supan et al.

equity premium (Brooks, 2000). The size of this effect depends on the size of the equity premium itself; in Brooks (2002) the effect is rather small.

One explanation for the equity premium that has a bearing on future expectations of the excess return on stock is the "junior-can't-borrow" argument of Constantinides et al. (2002). This states that the size of the equity premium is determined by the inability of the young to participate in the stock market. As a result, stock is held mainly by older workers, who demand a relatively high risk premium, because their human capital hedge is lower. If the number of older households increases relative to the number of young and middle-aged households, the reluctance of the elderly to hold equity will push up the equity premium. There are two versions of this story: the asset-market-meltdown hypothesis, which predicts a fall in stock prices (Poterba, 2001; Abel, 2001), and a predicted plunge of the risk-free rate (Brooks, 2003a). Brooks shows that the size of the change in the risk premium is dampened by the presence of a DB social security system, which creates a hedge for retired households and allows them to take on more risk. However, in his stylized model the risk-free rate still falls by 5%-points below its steady-state value around 2020, if households are borrowing-constrained – even with a DB social security system in place.

The empirical evidence for an effect of age structure on rates of return is mixed. Bakshi and Chen (1994) found an effect of about 60 basis points on the US risk premium following a 1% change in the mean age. However, Poterba (2004) found little empirical evidence of demographic structure on asset returns in the United States. Goyal (2004) found a positive effect of *changes* in the size of the 45-64 cohort on excess stock returns and a negative effect of changes in the size of the 65+ cohort. Ang and Maddaloni (2005) confirmed the result of Bakshi and Chen for the US, but they found that for most other countries the fraction of retired household has a negative effect on the equity premium. The effect is especially strong for countries with a large DB social security system. Summarizing, it seems that there is little discussion about the rate of return on productive capital falling as a result of population ageing. The "consensus" estimate is a fall of 1–1.5%-points in case of an increase in the dependency ratio of 30%-points.

3.3.4 Spillovers from productivity to asset market returns

Labor-augmenting technology shocks have a positive short-run effect on marginal capital productivity. For $\sigma \approx 0.5$ and $s_L \approx 0.7$ where s_L denotes the cost share of labor and σ the substitution elasticity, the elasticity is about 1.4. However, the net effect of TFP shocks on rates of return also depends on the labor response. King and Watson (1996) and Beaudry and Guay (1996) show that output shocks have a negative effect on real interest rates, with a half-elasticity of about −0.3. Galí (1999) claimed that TFP shocks in the short run lead to a reduction in hours worked, which severs any short-term link between technology shocks and capital returns. This argument was taken further by Beaudry and Portier (2005, 2006), who demonstrated that stock prices lead TFP growth by a few years, suggesting that news of TFP innovations precedes the actual increase in productivity. Their results are supported by Avouyi-Dovi and Matheron (2006), who showed that both for the US and the Euro area stock prices are negatively correlated with productivity at high frequencies, and positively at periods of between six quarters and eight years. For the US, these results are statistically significant (and less so for the Euro area).

It can be argued that the observed positive link at business cycle frequencies between productivity innovations and equity returns must also hold in the long run. Indeed, if we accept the constancy of labor's share in GDP as a stylized fact, then the long-run development of capital income and wage income must occur at the same pace. Along this line, Benzoni et al. (2007) argued that wage income and dividend income are cointegrated. Using an Augmented Dickey-Fuller (ADF) test, they obtained a t-statistic for the cointegrating vector of −4.1, with a coefficient of −0.26 over the period 1929–2004. This suggests that wage income and dividend income are strongly positively correlated at period lengths of four years and over. However, over the period 1947–2004, the ADF statistic drops to −2.4, which is not significant at 5%.

Bohn (2009) used a VAR model to estimate 30-year correlations between productivity and capital returns. He reported a positive correlation of between 30% and 66%, depending on the specification of the VAR and the cointegrating vector. In addition, the residual volatility in capital returns, conditional on productivity, is only a bit higher than that of production of productivity itself, after correction for a growth trend.

However, this correlation between stock returns and productivity takes two years or more to materialize. There is also empirical evidence that dividends and productivity levels are cointegrated. This implies that the relation between stock returns and productivity growth should also hold in the long run.

Section 2.4 describes how large negative shocks in GDP, or TFP, most often coincide with large negative shocks in stock returns. While there is no causal link from TFP to stock returns in this case, it is still useful to utilize the observed correlation between these two types of events in stochastic simulation of risk factors. The inventory made by Barro and Ursúa (2008), in table C2, suggests that stock prices fall on average more than GDP. We assume that a fall in GDP of 20% is associated with a fall in stock prices of 35%, on average.

We conclude that productivity shocks have a positive effect on contemporaneous *realized* excess capital returns, with a half-elasticity of around unity. They have a negative short-term effect on real interest rates, with a half-elasticity of around −0.3. Shocks to (unobserved) structural TFP growth have a positive effect on excess stock returns. However, this correlation between stock returns and productivity takes two years or more to materialize. There is also empirical evidence that dividends and productivity levels are cointegrated. This implies that the relation between stock returns and productivity growth should also hold in the long run.

3.4 Rare Disasters

Aggregate equity returns have an annual standard deviation of about 16%. A fall of the stock market of 50%, as happened in 2008, should occur once every 30,000 years if returns are normally distributed. The actual frequency of such crashes is, however, substantially higher.

Events like financial crises are commonly interpreted as a system failure, with nonlinear effects of the original shock. As a result, these 'rare disasters' cannot conveniently be modeled as a volatility spike. First, the effects of the shocks are invariably negative. Second, these effects appear to have much greater persistence than 'common' volatility shocks, which have a half-life of one year. Cerra and Saxena (2008) and Reinhart and Rogoff (2009) show that the aftermath of a financial crisis encompasses more than five years. Third, the size and duration of their effect is variable.

In terms of the goals of the present study, the issue is how rare disasters can be modelled in a convenient way. Barro (2006) and Barro and Ursúa (2008) modeled rare disasters in terms of a random event with low probability. In their dataset of a series of consumption and GDP disasters, due to financial crises, war, and so forth, disasters occur with a probability of about 3.5% per year, with a mean effect on GDP of 20%, and a mean duration of three years. In a vast majority of cases, equity prices fall during a disaster.[15] Conditional on a fall, equity prices drop by 30% on average. By including the possibility of consumption disasters in the stochastic discount rate, Barro (2006) obtained a realistic equity premium.[16] [17]

The notion of rare disasters is related to the discussion of unit roots in GDP in Section 2.1.1. The analyses of Reinhart and Rogoff (2008), Cerra and Saxena (2008) and Reinhart and Rogoff (2009) suggest that financial crises are a source of structural breaks in GDP growth. Barro et al. (2009) generalized the analysis in Barro and Ursúa (2008) by allowing for both permanent and transitory effects of disasters and for a spread of disasters over several periods. A disaster causes a permanent fall in consumption, plus 'undershooting' of the long-run outcome. They find that allowing

---

15 Chile, in the seventies, and Argentina, in the eighties, are the exceptions.

16 Note that in this "reduced-form" approach the number of risk factors does not increase. Instead, the systemic risk affects the shape of the risk distribution.

17 The notion of fat tails in asset returns also arises in Weitzman (2007), who points out that investors do not know the exact distribution of returns, and must estimate the relevant parameters from a finite sample. This typically results in t-distributions for the predicted returns, rather than normally distributed prediction errors. As a result, the equity premium may become arbitrarily large (Geweke, 2001).

for partial recovery after a disaster substantially raises the level of risk aversion needed to explain the equity premium.

Rare disasters occur with a probability of between 1% and 3%. They have an impact on consumption and GDP that is several times larger than that of standard business cycle risks. As a result, these risks have a great impact on risk premiums and rates of return. Rare disasters usually also have a large negative impact on stock returns. As a result, stock returns have a fatter lower tail than what would follow from a normal distribution. Systemic risk can be represented by endowing the relevant risk factors (productivity and asset returns) with a fat lower tail.

4. Simulation Results

The literature survey of the previous section outlined the state of the art on four important macroeconomic risks: demography, growth, financial markets and rare disasters. We also looked at causal links between these risk factors – for instance, how demographic developments influence the productivity and the rate of return on capital. Broer (2010) formalizes the insights of the previous section in a VAR model, using both own estimates and parameters values from the literature.

The drawback of summarizing the above review in a numerical model is of course that it downplays the uncertainty and discussion that is still ongoing in the literature. The formulation of a mathematical model necessitates a judgment call on issues that may not yet have been settled. A compounding problem is that the sheer size of the model makes it unwieldy to present a sensitivity analysis on every unresolved issue.

On the other hand, it may be hard to keep tabs on all of the practical implications of all aspects of the discussion without resorting to some sort of numerical aggregation. By putting the main insights and parameters into a model, it is easier to appraise which of the different risk factors is more likely to be important in practice. Ideally, the simulation study presented in this section should give the reader a feeling for the main risk sources of the coming decades.

This section presents a number of graphs that display the development of the distribution of several variables over time, based on 10,000 Monte Carlo runs. The graphs show the development over time of the median (the light grey solid line), the

95% interval (dark grey), and the 99% interval (light grey, again). Let us start by describing productivity risk.

4.1 Productivity

Figures 3 and 4 present the outcomes for the labor productivity process. In the long run, labor productivity growth is on average 1.9%. However, the recovery of productivity growth from the historically low rates around the turn of the century takes about two decades. First, during the next two decades the demographic shift lowers expected productivity growth by about 0.1%. Moreover, catching up with productivity growth in the United States takes time. Figure 3 shows that there is considerable uncertainty about the size of productivity growth in any given year. In addition, the lower tail of productivity growth is stretched out considerably, as a result of the occurrence of rare disasters, as discussed in Section 2.4.

With small probability, shocks of up to –10% may occur, and recovery takes time.[18] Shocks in productivity growth are persistent, as displayed in Figure 4. This characteristic of the model corresponds well with the increasing variance profile of productivity levels for the Netherlands (as presented in Section 2.1.1). As a result of the unit root in productivity, long-run productivity levels are rather uncertain. In the space of a century, productivity may be anywhere between the same and eight times the present level (implying average annual growth rates between 0% and 2.1%).[19]

---

18 These shocks may occur either in the Netherlands or, independently, in the United States. In the latter case, the flow of innovations from the US peters out for a few years.

19 The distribution is skewed towards lower productivity as a result of the rare disasters, as well as the log-linearity of the productivity distribution.

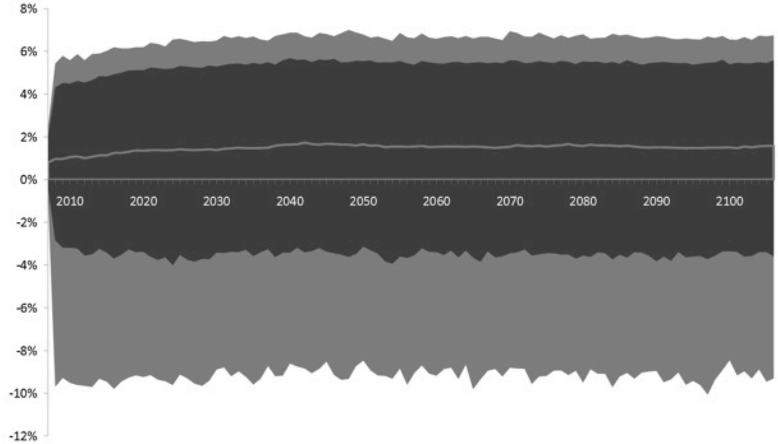

Figure 3: Labor productivity growth distribution. Graph shows the median, the 0.5%, 2.5%, 97.5%, and the 99.5% lines.

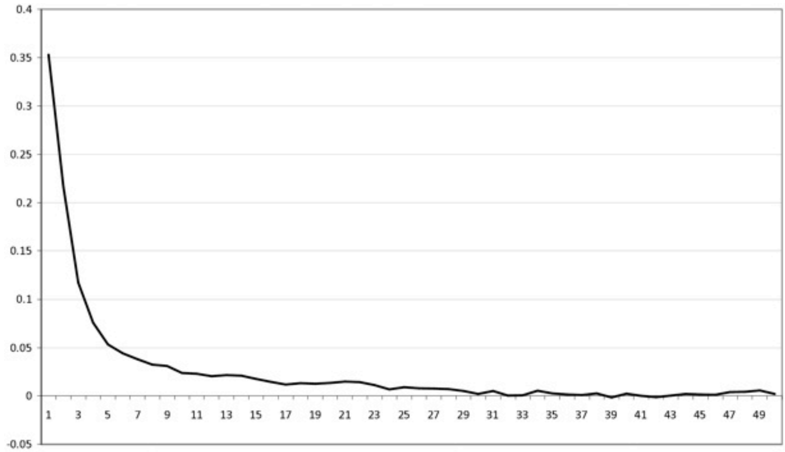

Figure 4: Autocorrelations in labor productivity growth

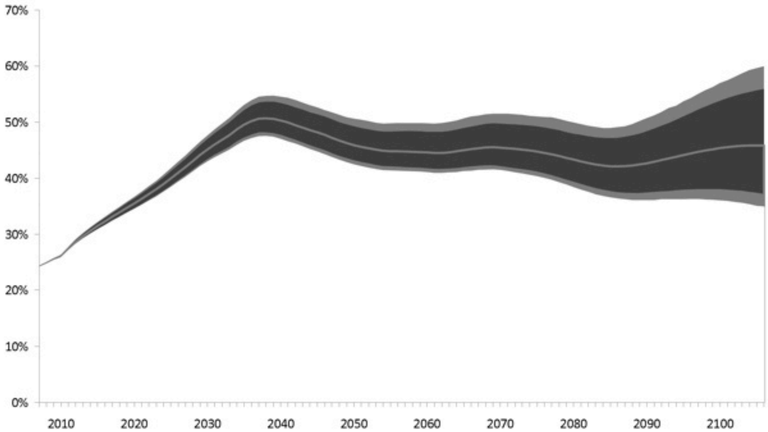

Figure 5: Dependency ratio distribution. Graph shows the median, the 0.5%, 2.5%, 97.5%, and the 99.5% lines.

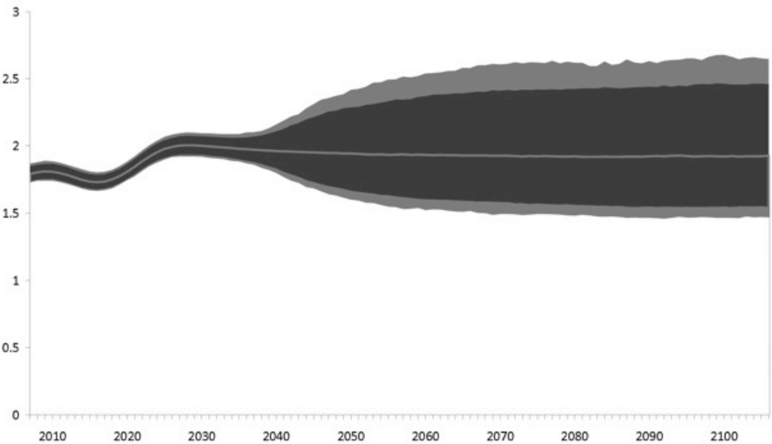

Figure 6: Total fertility distribution. Graph shows the median, the 0.5%, 2.5%, 97.5%, and the 99.5% lines.

4.2 Demographics

Figures 5–7 provide an overview of the uncertainty in demographics. Figure 5 shows that the dependency ratio will almost certainly increase in the medium run. The long-run development is less certain. Dependency ratio uncertainty increases slowly over the next thirty years, due to uncertain mortality and migration, as the future workers over that period have already been born. There is a boost in dependency ratio uncertainty when cohorts of currently unborn workers start to enter. A second increase in uncertainty occurs as these workers retire.

Figure 6 shows that total fertility may end up anywhere between 1.5 and 2.7 children per woman.[20] In addition, uncertainty is boosted by the effect of net migration. Strong productivity growth leads to a larger inflow of young age cohorts, which lowers dependency ratios somewhat. Our simulations suggest that migration will depress the dependency ratio by about five %-points.

The effect of future mortality rates on the life expectancy of a 65-year old is given in Figure 7. Life expectancy increases due to the decline in expected mortality. However, uncertainty is sufficiently large that, in the medium run, an increase in mortality is not inconceivable. In fifty years, life expectancy of a 65-year old may have risen to 90, but it may also have remained constant.

---

20 Total fertility is an artificial characteristic, as it is computed as a cross-section of fertility rates at a given point in time. However, in the long run this rate coincides with the completed fertility of a cohort, if the latter is stationary. The underlying uncertainty is the subject of Appendix A to Broer (2010), which shows the confidence intervals of the two state variables that determine fertility developments. The point is that there is substantial uncertainty in the *current* value of these state variables, because fertility behaviour of current cohorts has so far been only partially observed.

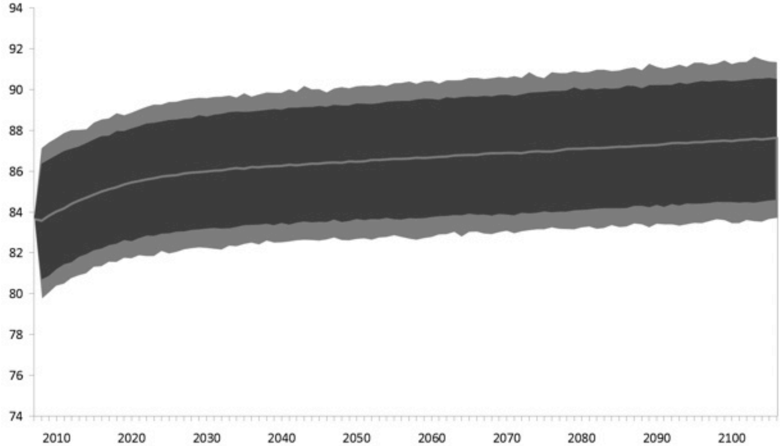

Figure 7: Expected life of people of age 65, distribution. Graph shows the median, the 0.5%, 2.5%, 97.5%, and the 99.5% lines.

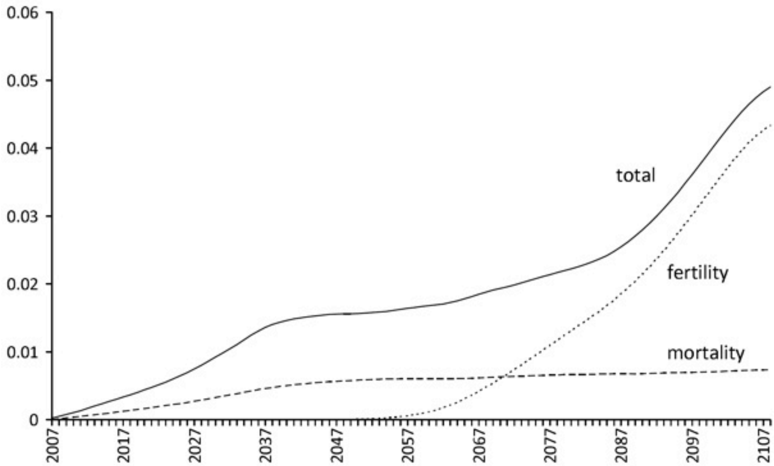

Figure 8: Standard deviation of the dependency ratio with all risks, and with just fertility- and mortality risk.

We assess the role that fertility and mortality play in the uncertainty surrounding the dependency ratio by using a method that we will employ several times in this paper.[21] For this method, we simulate future paths of the economy with only a single (group of) risk turned on, while the rest has their variance reduced to zero. In this case, for instance, we generate a number of model paths with only the risks that pertain to mortality turned on. We do the same for a version with only fertility risk active. In each of these simulations, we can retrieve a distribution for the dependency ratio.

Generally speaking, it is not true that the uncertainty in these sub-distributions sums to that in the main scenario. This is because of the nonlinearities involved in the model and in computing the dependency ratio itself. Nonetheless, it can be instructive to look at these sub-distributions to ascertain which factor is most important. This is done in figure 8, which plots the standard deviation of the dependency ratio in three cases: the main model with all risks active, and two subversions with only fertility- and only mortality risk. With only fertility risk, the first fifty years see virtually *no* movement in the dependency ratio. This makes sense, since it takes at least twenty years for children to even be counted in the dependency ratio, and the uncertainty in fertility sharply increases after about thirty years (figure 7). Up to that time, mortality risk and migration risk are the main drivers of uncertainty in the dependency ratio.

21 See, for instance, Section 6.4 below on the role of different risks in the variability of the internal rate of return of different pension schemes.

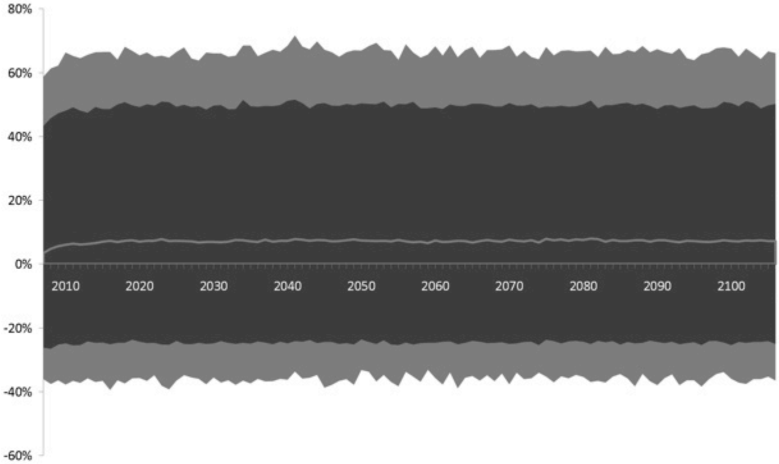

Figure 9: Real rate of return to equity distribution. Graph shows the median, the 0.5%, 2.5%, 97.5%, and the 99.5% lines.

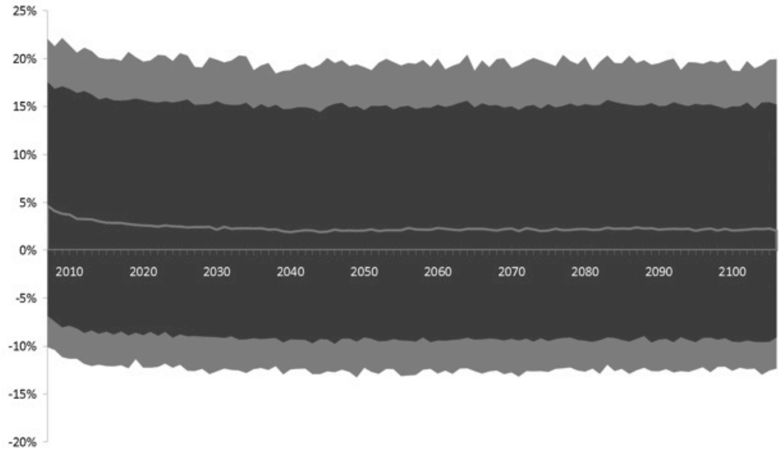

Figure 10: Real rate of return to bonds distribution. Graph shows the median, the 0.5%, 2.5%, 97.5%, and the 99.5% lines.

4.3 Asset Returns

In Figure 9, the real rate of return to equity shows considerable variation, at a standard deviation of 18.5%, a bit higher than in Campbell and Viceira (2005) because of the spillover from other risk factors, including the rare disasters. The large standard deviation of returns makes it difficult to discern from the figure that the median return on equity falls from 7.9% to 7.1% during the first 25 years. The mean return falls from 9.1% to 8.4% during this period, after which it stabilizes around 8.6%.[22] The rare disasters spread out the lower tail of the return distribution, but because the lognormal distribution is positively skewed, the effect of this on the graph is not conspicuous. The corresponding autocorrelation function (not shown here) shows only small negative correlations (that is, little mean reversion). This finding is also apparent from the annualized standard deviations of equity returns (not shown here), which indicate only a modest reduction in risk over longer horizons.[23]

The distribution of the real return on five-year bonds is similar to that on equity, with a slight decline in the rate of return in the first few years. The same comment applies to short-term bills. Real bond returns feature hardly any autocorrelation, but the return on short-term bills is correlated over a period of about five to ten years. The distributions of their returns are in Figures 10–11. Bonds converge to

22 The deviation between the mean and median originates from the (approximate) log-normality of the distribution of the rate of return on equity, combined with the large standard deviation.

23 This finding contrasts with the results presented in Campbell and Viceira (2005), who found a substantial fall in equity risk with increasing horizon. We were unable to reproduce their results, however: a numerical check of their variance profiles using their published parameter estimates yields a completely flat annualized equity variance profile.

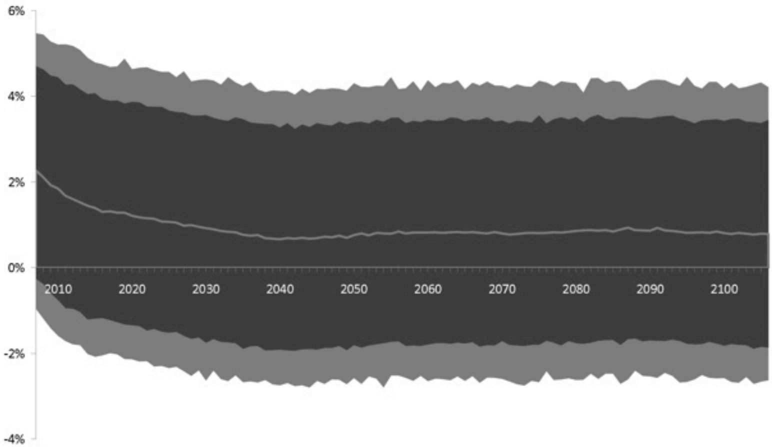

Figure 11: Real rate of return to one-year deposits distribution. Graph shows the median, the 0.5%, 2.5%, 97.5%, and the 99.5% lines.

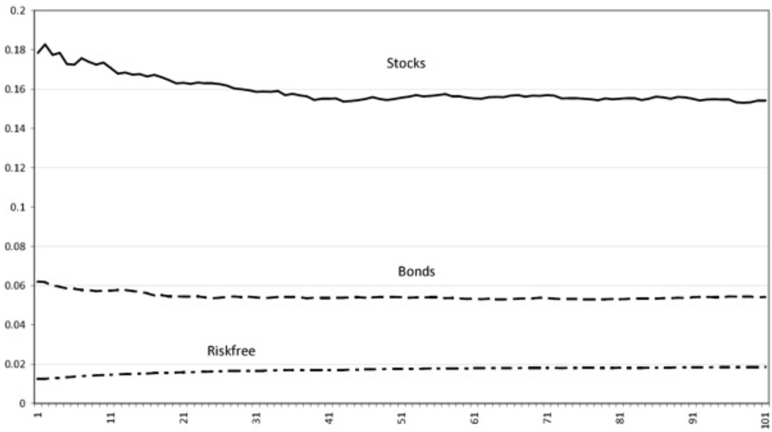

Figure 12: Annualized standard deviations on three assets. The period over which the standard deviation is computed is on the horizontal axis. A declining standard deviation indicates mean reversion over time.

a mean (real) rate of return of around 2.3%, with a standard deviation of 6.2%. Bills go to 0.8% average real return, with a 1.3% standard deviation.

The rate of inflation does not increase over time. Our calculation indicates only a minor upward drift. However, the margin of uncertainty is considerable. This is partly caused by the substantial persistence in inflation, which boosts inflation risk. As a result, the return on nominal assets is rather uncertain over long time spans. To compare the riskiness of the different assets over longer time periods, consider Figure 12, which shows the annualized standard deviations for the three assets under consideration. Even though there is some mean reversion in stocks – and therefore some reduction in the long-term variability – the order of riskiness of the three assets remains intact over longer time periods.[24]

In the model, stock returns and productivity growth are cointegrated (see section 3.3.4). We show the size of this effect by computing the correlation between log stock returns and log productivity growth over different horizons, a kind of correlation term structure. The solid line in figure 13 shows that this correlation is low at horizons as short as a year, and that it increases to about 0.18 at a 30-year horizon. In the figure, we also plot the correlation between log productivity growth and the log return to a portfolio of stocks, bonds and safe assets that is used in section 5.2 below. This

24 A quick scan through the literature shows that the time profiles of standard deviations are highly dependent on the data that are being used. The time profiles in Campbell and Viceira (2002) based on annual data, those based on quarterly data, and those in Kuipers (2008) all deviate from the time profiles in Figure 12. Moreover, in only one of the four datasets are equity returns as risky as nominal bonds at a horizon of less than 100 years.

correlation is slightly lower, as the higher stock returns do not lift the returns to the other assets on a one-for-one basis.

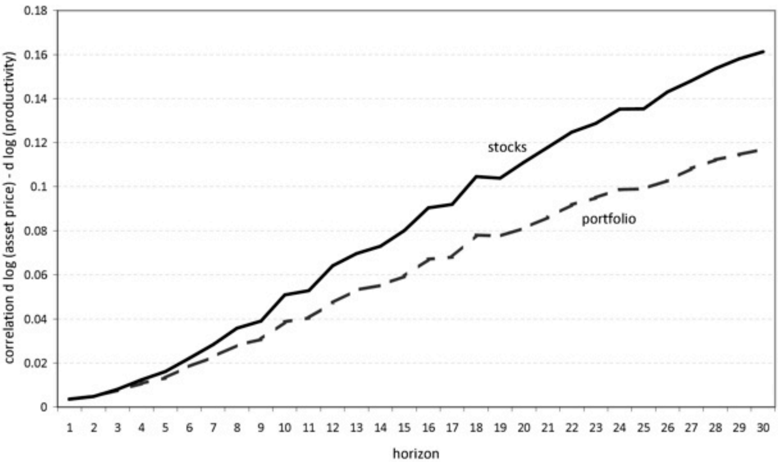

Figure 13: Correlation coefficient between the change in log asset price (including dividends) and the change in log productivity, where the change is computed over periods of different lengths. This horizon is on the horizontal axis, from one to 30 years. The solid line shows correlation between stock returns and productivity growth, the dashed line shows correlation between the pension portfolio return (see section 5.2) and productivity growth.

5. Macroeconomic Risks and Pension Schemes

Whether or not the goings-on at the financial markets have any effect on a particular individual's pension scheme depends on whether or not it is funded; more generally, the design of a pension scheme will determine to which risks it is most vulnerable.

Many different schemes are on offer in the pension world: both contributions to the pension scheme and benefits from the scheme can be contingent (or not) on prices, wages, individual labor supply, rates of return, and so forth. However, most of the schemes can be classified using just three dimensions (Lindbeck and Persson, 2003): funded or unfunded, DB or DC, and actuarial or non-actuarial. As the latter category mostly pertains to labor supply issues, we restrict our attention to the first two aspects. Indeed, these largely determine which risks play a role in a pension scheme and which of the participants will bear the brunt of the risks. In general terms, we can state that unfunded schemes will be susceptible to demographic and productivity risks, while participants in funded schemes should worry more about financial risks. But we can be more specific than that, given that we have parameterized the size of, and relationships between, all of these risks and can run simulations on them. These simulations will be the subject of the next section. This section proceeds by outlining the rules to the pension schemes that we employ, and explaining how the risk model helps in defining the forward-looking premiums of funded pensions.

5.1 Labor income

All pension schemes in this paper in some way take their cues from labor income. Premiums are a fraction of the labor income of the participant – and benefits, in some cases, are indexed to the wage level. We discussed developments in the economy's total factor productivity above, but the link between wages and TFP is not direct. Indeed, the variability in productivity likely overstates the variability in wages, as wages are thought to react sluggishly to innovations in productivity. A possible reason for this is that wages are the result of a central bargaining procedure that may smooth out high-frequency shocks over time.

It is also possible that productivity estimates themselves are influenced by the fact that the labor market is not a spot market. If firms decide to hoard labor in cyclical downturns, then perceived falls in productivity may actually be temporary decreases in effort, due to a lack of demand (see, for instance, Burnside et al., 1993). It is unlikely that workers will receive lower wages while labor is hoarded.

The CPB model SAFFIER (Kranendonk and Verbruggen, 2007, p.23) takes account of this phenomenon by using a two-step approach to wage formation. A long-term nominal wage w_t^* is computed as a function of prices, productivity and parameters from the labor market and the social security system:

$$w_t^* = p_t \cdot h_t \cdot f(u, r, \lambda) \ ,$$

with *p* the price level and *h* labor productivity. Unemployment *u* and the replacement rate *r* are not modelled in the current effort; the wedge λ is affected by pension premiums and may be relevant (but will be ignored for now).

The actual wage *w* only slowly converges to the long-term wage. It follows from

$$\Delta \ln w_t = \Delta \ln p_t + 0.34 \Delta \ln h_t - 0.4 (\ln w_{t-1} - \ln w_{t-1}^*) + g(\cdot) \ ,$$

with g a function of labor market and social security parameters. For the current work, we set g equal to zero. Note that in this formulation, price shocks influence wages one for one, but that only one-third of productivity shocks is absorbed in wages in the same year. The rest is taken in gradually at a 40% rate of convergence. Note that if shocks in h_t have a non-zero mean (as they do in the current model), the expected value for w_t will always be lower than w_t^*; that is, actual wages are expected to be below productivity.

Note also that labor supply, in our model, does not respond to these wages. We take data on current age-dependent participation and projections for future participation profiles from CPB (2006), and set labor supply exogenously. A graph of participation in 2009 appears in Figure 14.

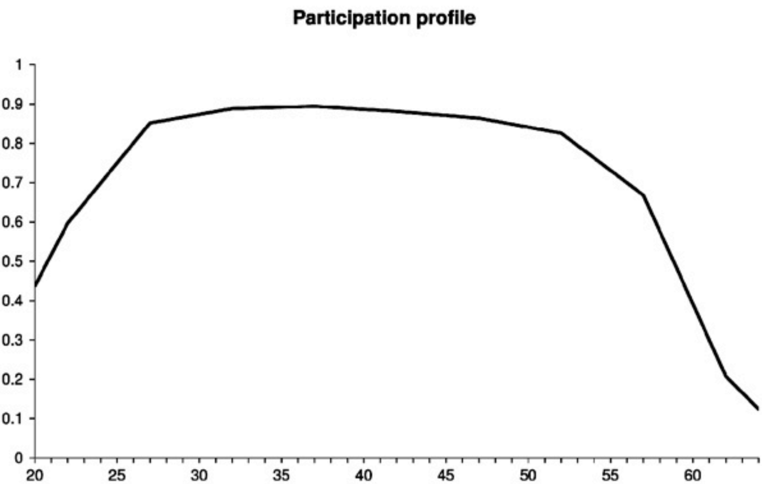

Figure 14: Age-dependent labor participation in 2009.

5.2 Rules for Pension Schemes

PAYG schemes take a percentage of the labor income of participants and redistribute this money to retirees. As these schemes are unfunded, they do not need to form an opinion on future events. Rather, they can set their premium rate or their replacement rate (which defines benefits) to balance their current budget. Schemes can be classified depending on the instrument they choose: in the DC case, the premium rate is fixed and benefits vary as a result of fluctuations in the dependency ratio. In the DB case, the replacement rate is set and the premium rate changes in response to demographics. We use these two simple variants as our stylized PAYG schemes; both of them start out with a replacement rate of 50% (that is, current retirees receive the wage equivalent of someone who works 0.5 FTE).

Funded schemes do not redistribute premiums immediately but rather invest the money. The payoff to these investments is used to finance future benefits, which may be indexed to wages. To even consider such a scheme, pension fund administrators must have expectations regarding future financial returns, and developments in both the mortality rate and the wage rate. In section 4.3, we show how consistent expectations for these variables can be derived from the same model that is used to simulate future developments in this exercise.

Taking as given that these expectations exist, we can compute the actuarial costs of a yearly annuity, starting at age 65 and continuing until death, for someone of age $x < 65$. This is simply the stream of payments, conditional on being alive, discounted with the expected mortality rate before 65 and the expected rate of return. We fix the portfolio decision of the pension fund so that it includes 50% shares, 40% bonds and 10% in the safe asset. On the basis of this portfolio, the actuarially fair price of an annuity can be computed for each age group between 20 and 64. Note that we do not change the weights in this portfolio, even though the age structure of the fund's members may change. This may lead the pension scheme to appear to be riskier than those that do change portfolio weights in response to their membership's age structure.

Exactly how large an annuity is offered per euro of premium paid depends on the rules of the pension fund. We work with a collective fund that uses an average premium approach, which means that the pension rights are allotted proportional to the premium paid, regardless of the age of the participant. While this is fair when seen over the span of an entire career, young participants get less than fair value, while participants close to retirement are being subsidized. Because we fix labor supply to the schedule in figure 14, we do not analyze how this distortion affects behavior.

At the time of payment, the participants thus obtain annuity rights that are equal, on average and in expected value, to the amount paid in. In the period between payment and benefits, however, the risks that we discussed earlier will generally open up a wedge between the value of the acquired rights and the value of the portfolio held by the fund. Rates of return may be higher or lower than expected, or mortality may be different than planned. The reaction to this difference in value again divides funds on the DB–DC dimension.

In the relatively simple world of DC, everybody's rights are immediately adjusted proportionally so that the total value of these rights is once again equal to that of the portfolio of assets. This means that retirees see their benefits change, while active workers perceive a change in their future benefits. Current premiums are only affected insofar as the shock has also changed expected values for the rates of return or mortality.[25]

With a DB pension scheme, acquired rights are defined in terms of future wage levels. Participants accumulate a fraction b over their career, and when they are retired a sum of b times the wage level is paid in a yearly annuity. The accumulation in b is proportional to the amount of premium paid in, and the coefficient of proportionality is such that in expected value, the acquired rights are worth as much as the amount paid in (again, this holds only for active participants as a whole, so subject to the average premium approach). But in contrast to the DC fund, once these rights have been accumulated, they do not change. Rather, any unforeseen change in the actuarial value of the accumulated rights (due to realizations in the mortality rate, the rate of return or the wage rate) is evened out by raising or lowering the premium rate on the active participants.

Different assumptions can be made about the speed with which a gap between assets and liabilities of the pension scheme will be closed. We have chosen to let the fund smooth out shocks over time, by only adjusting the gap between assets and obligations by 20% of its value in each period. In the long run, this is enough to keep the fund solvent, without leading to unrealistic variability in the premium rate.

The funded DB scheme is the only one of our four prototype schemes that bases its policies on the funding ratio. Our model

---

25 Note that this makes the fund not strictly DC, since what is defined is not the rate of contribution, but rather the pension buildup rate, or the amount of pension rights acquired per hour worked. When the price of these rights fluctuates due to changing insights into mortality or rates of return, the contribution rate reflects this.

assumes that the average rate of return on the portfolio acts as discount rate, similar to US practice but different from the Dutch case. We have refrained from mimicking the Dutch case as closely as possible. This would introduce so many institutional differences between the funded DB scheme and the other schemes, that a comparison would no longer yield any insight into the role of the PAYG/funding dimension and the DB/DC dimension.

There is another practical problem with funded DB schemes, which is that premiums have a natural lower bound of zero. Because in the case of a large, positive, shock to assets it is quite possible that the above rules specify a negative premium rate, we add an additional rule. If 20% of the gap between obligations and assets cannot be closed by lowering the premium alone, the premium is set at 0.1%.[26] The remainder of the 20% is then closed by increasing the accumulation coefficient for the rights that are being built up during this period. Thus, by lowering the price of pension rights, the effects of the measure accrue to the same group that benefits from the lower premium, and they are distributed in the same manner.

With this final rule, we have defined four "pure" pension schemes: PAYG DC, PAYG DB, funded DC and funded DB. Each of these schemes leads to a series of premium payments, followed by a series of benefit receipts. With the current rules, both of these series are stochastic in all of the schemes: they are subject to one or more of the risk sources detailed in Section 2. Together, after correction for discounting, these series determine the net benefits that a generation derives from participating in the pension scheme.

26 Keeping the premium rate positive is essential to ensure being able to compute an internal rate of return; see Section 5.2.

The question is how to best measure this benefit. We do this in two ways, by looking at the streams directly and by aggregating them in an internal rate of return. This approach is detailed in Section 4.4. First, the next section looks at the expectations that are necessary to compute the value of future pension rights, and whose existence we have so far assumed.

5.3 Valuing Future Benefit Streams

So far in this paper, we have surveyed the literature on macroeconomic risks and their relations, and laid down rules for four different pension schemes. The results of the literature survey have been formalized in a VAR model by Broer (2010), and we used this model to simulate paths for economic variables of interest. The ability of a VAR model to generate an unlimited number of paths for the future variables (taking account of mutual interdependencies) is one of its useful aspects. But there is a second benefit: at each point in every scenario, the model also specifies the expectations that hold for all future values of all variables (see Box 3).

Expectations on the future value of rates of return, mortality, and wages are necessary inputs to running a funded pension scheme. We thus see that the formalization of risk distributions in a VAR model helps us in this respect. We can get the expected future values of the rate of return on all different assets directly by employing the method in Box 3. Unfortunately, for some aspects of the model, getting the expectations straight from the data generating process proves to be unwieldy. This is the case for anything that involves the demographic sub model. In this model, the state vector consists (among others) of the sizes of all cohorts, making it over 120 variables long. The data requirements for keeping track of this vector are quite large, and we find that the alternative of using the current dependency ratio and past dependency ratios as explanatory variables works almost as well. As all of the influence of demography on economic variables of interest happens to run through the dependency ratio, we rather estimate a small AR(3) model on it. Specifically, we let the agent project the dependency ratio d_{t+1} according to

$$d_{t+1} = \phi_t + \psi_t^0 d_t + \psi_t^1 d_{t-1} + \psi_t^2 d_{t-2} \ .$$

The time-varying coefficients ϕ_t and ψ_t^i were estimated on the runs of the VAR model itself, so that the expectation includes predictable swings in the dependency ratio, while also taking account of the current state of demography. Expectations for d can, in turn, be used in expectations for economic growth and the return to capital. We use a similar approach (but with an AR(2) model) to predict the number of workers in future periods, which feeds into equilibrium rates of return. Future mortality rates (including their predicable decline due to a time trend) are extrapolated straight from the VAR model.

For the path of productivity we formulate a Kalman filter that accounts for the fact that observed productivity is different from the underlying state variable. Thus, the agent may adjust her estimate of the state variable in period t when new information becomes available in period $t+1$. We use a filter with a fixed Kalman gain matrix, so that we assume that the covariances of the errors in the system are known and do not have to be estimated (see, for instance, Welch and Bishop, 1995).

Box 3: Expectations in a VAR model

Writing the model as

$$x_t = Ax_{t-1} + b + \epsilon_t$$

for the state vector x, a matrix A, an intercept b and errors ϵ_t (the latter a vector with expectation zero), we can start with the current state vector and use our knowledge of *A*, *b* and the distribution of ε to generate a Monte Carlo path for the economy that specifies all values of *x* in the future. Additionally, at each future point t we can write the expected continuation of this path as

$$\mathrm{E}(x_{t+1} \mid x_t) = Ax_t + b$$

$$\mathrm{E}(x_{t+2} \mid x_t) = A\mathrm{E}(x_{t+1} \mid x_t) + b\text{ ,}$$

$$= A \cdot Ax_t + (A + I)b$$

and so forth.

5.4 Summarizing the Performance of Pension Schemes

Stochastic simulations produce an enormous amount of data. Even with our (cohort-specific) representative agents, there is a premium- or benefit flow for each year of birth, in each simulated year. To assess the variability of the streams in different pension schemes, we employ two approaches. The first is to select an age at which participants pay premiums, and an age at which they receive benefits (we pick 45 and 70). We then monitor the distribution of these two values over the projection period, with the idea that they are representative of the other premium- and benefit streams. This has the advantage that the variable that is studied is one that is readily observable in real life. Fluctuations in an amount of premium paid are easy to understand without much effort.

A second approach is to aggregate the entire set of premiums paid and benefits received over the lifetime in one measure. For this we use two concepts. The first is the internal rate of return (IRR). To compute the IRR of a series of payments and benefits, we solve for r in the equation

$$x_t + \frac{x_{t+1}}{1+r} + \frac{x_{t+2}}{(1+r)^2} + \ldots + \frac{x_{t+T}}{(1+r)^T} = 0 \text{ ,}$$

where x_t are net payments (when negative) or benefits (positive). Benefits are discounted with the (realized) rate of mortality. This is a polynomial equation of order T , and as such it can have multiple roots (that is, multiple solutions for r). However, if the series x_t changes its sign only once, then it has a well-defined single root; this is fortunately the case for all pension systems under consideration, in which an uninterrupted series of payments is followed by a series of benefits. The only exception is when the agent dies before reaching the age of eligibility, in which case there are only payments and r is undefined. We rule this out by computing the IRR conditional on the agent making it to the first year of eligibility. In fact, we use a representative-agent setup and compute one IRR for each cohort. We take the probability of living until age 65 as one, and from then on multiply the benefits payments with the survival probability for each year.

An alternative, but related, measure is the net present value of the pension scheme (NPV). It uses the same formula as the IRR,

$$NPV = x_t + \frac{x_{t+1}}{1+r} + \frac{x_{t+2}}{(1+r)^2} + \ldots + \frac{x_{t+T}}{(1+r)^T} \text{ ,}$$

with the difference that the interest rate is fixed. The number computed is the present value of payments and benefits, discounted by the probability of survival. The x_t s are again negative for premiums and positive for benefits. We use this measure when we look at diversification across different pillars in section 5.5 because it has the advantage of being a linear combination of all payment streams, rather than a nonlinear function.

Knell (2010) discusses the relationship between the IRR and the implicit tax rate, which is NPV divided by lifetime income. He shows that the properties of both measures in the face of demographic fluctuations are basically identical when the interest rate in the NPV calculation is equal to the growth rate of the economy. Note that the NPV in our case is a measure that is specific to a cohort; Fenge and Werding (2003), among others, discuss the concept of Net Pension Liabilities (also in present value) which is a measure of the pension system as a whole.

Note that we use a fixed interest rate to compute the NPV, rather than a stochastic discount rate that depends on the state of the economy. This implies that we do not take into account the insurance aspect of pension schemes, where income streams are valued more when they take place in times with high marginal utility of consumption. As with the IRR, our NPV measure does not value the variation in the payment series, but rather computes a kind of average. Our reasons for choosing this setup are pragmatic: the current concepts of IRR and NPV work without further assumptions, while a more elaborate appraisal of the pension schemes would need a fully specified model of the optimizing agent, her cash flows, consumption and her utility function. Though this is not impossible to do in principle, we opt instead for the clarity and intuitiveness of these simple measures.

A further problem with both of these measures is that we do not have historical payment data on people who have started paying into the pension fund before the first year of our simulations. Also, the shocks in their past are already realized, and do not contribute to the variability of the IRR and the NPV. In this paper, we therefore compute these measures only for people whose entire labor market career we can follow. That is, we present IRRs for generations born in 1987 and after, which means people who are not yet 20 when we start our simulations in 2007.

As expected, the IRR gives a quick and easy-to-understand look into the (cohort-specific) return characteristics of each type of pension fund. Both the level of the return and its variability can be inferred from our simulations. The measure therefore serves as a complement to the information on singled-out premium and benefit series.

6. Simulation Results for Pensions

This section runs four different pension schemes through a stochastic simulation and summarizes their performance using readily observable premiums and benefits and a measure of the return to investment. It is important to note from the outset that there is no channel of causation running from these pension schemes back into the economy. We assume fixed labor supply and do not let anything that happens to premiums or benefits affect the VAR model that gives us the simulated path we are using. The simulated paths are an input to this exercise and do not change by anything we do here.

The rules for the four different pension schemes were discussed at length in Section 5. We set the replacement rate of all schemes at 50% of a full-time wage in the first period. Funded schemes start with their current members having accumulated rights as though the participation rates had remained constant over the past 45 years, and participation with the pension scheme was constant. The accrual rate of funded schemes is 1.8%, which leads exactly to the desired replacement rate, using the participation profile in figure 14.

Stochastically simulating the different pension schemes allows us to approximate the distribution of many of their characteristics. While these characteristics, such as premium rates and benefit levels, do not directly tell us anything about the welfare levels of agents in the economy, they may be of intrinsic interest to policymakers and oversight authorities. We graph the distributions of several characteristics in this section. The number of Monte Carlo runs is 10,000.

6.1 Pension Policy Variables

For the PAYG scheme in which a (defined) benefit is linked to the wage level, the parameter that is uncertain is the premium rate. It depends entirely on demographic variation, as changes in the productivity level affect both premiums and benefits in the same manner. The distribution of the premium in the DB–PAYG scheme is given in figure 15. Note that demographics force an upward trend in the median, but that the premium stabilizes after 2040. The uncertainty gets larger over time as demographic uncertainty increases.

The second variable, the base premium of the funded pension scheme, is rather more stable. This is the premium rate that represents the value of new pension obligations in the year they are incurred. The expected value of mortality rates is consistent with the model, and some decline is expected over the long periods that are covered by the computation. When people pay a fraction τ of their wage income into the pension fund, on average they are receiving pension rights of the same value in return. The actuarial calculation behind these rights is driven by demographics and by the market for financial assets. Of course, the averaging of the value of pension rights over different generations does mean that the older employees underpay, while the younger employees get less than they pay for. The base premium for funded pensions is in figure 16.

The median of the base premium rate for funded pensions is observed to increase in the first decades of the simulation. This is the result of two things: a decrease in expected mortality and a decline in expected interest rates. The latter follows from the deterioration of the dependency ratio (see section 3.3.3; a higher dependency ratio decreases interest rates both directly and through a retarding effect on productivity growth).

The calculation of the base premium rate is based on the fund's expected rate of return, which includes not only the interest rates on bonds and safe deposits, but also the return to stocks. While the first two are expected to decline (figures 10, 11), the latter increases initially (figure 9). The combined effect is a mild drop in the expected return to the entire portfolio. This is in the range of 0.2 %-points from 2010 to 2040.

How much of the increase in the premium rate can be attributed to the decline in expected returns? Between 2015 and 2035, the median base premium rate for funded pensions rises by exactly one %-point. About 40% of this increase is caused by lower expected rates of return, the rest by decreased expected mortality.

Note that the uncertainty surrounding premium rates for funded pensions is much higher than that surrounding the base rate for PAYG pensions, particularly in the first few decades. The actuarial computations behind the former are influenced by the movement of return variables on financial markets. These financial market variables are much more uncertain in the short term than the demographic variables that determine the uncertainty surrounding the PAYG premium rate. In the (very) long run, however, the two uncertainties are of comparable size (from today's perspective).

For the funded pension scheme, we calculate the coverage ratio as the value of their assets divided by the present value of their obligations. The obligations are understood to include indexation to the economy's real wage increases. The coverage ratio fluctuates in bands, as its value will be the main driver of pension fund behavior. It is in fact possible to describe the fund's actions as an effort to stabilize the coverage ratio. Figure 17 shows the

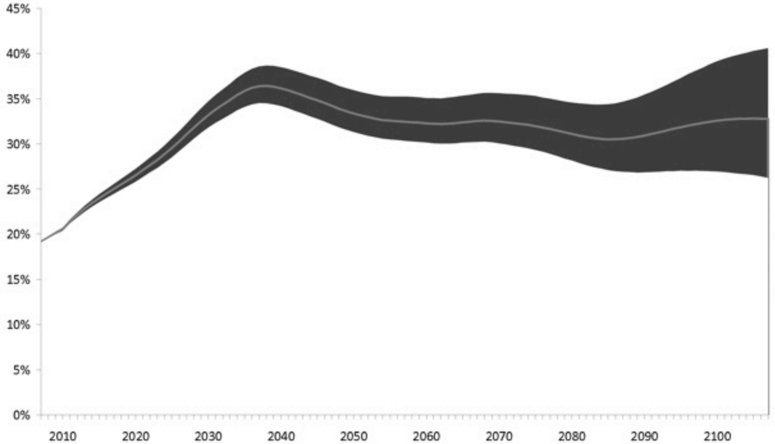

Figure 15: The simulated distribution of the premium rate for DB PAYG over time: the median, the 2.5% and 97.5% lines.

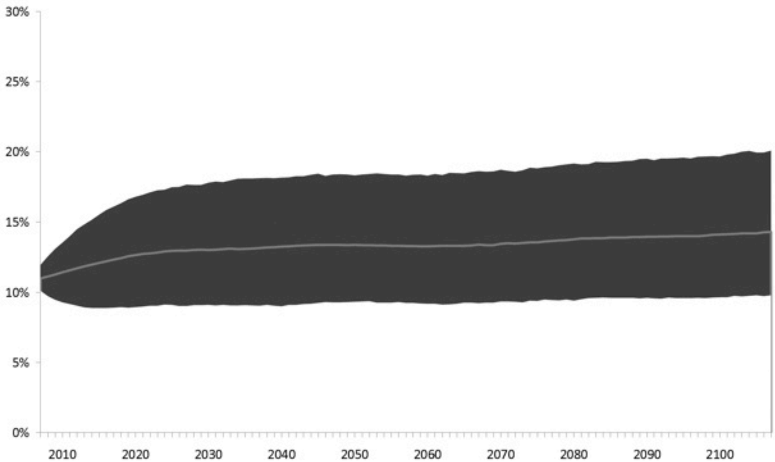

Figure 16: The simulated distribution of the funded base premium rate over time: the median, the 2.5% and 97.5% lines.

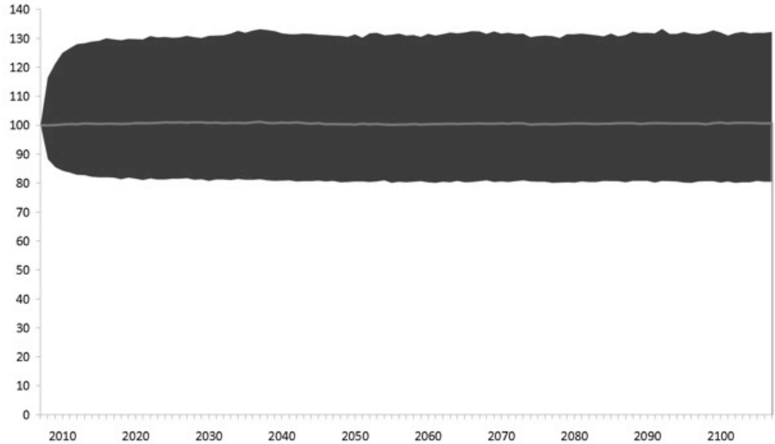

Figure 17: The simulated distribution of the real coverage ratio of the DB pension fund over time: the median, the 2.5% and 97.5% lines.

distribution of the real coverage ratio for the funded DB scheme. This is the only interesting case, as the funded DC scheme immediately adjusts rights to set the coverage ratio back to 100. Pension funds start at a real coverage ratio of 100% in 2007. Very quickly after that, the bands of uncertainty widen.

Finally, we look at the remaining two instruments of the pension funds. The variable indexation parameter is used to cut or increase pension rights, depending on the state of the coverage ratio in DC funded systems. When this parameter is one, pension rights are increased by exactly the (nominal) wage increase. When it is below one, existing pension rights decline in value; when it is over one, they increase. Figure 18 has the distribution.

In funded DB systems, the indexation parameter is constant at one but the premium is variable. Extra premium gets added to the base premium level of figure 16 (above) if the fund suffers from unforeseen setbacks. When results are better than expected, the

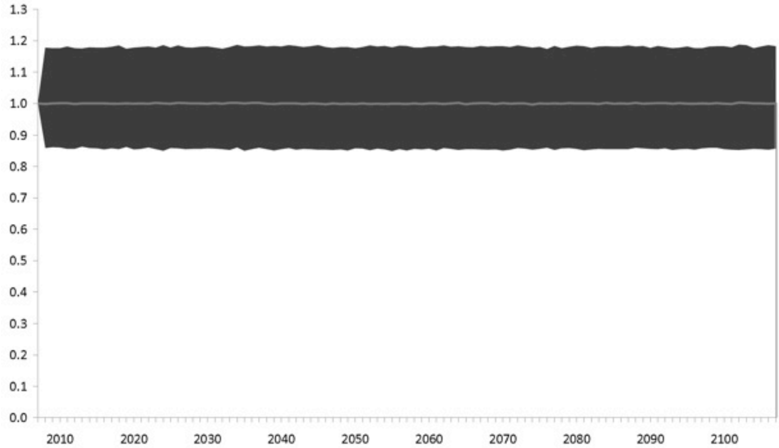

Figure 18: The simulated distribution of the benefit indexation of the DC pension fund over time: the median, the 2.5% and 97.5% lines.

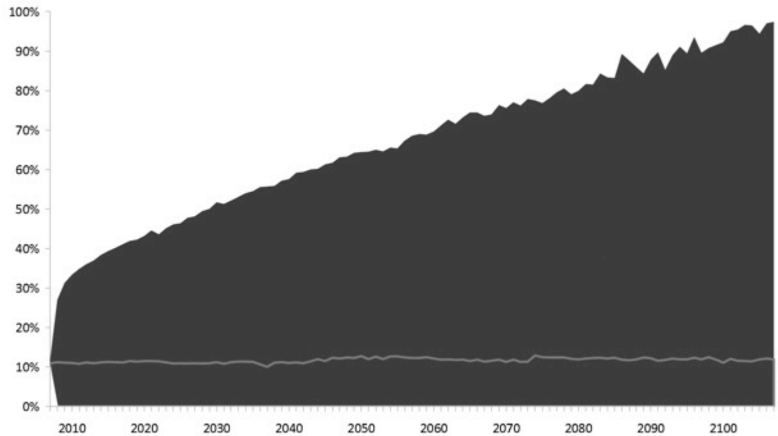

Figure 19: The simulated distribution of the total premium rate of the funded DB pension over time: the median, the 2.5% and 97.5% lines.

premium rate drops below the base level, with a minimum value of 0.1%. The distribution is in figure 19. Note that the uncertainty bands of this distribution widen over time. This is a consequence of the shrinking of the premium base that is due to population ageing (as reflected in the increase in the dependency ratio; see also figure 3). Using the premium paid by active members to compensate shocks in retirees' rights means that the dependency ratio matters for the size of the premium adjustment.

What the funded pensions are providing is a sharing of risks between generations. Shocks on the financial markets do not affect the pension benefits of those that have already retired, unless they lead to a severe deterioration of the coverage ratio. In principle, this should make the agents in the model better off. The performance of the two pension schemes is evaluated in the next two sections.

6.2 The Distribution of Premiums and Benefits

In the sections below, we summarize the effects of pension systems over the life cycle by calculating the internal rate of return. This section, however, briefly sketches the distributions of some easy-to-observe quantities, such as the premiums or benefits paid or received at a specific age. While these numbers naturally paint an incomplete picture of the total pension system, they do offer some advantages: since they have real-world counterparts with which the reader may be familiar, there is no issue of interpretation that may cloud the insights of the model. Also, our aggregate measures of the following sections necessarily look at the entire lifetime of a person. But to the current 70-year old, the level of benefits that somebody of that age may expect is probably of much more interest. Finally, for generations that have started paying into a pension system before the first year of our simulations, we have had to approximate the premiums paid and rights accumulated in earlier years. This involves certain assumptions, and to keep these out of our analysis we generally look at cohorts that start working after the first year of our simulation. The distributions in this section, however, do involve people over the age of 20 in the first year of simulations.

Figure 20 shows the distributions of the premium rate in the four schemes with the median, the 5% and the 95% line. All of these systems start out at the same replacement rate, but over time drift apart in this respect. The DC schemes generally keep the premium rate around the same level, whereas DB schemes do the opposite. The fanning out of the 5%- and 95%-lines indicates the increasing uncertainty (from the point of view of the starting year of the simulation) about the premium rate. Apart from the difference in DC and DB, we can see that the premium rate under funded systems is more uncertain, but lower. Uncertainty is especially striking in the case of funded DB, where unexpected shocks are wholly absorbed by the premium rate. But in funded DC, there is also uncertainty about future premium rates. This is caused by the uncertain changes in expected rates of return and mortality. These affect the base rate that is charged for a funded DC pension (see also figure 16). The uncertainty in PAYG premiums comes wholly from demographics. We see that this is accumulated uncertainty: in the short run, the premium amount is perfectly predictable, but the bands widen for the further future.

Figure 21 shows the distributions of the replacement rate at 70 years of age in the four pension systems. Note that DB actually means a fixed replacement rate in the PAYG case. As explained in section 4.2, in the case of funded pensions this is not true. When a shock arrives that would call for a negative premium, extra pension rights are accumulated instead. This makes the replacement rate a

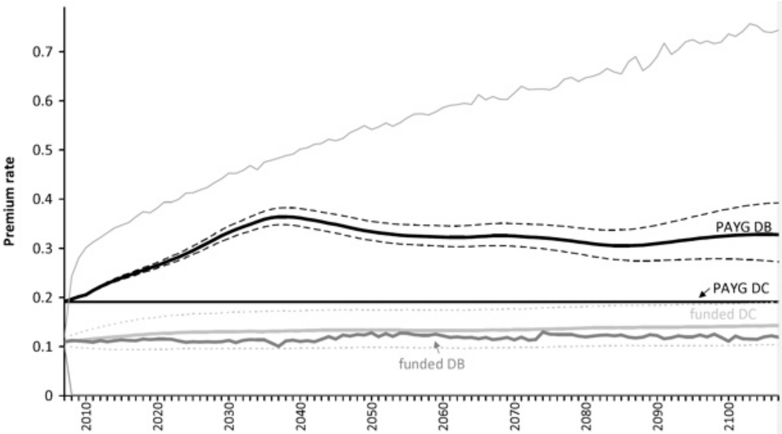

Figure 20: Distributions of the premium rate in the four pension systems. For each system except PAYG DC, this figure shows the 5% line, median and 95% line. PAYG DC just has the median. On the horizontal axis is the year.

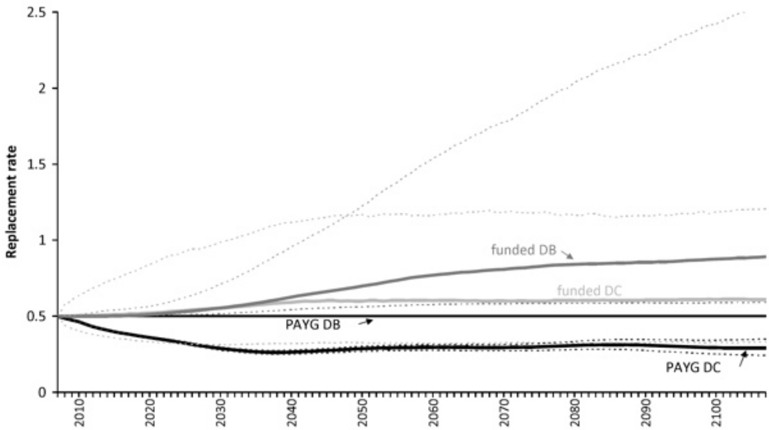

Figure 21: Distributions of the replacement rate at age 70 in the four pension systems. For each system except PAYG DB, this figure shows the 5% line, median and 95% line. PAYG DB just has the median. On the horizontal axis is the year.

function of past shocks, which explains the non-trivial distribution. It is, however, true that the replacement rate of DB funded pensions never falls below 0.5, since only positive shocks affect the accumulation of pension rights.

6.3 The Internal Rate of Return

As discussed in section 4.4 above, premium rates and replacement ratios give only a partial picture of the pension system. To give a broader picture, we complement this now with information on the internal rate of return.

Elements of the distributions of the IRR in a pay-as-you-go scheme and in a funded scheme are in figures 22 through 25. The first generation that we follow from age 20 on is the cohort born in 1987. Earlier cohorts enter our model when they have already spent some time on the labor market, which makes it hard to compute the IRR over the entire pension career.

The figures indicate that even though their size becomes very different over time, the two PAYG schemes show very similar IRRs, in terms of both the median and the distribution. This is hardly surprising, since (as we shall see in section 6.4) the relevant shocks affect the returns to these schemes in a very similar way.

Funded schemes offer returns that are both higher on average and more uncertain. It is instructive to note that the IRR characteristics of the funded DC scheme are exactly the same for the generation born in 1987 as for that born sixty years later. This indicates a lack of predictable developments in its determinants, which means mostly the returns on financial markets. On the other hand, we can see a small increase in the variance of the IRR of the funded DB scheme, which has to do with the changing dependency ratio. This ratio is important because mismatches between the value

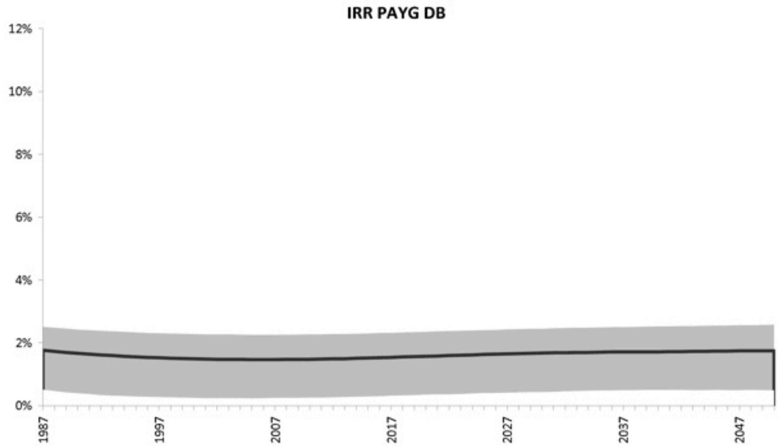

Figure 22: The internal rate of return to a DB PAYG system, median, 5% and 95% lines. On the horizontal axis is the year of birth.

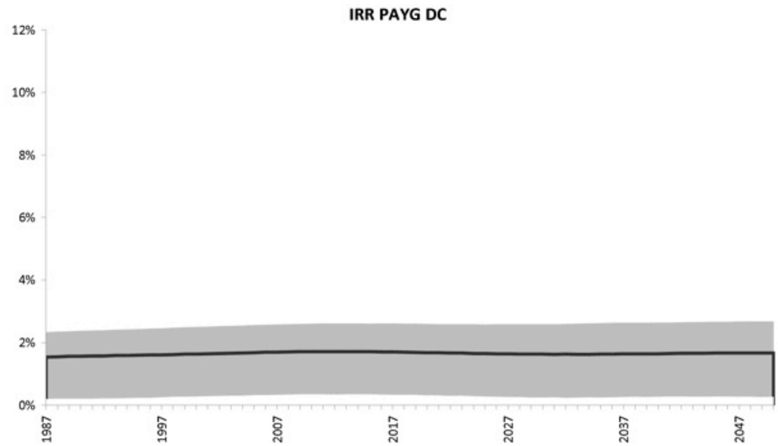

Figure 23: The internal rate of return to a DC PAYG system, median, 5% and 95% lines. On the horizontal axis is the year of birth.

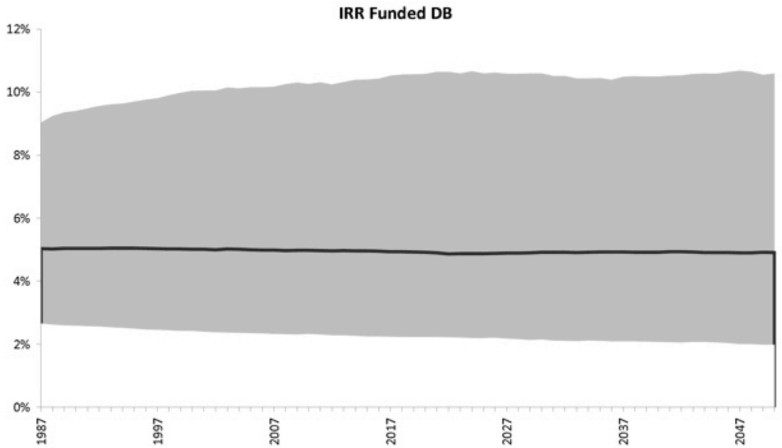

Figure 24: The internal rate of return to a DB funded system, median, 5% and 95% lines. On the horizontal axis is the year of birth.

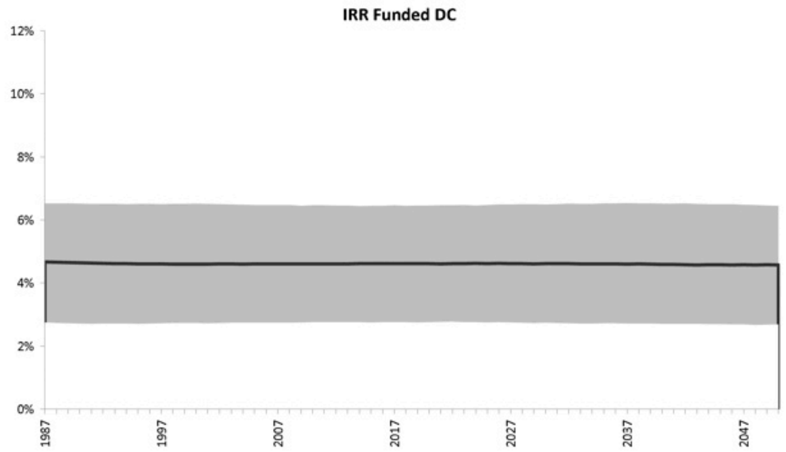

Figure 25: The internal rate of return to a DC funded system, median, 5% and 95% lines. On the horizontal axis is the year of birth.

of assets and obligations are borne by workers only.

What is more striking is the size of the risk spread in figure 24 (that of IRRs to the funded DB scheme), compared to the other schemes. What explains the larger variance in the rates of return to this scheme? There are two main reasons. Firstly, the funded DB scheme concentrates participants' risks in the period in which they work. In a funded DC scheme, members are subjected to financial risks during their retirement as well. Because of this, the IRR in DC systems is the average of a larger number of risks, which decreases its variance. Secondly, the DB funded pension scheme is the only scheme that is susceptible to both financial risks and to demographic risks. The dependency ratio is important in this scheme because the risks on all of the fund's obligations are borne by just the workers. This leverages the exposure to financial risk.

Interestingly, the variation of the IRR in the funded DB scheme is inversely proportional with the amount of smoothing that the fund itself carries out. In these simulations, we assume that only 20% of the difference between obligations and assets is eliminated each year; increasing this to 40% causes a drop in the variation of the IRR. A reduction to 10% increases the variation in IRR.

6.4 Separating out sources of risk

In our stochastic simulations there are four clusters of variables that can be thought of as distinct sources of risk. They are demographics, productivity, financial markets, and systemic risk. In VAR terms, every state variable in each of these clusters is subject to a shock (or rather an error term) in each period, which immediately impacts on its value. How this shock then propagates through the economy is implied in the parameters of the model, of which section 2 gives an overview.

In this section we observe the influence of the four different sources of risk by running four alternate sets of simulations in which only one of the sources of risk is active. For instance, we can set the variance of the shocks to productivity and the financial markets to zero while leaving that of the shocks to demographics at its estimated value. After simulating a large number of paths we observe how the economy and our pension schemes behave in this hypothetical situation. Note that in the above example, it is generally not true that the variance in the un-shocked variables themselves turns to zero. Via the many spillovers discussed in section 2, shocks in any one cluster have an impact on the other two as well. But by isolating the different sources we can trace the beginning of every shock and thus find out what kind of uncertainty is dominant in our economy's future, and against which kind of uncertainty each pension scheme offers protection.

Table 1 shows what this means for the IRRs of the four pension schemes discussed so far. The table lists the distribution of the internal rate of return to a pension fund for someone who enters the labor market in 2007. For each pension system, there are five rows: one with all risk sources operating simultaneously (this is the benchmark case) and four rows in which only one risk source is active, while the variance of the other risk sources is set, counterfactually, at zero.

The Table shows how different kinds of pension systems are susceptible to different kinds of risks. Financial risks dominate the variation in returns to funded systems, but play no role in the returns to unfunded systems. Demographic risks turn out to be relatively unimportant for all kinds of pension schemes. Indeed, the main drivers of variation in the returns to PAYG systems turn out to be fluctuations in productivity and, especially, the risk of rare disasters. The effect of disasters works through productivity, of course, but as a source of uncertainty it is modeled separately. Note also the asymmetric nature of the distribution in the case of disasters.

It is also clear that the different kinds of pension schemes react differently to various sources of risk. This begs the question whether it is possible to generate gains from diversification by using two types of pension schemes in a hybrid combination. This will be the subject of the next sub-section.

	DB funded				***DC funded***		
	2.5%	*median*	*97.5%*		*2.5%*	*median*	*97.5%*
All risks	2.5%	5.4%	10.6%		2.2%	4.6%	6.8%
Financial	2.9%	5.8%	11.3%		2.8%	4.9%	7.1%
Productivity	4.6%	5.1%	5.8%		4.3%	4.6%	4.8%
Demography	5.1%	5.2%	5.3%		4.6%	4.6%	4.7%
Disasters	3.3%	5.0%	5.4%		2.9%	4.4%	4.7%
	DB PAYG				***DC PAYG***		
All risks	0.3%	1.7%	2.6%		-0.1%	1.5%	2.4%
Financial	2.2%	2.2%	2.2%		1.9%	1.9%	1.9%
Productivity	1.6%	2.1%	2.7%		1.3%	1.9%	2.5%
Demography	2.0%	2.1%	2.2%		1.8%	1.9%	2.1%
Disasters	0.3%	1.8%	2.3%		0.0%	1.6%	2.1%

Table 1: The internal rate of return (median, and two quantiles) to the four different pension systems when different risk sources are the only ones active. The year of birth is 2007 for all figures in this table.

6.5 Pension Diversification

Up to now, we have worked with four "pure" combinations of DB, DC and funded, pay-as-you-go pension systems. In many countries (including the Netherlands), actual pension systems are usually a hybrid combination of these extremes. In the usual parlance, pension provision is offered in multiple *pillars*, each providing a share of pension income, and each built using different principles. No doubt, the idea is that such diversification offers a form of guarantee against the kinds of risks that have been described in this paper.

Assessing whether this diversification actually works is a complicated business, even with the tools and insights we have developed so far. The theory of finance that is usually employed to assess the risk-return tradeoff assumes that assets can be traded continuously and that prices adjust in the market. With pension schemes, however, such tradability of rights does not exist. Often, there is not even a way to marginally adjust the holdings of the pension "asset". It is thus incorrect to use the return characteristics of the different pension schemes as though they belonged to a tradable asset.

We can, however, use our current data to construct hybrid forms of the four "pure" pension schemes as a guide to which hybrids would be worth looking at. The idea is outlined in figure 26. The two dimensions of pension provision that we have looked at, and their four combinations, are depicted as sides of a square. These corners can be combined two by two to form four hybrids, indicated by the dots on the sides of the square. They are either combinations of DB and DC, both funded (DBDCF) or unfunded (DBDCP), or they are combinations of funded and PAYG, both DB (DBFP) or DC (DCFP). In all cases, we add the premium payments of

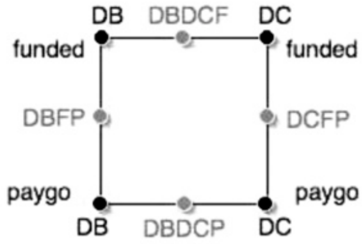

Figure 26: The four "pure" pension schemes and the four hybrids.

the two constituent schemes and divide by two, and do the same with the benefit payments.

For each of these four hybrids, we compute the net present value (NPV) of each of the schemes, which is a concept similar to IRR, but based on a fixed interest rate (see section 5.4). The advantage of the NPV over IRR is that it is a linear measure (that is, the weighted combination of two schemes has an NPV that is the same, weighted, combination of the two individual NPVs). This allows the use of simple statistics to assess diversification, which is not possible in the case of IRRs (see Box 4).

	Correlation NPV pension scheme	
	Birth year 1987	Birth year 2037
PAYG DB+DC	0.993	0.934
Funded DB+DC	0.591	0.501
PAYG + Funded DC	0.389	0.466
PAYG + Funded DB	0.372	0.325

Table 2: Correlation rates between the net present value of pension schemes, for two cohorts. A low rate of correlation suggests a high potential for diversification.

Box 4: The IRR of weighted pension schemes

Computing the IRR of a series is a nonlinear transformation, and as such it should be no surprise that the IRR of a linear combination of pension schemes does not equal the same linear combination of the original IRRs. In general, it turns out that the IRR of combined schemes is *higher* than the (weighted) average of the constituents. This can easily be seen for the following simple example:

Suppose we have two series $\{x_1, x_1, x_2\}$ and $\{x_1, x_1, x_3\}$ for which the following holds:

$$-x_1 - \frac{x_1}{1+r_A} + \frac{x_2}{(1+r_A)^2} = 0 \text{ and}$$

$$-x_1 - \frac{x_1}{1+r_B} + \frac{x_3}{(1+r_B)^2} = 0 .$$

Clearly, r_A and r_B are the IRRs of these two series with x_1 the premiums and x_2 and x_3 the benefits. Now take the average of the two schemes, $\{x_1, x_1, (x_2 + x_3)/2\}$. If we call the IRR of this series r, it turns out that

$$(1+r)^2 + 1 + r = 1 + \frac{1}{2}(r_A + r_B) + \frac{1}{2}\left((1+r_A)^2 + (1+r_B)^2\right).$$

We can now show that r must exceed the average of r_A and r_B: write $r_B = r_A + 2\Delta$ and $r = r_A + \Delta$. Then the left-hand side of the above equation is equal to the right-hand side minus a term Δ^2. Of course, our pension- and benefits series are much longer, but this only increases the importance of the nonlinear terms. This is why simple rules of variance do not apply to IRRs.

A numerical way to assess the benefits of diversification is to look at the correlation coefficient between the NPVs of two different pension schemes (across different sample paths). The lower this coefficient, the higher are the potential for gains from diversification. Table 2 shows these coefficients for two different birth years, one cohort entering the labor market right at the beginning of our simulations, one entering fifty years later. The table suggests that the most valuable diversification happens along the PAYG-funded axis. On the DB side, the diversification (from today's point of view) gets more valuable for future generations.

To get a sense of scale, it may be useful to compare the size of the pension system's NPV to the benefits of that system. This can be done by valuing the NPV at the moment the agent turns 65. At that moment, he has on average 21 years left to live. We annuitize the NPV for that period against the same interest rate that was used to compute it, 0.8%, to get a sense of the equivalent annual amount, and express it as a percentage of the (expected) benefit amount.

So do these hybrid pension systems offer an advantage in terms of diversification gains? We can compute the theoretical standard deviation of the payoffs in the case of *zero* diversification gains. Comparing the *actual* sigma of the hybrid pension scheme then gives an indication of the size of these gains compared to the average benefits of the scheme. For people born in 1987, the results are shown in table 3.

In Table 3, the first column gives the standard deviation as a percentage of the expected pension benefit. This confirms that funded pensions are riskier than the unfunded, at least at this horizon. The next column gives the theoretical standard deviation of the hybrids, conditional on the fact that these hybrids offer no diversification advantage. The actual standard deviation follows,

	sigma		max sigma	sigma	Diversification gain
DB funded	81%	DBDCP	25%	25%	0%
DC funded	72%	DBPF	53%	48%	6%
DB payg	23%	DCPF	56%	51%	5%
DC payg	27%	DBDCF	76%	68%	8%

Table 3: Mean and standard deviation of the net present value of four pure and four hybrid pension schemes, annuitized at 65 and expressed as a percentage of the expected benefits.

and the difference between the two is in the last column. We observe that there exists no useful diversification between the different unfunded systems. The highest gains are between the two DB pension schemes and between the two funded schemes.

As regards the gains from diversification, we find implicit correlation coefficients that are quite large: between 0.5 and 1. With nigh perfect correlation between the internal rates of return of different pension schemes, the risk reduction that can be achieved by mixing different pension schemes seems to be fairly small.

7. Conclusion

This paper takes stock of the available empirical information about the size of the main risks associated with the returns to social security (viz. demographic risks (fertility, mortality, migration), productivity risk, asset return risks (inflation, bond returns, and equity returns) and systemic risk, i.e. the risk of a rare event (with probability between 1 and 3%)) in which extremely large drops in productivity and equity prices coincide.

We find that the systemic risk is an important risk factor in terms of future consumption levels. Productivity in normal times, i.e. when there is no rare event, is similarly very important. As a rule-of-thumb, the latter risk increases by about 2% per year. There is little evidence of a deceleration of the rate of growth of this type of productivity risk over time.

Demographic risks are found to be modest in terms of their effect on the old-age dependency ratio. The dependency ratio does show substantial fluctuation in the next decades, but the larger part of this fluctuation is predictable. Asset return risk declines with the length of the investment period as a result of mean reversion. We find the effect to be limited, however, and extending only for a couple of years.

Overall spillovers between different risk factors are of little importance. The most significant is the spillover from demographic shifts to financial markets. Indeed, due to ageing, the rate of return is predicted to fall by 1%-point. Interactions between productivity risks, on the one hand, and demographic risks and financial risks, on the other, are found to play a modest role. However, productivity and equity return risks are very much related, particularly in the long run.

Comparing pension schemes, we find that the PAYG has the lowest implicit rate of return, mainly reflecting the prospected ageing of the population. Both the DB funded and the DC funded schemes that characterize the second- and third-pillar scheme in the Netherlands earn a higher average rate of return. Yet, the rates of return on the funded schemes are more risky. This holds especially in the short- and medium term. At longer horizons, the relative stability of the returns on the PAYG scheme decreases on account of productivity risk that does not exhibit mean reversion.

An interesting exercise is to attribute the riskiness of the returns on the different pension schemes to the four types of risks we distinguished. For the PAYG scheme, we find that the variability of the implicit rate of return stems from variability of productivity risk, including the risk of an extreme event. The effect of demographic risk on the return to a PAYG pension system turns out to be second-order only, however.

With regard to the DB funded scheme, the variability of the corresponding rates of return is determined almost fully by financial risks and risks of an extreme event. Again, the contribution of demographic risks is minor. The same holds true for the contribution of productivity risks.

As regards the gains from diversification, we find implicit correlation coefficients that are quite large: between 0.5 and 1. With nigh perfect correlation between the internal rates of return of different pension schemes, the risk reduction that can be achieved by mixing different pension schemes seems to be fairly small.

REFERENCES

Abel, A.B., 2001, Will bequests attenuate the predicted meltdown in stock prices when baby boomers retire?, *Review of Economics and Statistics* 83, 589–595.

Acemoglu, D. and Johnson, S., 2007, Disease and development: The Effect of life expectancy on economic growth, *Journal of Political Economy*, 115, 925-985

Alders, P.A. and Broer, D.P., 2005, Ageing, fertility, and growth, *Journal of Public Economics*, 89, 1075-1095.

Ang, A. and Maddaloni, A., 2005, Do demographic changes affect risk premiums? Evidence from international data, *Journal of Business* 78, 341–379.

Avouyi-Dovi, S. and Matheron, J., 2006, Productivity and stock prices, *Financial Stability Review* 8, 81–94.

Bakshi, G.S. and Chen, Z., 1994, Baby boom, population aging, and capital markets, *Journal of Business* 67, 165–202.

Barro, R.J., 1996, Determinants of economic growth: A cross-country empirical study, *NBER working paper 5698*, NBER.

Barro, R.J., 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics* 121(3), 823–866.

Barro, R.J. and Becker, G., 1989, Fertility choice in a model of economic growth, *Econometrica*, 76, 481-501.

Barro, R.J. and Ursúa, J.F., 2008, Macroeconomic crises since 1870, *Brookings Papers on Economic Activity* 1, 255–350.

Barro, R.J., Nakamura, E., Steinsson, J. and Ursúa, J., 2009, Crises and recoveries in an empirical model of consumption disasters, *mimeo*, Columbia University.

Beaudry, P. and Guay, A., 1996, What do interest rates reveal about the functioning of real business cycle models?, *Journal of Economic Dynamics and Control* 20(9-10), 1661–1682.

Beaudry, P. and Portier, F., 2005, The "news" view of economic fluctuations: Evidence from aggregate Japanese data and sectoral U.S. data, *Journal of the Japanese and International Economies* 19(4), 635–652.

Beaudry, P. and Portier, F., 2006, Stock prices, news, and economic fluctuations, *American Economic Review* 96(4), 1293–1307.

Benati, L., 2007, Drift and breaks in labor productivity, *Journal of Economic Dynamics and Control* 31(8), 2847–2877.

Benzoni, L., Collin-Dufresne, P. and Goldstein, R.S., 2007, Portfolio choice over the life-cycle when the stock and labor markets are cointegrated, *Journal of Finance* 62(5), 2123–2167.

Birchenall, J.A.,2007, Escaping high mortality, *Journal of Economic Growth*, 12(4), 351-387

Bodie, Z., 1976, Common stocks as a hedge against inflation, *Journal of Finance* 31, 459–470.

Boersch-Supan, A., 2003, Labor market effects of population aging, *Labour* 17, 5–44.

Boersch-Supan, A., Ludwig, A. and Winter, J., 2006, Aging, pension reform, and capital flows: A multi-country simulation model, *Economica* 73, 625–658.

Bloom, D., Canning, D. and Sevilla, J.P., 2004, The effect of health on economic growth: A production function approach, *World Development*, 32(1), 1-13

Bohn, H., 2009, Intergenerational risk sharing and fiscal policy, *mimeo*, UC Santa Barbara.

Broer, D.P., 2010, Macroeconomic risks and pension returns, *CPB memorandum241*, CPB.

Brooks, R., 2000, What will happen to financial markets when the baby boomers retire?, *IMF working paper 00/18*, IMF.

Brooks, R., 2002, Asset-market effects of the baby boom and social-security reform, *American Economic Review, Papers and Proceedings* 92, 402–406.

Brooks, R., 2003a, The equity premium and the baby boom, *Econometric Society North American Winter Meeting 155*, Econometric Society.

Brooks, R., 2003b, Population aging and global capital flows in a parallel universe, *IMF Staff Papers* 50, 200–221.

Burnside, C., Eichenbaum, M. and Rebelo, S., 1993, Labor hoarding and the business cycle, *Journal of Political Economy* 101(2), 245–273.

Campbell, J.Y., 2003, Consumption-based asset pricing, in G.M. Constantinides, M. Harris and R.M. Stulz (eds), *Handbook of the Economics of Finance*, vol. 1B, North-Holland, ch. 13, pp. 803–888.

Campbell, J.Y. and Cochrane, J.H., 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107(2), 205–251.

Campbell, J.Y. and Mankiw, N.G., 1987, Are output fluctuations transitory?, *Quarterly Journal of Economics* 102, 857–880.

Campbell, J.Y. and Shiller, R.J., 2001, Valuation ratios and the long-run stock market outlook: An update, *NBER working paper 8221*, NBER.

Campbell, J.Y. and Viceira, L.M., 1999, Consumption and portfolio decisions when expected returns are time varying, *Quarterly Journal of Economics* 114, 433–495.

Campbell, J.Y. and Viceira, L.M., 2002, *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, Oxford University Press, Oxford.

Campbell, J.Y. and Viceira, L.M., 2005, The term structure of the risk-return tradeoff, *CEPR discussion paper 4914*, CEPR.

Campbell, J.Y., Chan, Y.L. and Viceira, L.M., 2003, A multivariate model of strategic asset allocation, *Journal of Financial Economics* 67(1), 41–80.

Campbell, S.D. and Diebold, F.X., 2009, Stock returns and expected business conditions: Half a century of direct evidence, *Journal of Business and Economic Statistics* 27(2), 266–278.

Carlaw, K.I. and Lipsey, R.G., 2003, Productivity, technology and economic growth: What is the relationship?, *Journal of Economic Surveys* 17, 457–495.

Cerra, V. and Saxena, S.C., 2008, Growth dynamics: The myth of economic recovery, *American Economic Review* 98(1), 439–457.

Christiano, L.J. and Eichenbaum, M., 1990, Unit roots in real GNP: Do we know and do we care?, *Carnegie-Rochester Conference Series on Public Policy* 32, 7–62.

Clark, X., Hatton, T.J. and Williamson, J.G., 2002, Where do U.S. immigrants come from, and why?, *NBER working paper 8998*, NBER.

Cochrane, J.H., 1988, How big is the random walk in GNP?, *Journal of Political Economy* 96, 893–920.

Cochrane, J.H., 1994, Permanent and transitory components of GNP and stock prices, *Quarterly Journal of Economics* 109, 241–265.

Cochrane, J.H., 2007, The dog that did not bark: A defence of return predictability, *Review of Financial Studies* 21(4), 1533–1575.

Cochrane, J.H., 2008, Financial markets and the real economy, in R. Mehra (ed.), *Handbook of the Equity Risk Premium*, Elsevier, ch. 7, pp. 237–325.

Cogley, T., 1990, International evidence on the size of the random walk in output, *Journal of Political Economy* 98(3), 501–518.

Constantinides, G.M., Donaldson, J.B. and Mehra, R., 2002, Junior can't borrow: A new perspective on the equity premium puzzle, *Quarterly Journal of Economics* 117, 269–296.

Cook, S., 2008, More uncertainty: On the trending nature of real GDP in the US and UK, *Applied Economics Letters* 15(9), 667–670.

de Hek, P. and van Vuuren, D., 2008, Wage and productivity over the worker's life cycle, *CPB Document forthcoming*, CPB.

Diebold, F.X. and Senhadji, A.S., 1996, The uncertain unit root in real GNP: Comment, *American Economic Review* 86, 1291–1298.

Dimson, E., Marsh, P. and Staunton, M., 2008, The worldwide equity premium: A smaller puzzle, in R. Mehra (ed.), *Handbook of the Equity Risk Premium*, Elsevier, ch. 11, pp. 467–514.

Doepke, M., 2004, Accounting for fertility decline during the transition to growth, *Journal of Economic Growth*, 9(3), 347-383.

Domeij, D. and Flodén, M., 2006, Population aging and international capital flows, *International Economic Review* 47, 1013–1032.

Donaldson, J. and Mehra, R., 2008, Risk-based explanations of the equity premium, in R. Mehra (ed.), *Handbook of the Equity Risk Premium*, Elsevier, ch. 2, pp. 37–99.

Dostie, B., 2006, Wages, productivity and aging, *IZA discussion paper no. 2496*, IZA.

Fama, E.F., 1965, The behavior of stock-market prices, *Journal of Business* 38 (1), 34–105.

Fama, E.F. and French, K., 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25(1), 23–49.

Fama, E.F. and French, K.R., 1988, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246–273.

Fama, E.F. and French, K.R., 2002, The equity premium, *Journal of Finance* 57 (2), 637–659.

Fama, E.F. and Schwert, G.W., 1977, Asset returns and inflation, *Journal of Financial Economics* 5(2), 115–146.

Fehr, H., Jokisch, S. and Kotlikoff, L.J., 2004, Fertility, mortality, and the developed world's demographic transition, *Cesifo working paper 1326*, Cesifo.

Fehr, H., Jokisch, S. and Kotlikoff, L.J., 2005, Will China eat our lunch or take us out to dinner? Simulating the transition paths of the U.S., EU, Japan, and China, *NBER working paper 11668*, NBER.

Fenge, R. and Werding, M., 2003, Ageing and fiscal imbalances across generations: Concepts of measurement, *CESifo working paper 842*, Munich.

Fogel, R.W., 1997, New findings on secular trends in nutrition and mortality: Some implications for population theory, in M.R. Rosenzweig and O. Stark (eds.), *Handbook of Population and Family Economics*, ch. 9, pp. 433-481.

Galí, J., 1999, Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations?, *American Economic Review* 89, 249–271.

Galor, O., 2005, From stagnation to growth: Unified growth theory, in P. Aghion and S.N. Durlauf (eds.), *Handbook of Economic Growth vol. 1A*, Elsevier, ch. 4, pp. 171-293.

Galor, O. and Weil, D.N., 2000, Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond, *American Economic Review*, 90, 806-828.

Geweke, J., 2001, A note on some limitations of CRRA utility, *Economics Letters* 71 (3), 341–345.

Goyal, A., 2004, Demographics, stock market flows, and stock returns, *Journal of Financial and Quantitative Analysis* 39(1), 115–142.

Goyal, A. and Welch, I., 2003, Predicting the equity premium with dividend ratios, *Management Science* 49, 639–654.

Goyal, A. and Welch, I., 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

Hári, N., de Waegenaere, A., Melenberg, B. and Nijman, T.E., 2008, Estimating the term structure of mortality, *Insurance: Mathematics and Economics* 42, 492–504.

Harris, J.R. and Todaro, M.P., 1970, Migration, unemployment and development: A two-sector analysis, *American Economic Review* 60, 126–142.

Hellerstein, J.K., Neumark, D. and Troske, K.R., 1999, Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equations, *Journal of Labor Economics* 17, 409–446.

Hoevenaars, R.P.M.M., Molenaar, R.D.J., Schotman, P.C. and Steenkamp, T.B.M., 2008, Strategic asset allocation with liabilities: Beyond stocks and bonds, *Journal of Economic Dynamics and Control* 32(9), 2939–2970.

Jorion, P. and Goetzmann, W.N., 1999, Global stock markets in the twentieth century, *Journal of Finance* 54(3), 953–980.

Kalemli-Ozcan, S., 2002, Does the mortality decline promote economic growth?, *Journal of Economic Growth*, 7(4), 411–439

Karemera, D., Oguledo, V.I. and Davis, B., 2000, A gravity model analysis of international migration to North America, *Applied Economics* 32, 1745–1755.

Keilman, N., Cruijsen, H. and Alho, J.M., 2008, Changing views of future demographic trends, in J.M. Alho, S.E.H. Jensen and J. Lassila (eds), *Uncertain Demographics and Fiscal Sustainability*, Cambridge University Press, ch. 2, pp. 11–54.

King, R.G. and Rebelo, S.T., 1999, Resuscitating real business cycles, in J.B. Taylor and M. Woodford (eds), *Handbook of Macroeconomics, vol. 1B*, Elsevier, ch. 14, pp. 927–1007.

King, R.G. and Watson, M.W., 1996, Money, prices, interest rates and the business cycle, *Review of Economics and Statistics* 78, 35–53.

Knell, M., 2010, How automatic adjustment factors affect the internal rate of return of PAYG pension systems, *Journal of Pension Economics and Finance* 9 (1), 1–23.

Kotlikoff, L. and Wise, D., 1989, Employee retirement and a firm's pension plan, in D.A. Wisse (ed.), *The Economics of Aging*, University of Chicago Press, Chicago, pp. 279–334.

Kotlikoff, L.J. and Gokhale, J., 1992, Estimating a firm's age-productivity profile using the present value of workers' earnings, *Quarterly Journal of Economics* 107, 1215–1242.

Kotlikoff, L.J., Smetters, K. and Walliser, J., 2001, Finding a way out of America's demographic dilemma, *NBER working paper 8258*, NBER.

Kranendonk, H. and Verbruggen, J., 2007, SAFFIER, A multi-purpose model of the Dutch economy for short-term and medium-term analyses, CPB document 144.

Krueger, D. and Ludwig, A., 2007, On the consequences of demographic change for rates of return to capital, and the distribution of wealth and welfare, *Journal of Monetary Economics* 54, 49–87.

Kuipers, B.J., 2008, Solvency II ongeschikt voor peensioenfondsen, Achtergrond, APG Groep, 2, 9-15.

Lee, R.D. and Carter, L.R., 1992, Modeling and forecasting U.S. mortality, *Journal of the American Statistical Association* 87, 659–671.

Lee, R.D., and Miller, T., 2001, Evaluating the performance of the Lee-Carter method for forecasting mortality, *Demography*, 38, 537-549.

Lee, R.D. and Tuljapurkar, S., 1994, Stochastic population forecasts for the United States: Beyond high, medium, and low, *Journal of the American Statistical Association* 89, 1175–1189.

Lehr, C.S., 2009, Evidence on the demographic transition, *Review of Economics and Statistics*, 91, 871-887.

Lettau, M. and Ludvigson, S.C., 2001, Consumption, aggregate wealth and expected stock returns, *Journal of Finance* 56, 815–849.

Lettau, M. and Ludvigson, S.C., 2009, Measuring and modeling variation in the risk-return tradeoff, in Y. Ait-Sahalia and L. Hansen (eds), *Handbook of Financial Econometrics, vol. 1*, Elsevier, ch. 11, pp. 617-690.

Lettau, M. and van Nieuwerburgh, S., 2008, Reconciling the return predictability evidence, *Review of Financial Studies* 21, 1607–1652.

Lettau, M., Ludvigson, S.C. and Wachter, J., 2008, The declining equity premium: What role does macroeconomic risk play?, *Review of Financial Studies* 21 (4), 1653–1687.

Lima, L.R. and de Jesus Filho, J., 2008, Further investigation of the uncertain trend in US GDP, *Applied Economics* 40(9), 1207–1216.

Lindbeck, A. and Persson, M., 2003, The gains from pension reform, *Journal of Economic Literature* 41(1), 74–112.

Ludwig, A., Krueger, D. and Boersch-Supan, A.H., 2007, Demographic change, relative factor prices, international capital flows, and their differential effects on the welfare of generations, *NBER working paper 13185*, NBER.

Manuelli, R.E. and Seshadri, A., 2009, Explaining international fertility differences, *Quarterly Journal of Economics*, 124(2), 771-807.

Matsen, E. and Thøgersen, Ø., 2004, Designing social security – a portfolio choice approach, *European Economic Review* 48, 883–904.

Mayda, A.M., 2005, International migration: A panel data analysis of economic and non-economic determinants, *IZA discussion paper 1590*, IZA.

Mehra, R. and Prescott, E.C., 2008, Non-risk-based explanations of the equity premium, in R. Mehra (ed.), *Handbook of the Equity Risk Premium*, Elsevier, ch. 3, pp. 101–115.

Murray, C.J. and Nelson, C.R., 2000, The uncertain trend in U.S. GDP, *Journal of Monetary Economics* 46, 79–95.

Murray, C.J. and Nelson, C.R., 2002,The great depression and output persistence, *Journal of Money, Credit and Banking* 34, 1090–1098.

Nahuis, R., Canton, E. and de Groot, H., 2000, Vested interests, ageing, and growth, *Technical Report 3*, CPB.

Nelson, C.R. and Plosser, C.I., 1982, Trends and random walks in macroeconomic time series. Some evidence and implications, *Journal of Monetary Economics* 10(2), 139–162.

OECD, 1998, Work force ageing: Consequences and policy responses, *working paper AWP 4.1*, OECD.

Perron, P., 1989, The great crash, the oil price shock, and the unit root hypothesis, *Econometrica* 57, 1361–1401.

Poterba, J.M., 2001, Demographic structure and asset returns, *Review of Economics and Statistics* 83, 565–584.

Poterba, J.M., 2004, The impact of population aging on financial markets, *NBER working paper 10851*, NBER.

Prskawetz, A. and Lindh, T., 2006, The impact of population ageing on innovation and productivity growth in Europe, *Research Report 28*, Vienna Institute of Demography.

Rappoport, P. and Reichlin, L., 1989, Segmented trends and non-stationary time series, *Economic Journal* 99, 168–177.

Reinhart, C.M. and Rogoff, K., 2008, This time is different: A panoramic view of eight centuries of financial crises, *NBER working paper 13882*, NBER.

Reinhart, C.M. and Rogoff, K.S., 2009, The aftermath of financial crises, *NBER working paper 14656*, NBER.

Rudebusch, G.D., 1993, The uncertain unit root in real GNP, *American Economic Review* 83(1), 264–272.

Salomons, R., 2008, A theoretical and practical perspective on the equity risk premium, *Journal of Economic Surveys* 22(2), 299–329.

Sen, A., 2004, Are US macroeconomic series difference stationary or trend-break stationary?, *Applied Economics* 36, 2025–2029.

Sickles, R.C. and Taubman, P., 1997, Mortality and morbidity among adults and the elderly, Ch. 11 in M.R. Rosenzweig and O. Stark (eds.), *Handbook of Population and Family Economics*, Elsevier, pp. 559-643.

Sinn, H.-W., 2000, Why a funded pension system is useful and why it is not useful, *International Tax and Public Finance* 7, 389-410.

Soares, R.R., 2005, Mortality reductions, educational attainment, and fertility choice, *American Economic Review*, 95(3), 580-601.

Storesletten, K., Telmer, C.I. and Yaron, A., 1999, The risk-sharing implications of alternative social security arrangements, *Carnegie-Rochester Conference Series on Public Policy* 50(1), 213-259.

Temple, J., 1999, The new growth evidence, *Journal of Economic Literature* 37, 112–156.

Valkanov, R., 2003, Long-horizon regressions: Theoretical results and applications, *Journal of Financial Economics* 68(2), 201–232.

Vougas, D.V., 2007, Is the trend in post-WW II US real GDP uncertain or non-linear?, *Economics Letters* 94, 348–355.

de Waegenaere, A., Melenberg, B. and Stevens, R., 2010, Longevity risk, *panel paper 14*, Netspar.

Wagener, A., 2003, Pensions as a portfolio problem: Fixed contribution rates vs. fixed replacement rates reconsidered, *Journal of Populations Economics* 16, 111–134.

Weil, D., 2007, Accounting for the effect of health on economic growth, *Quarterly Journal of Economics*, 122(3), 1265-1306.

Weitzman, M.L., 2007, Subjective expectations and asset-return puzzles, *American Economic Review* 97, 1102–1130.

Welch, G. and Bishop, G., 1995, An introduction to the Kalman filter, *Technical report*, University of North Carolina at Chapel Hill, Chapel Hill, NC.

Zivot, E. and Andrews, D.W.K., 1992, Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis, *Journal of Business & Economic Statistics* 10, 25–44.

SUMMARY OF DISCUSSION

By Korhan Nazliben

Risk Factors in Pension Returns

By Peter Broer, Thijs Knaap and Ed Westerhout (CPB Netherlands Bureau for Economic Policy Analysis, The Hague)

Chairman: Henk Don
Discussants: Peter Schotman and Hens Steenhouwer

Thijs Knaap opened the presentation session by stressing that he and the other authors of this panel paper have focused on the literature on macroeconomic risk factors mainly in order to understand the most important risks facing pension schemes in the future, and which setup of the pension sector best protects against these macroeconomic risks. In so doing, they take into consideration four major risk factors: productivity, demography, financial returns and rare disasters. Attention is paid to their possible interactions.

Knaap presented results on the role of the different risk factors, the propagation of shocks through time and possible interactions between them. He also discussed the impact of these risk factors on the returns of four prototype pension schemes, which can be distinguished along the lines of PAYG and funding, and DB and DC. He further presented some illustrative calculations on the diversification gains of pension schemes. Both of the discussants, Peter Schotman and Hens Steehouwer, argued that model risk is very relevant for the policies of pension funds. The discussion concentrated on the performance measurement of the pension schemes.

Peter Schotman opened the discussion by highlighting the current problems of DB and DC pension schemes. First, he outlined the scope for future research in designing hybrid structures to trade in order to solve the current problems of the pension schemes. He stated that pension schemes face several additional risk factors that are directly related to financial risk, particularly in the long run. However, from a statistical point of view, there is little knowledge about the long-run properties of risk factors. This uncertainty in estimations, originating from the statistics rather than economical structure of the model, creates a significant additional risk, in general. Schotman pointed out the possible consequences of this model risk from various points of view. In particular, the term structure of risk was discussed elaborately, and for comparative purposes some examples from the literature were mentioned. He stressed that a minor change in the long-run properties of state variables can have a major effect on long-run risk.

Additionally, Schotman argued the importance of persistence parameters, and how hard it is to estimate such parameters precisely from a practical point of view. He illustrated this with some examples from the literature, and explained why long-run interest rate risk increases strongly with persistence, and how volatilities increase due to parameter uncertainty. Moreover, he also mentioned another difficulty from a macroeconomic point of view: that the time-varying stochastic discount factor may not give us good insight into performance evaluation and comparison of pension designs. Accordingly, Schotman, unanimously with Bas Werker and Theo Nijman, expressed his skepticism as to whether NPV and IRR should be considered as appropriate criteria or not.

Hens Steehouwer (Ortec Finance) initially emphasized that scenario features such as volatilities and correlations are very important, since they have a large impact on model outcomes and thereby on decisions and diversification gains. He also added that the term structure of volatilities and correlations are crucially important, and may create significant estimation errors in modelling. Steehouwer also discussed the robustness of the model, comparing it with some results from the literature. With regard to evaluations of pension scheme performance, he recommended consideration of other alternative performance evaluation measures, such as a risk/return, value-based approach together with IRR approach. He also discussed the risk decomposition approach of the paper, and alternatively suggested decomposing risk by simulating the separate risk sources and distinguishing a separate diversification effect. He added that in such a model it is worth investigating the simulation of multiple combinations with different allocations to the four pension schemes, calculating the IRR scenarios and appropriate risk- and return measures in these IRR scenarios, and assessing the efficiency effects in risk-return graphs.

The discussion focussed on the use of NPV and IRR in order to evaluate pension schemes and the type of rate to be used to discount pension liabilities. Ed Westerhout added that mean returns that are calculated for the prototype pension schemes say little about the efficiency of these pension schemes, because the transition process that one faces when a pension scheme is replaced with a different scheme has not been taken into account. He therefore argued that it is not a good suggestion to calculate risk-return tradeoffs. The discussion concluded with general agreement that more risk factors should be considered in order to increase the accuracy of the models. At the same time, however, more research must be done to ascertain the nature and magnitude of these effects.

PUBLICATIONS IN THE PANEL PAPERS SERIES

1. *Saving and investing over the life cycle and the role of collective pension funds*
   Lans bovenberg, Ralph Koijen, Theo Nijman and Coen Teulings
2. *What does behavioural economics mean for policy? Challenges to savings and health policies in the Netherlands*
   Peter Kooreman and Henriëtte Prast
3. *Housing wealth and household portfolios in an aging society*
   Jan Rouwendal
4. *Birth is the sessenger of death – but policy may help to postpone the bad news*
   Gerard van den Berg and Maarten Lindeboom
5. *Phased and partial retirement: preferences and limitations*
   Arthur van Soest and Tunga Kantarci
6. *Retirement Patterns in Europe and the U.S.* (2008)
   Arie Kapteyn and Tatiana Andreyeva
7. *Compression of morbidity: A promising approach to alleviate the societal consequences of population aging?* (2008)
   Johan Mackenbach, Wilma Nusselder, Suzanne Polinder and Anton Kunst
8. *Strategic asset allocation* (2008)
   Frank de Jong, Peter Schotman and Bas Werker
9. *Pension Systems, Aging and the Stability and Growth Pact* (2008) Revised version
   Roel Beetsma and Heikki Oksanen
10. *Life course changes in income: An exploration of age- and stage effects in a 15-year panel in the Netherlands* (2008)
    Matthijs Kalmijn and Rob Alessie
11. *Market-Consistent Valuation of Pension Liabilities* (2009)
    Antoon Pelsser and Peter Vlaar
12. *Socioeconomic Differences in Health over the Life Cycle: Evidence and Explanations* (2009)
    Eddy van Doorslaer, Hans van Kippersluis, Owen O'Donnell and Tom Van Ourti
13. *Computable Stochastic Equilibrium Models and their Use in Pension- and Ageing Research (2009)*
    Hans Fehr
14. *Longevity risk* (2009)
    Anja De Waegenaere, Bertrand Melenberg and Ralph Stevens

15. *Population ageing and the international capital market* (2009)
    Yvonne Adema, Bas van Groezen and Lex Meijdam
16. *Financial Literacy: Evidence and Implications for Consumer Education* (2009)
    Annamaria Lusardi and Maarten van Rooij
17. *Health, Disability and Work: Patterns for the Working-age Population* (2009)
    Pilar García-Gómez, Hans-Martin von Gaudecker and Maarten Lindeboom
18. *Retirement Expectations, Preferences, and Decisions* (2010)
    Luc Bissonnette and Arthur van Soest
19. *Interactive Online Decision Aids for Complex Consumer Decisions: Opportunities and Challenges for Pension Decision Support* (2010)
    Benedict Dellaert
20. *Preferences for Redistribution and Pensions. What Can We Learn from Experiments?* (2010)
    Jan Potters, Arno Riedl and Franziska Tausch
21. *Risk Factors in Pension Returns* (2010)
    Peter Broer, Thijs Knaap and Ed Westerhout

Netspar

Risk Factors in Pension Returns

Pension schemes are vulnerable to a number of macroeconomic risks. This paper by Peter Broer, Thijs Knaap and Ed Westerhout (all CPB) assesses these risks. It reviews the literature on risks in demography, productivity and financial markets, paying attention to their development over time and possible interactions. Using this review, they construct a VAR model for the Netherlands to derive the implications of risks for the returns on defined-benefit (DB) and defined-contribution (DC) pension schemes that are either funded or PAYG-financed.