Netspar

Network for Studies on Pensions, Aging and Retirement

Netspar DISCUSSION PAPERS

Lans Bovenberg and Roel Mehlkopf

Optimal Design and Regulation of Funded Pension Schemes

DP 08/2013-026

Optimal Design and Regulation of Funded Pension Schemes

Lans Bovenberg* Roel Mehlkopf†

August 20, 2013

Abstract: This paper reviews the literature on the optimal design and regulation of funded pension schemes. We first characterize optimal saving and investment over an individual's life cycle. Within a stylized modeling framework, we explore optimal individual saving and investing behavior. Subsequently, various extensions of the model are considered, such as additional financial risk factors, stochastic human capital and more elaborate individual preferences. We then turn to the literature on intergenerational risk sharing, which suggests that a long-lived entity such as a pension fund or the government can yield ex-ante welfare gains by allowing non-overlapping generations to trade risk. The scope for this type of intergenerational risk sharing, however, is limited by the ability to commit generations to the contract. These commitment problems raise concerns with respect to sustainability and intergenerational fairness. We explore the role of solvency regulations to address these concerns about intergenerational fairness and discontinuity risk.

Key words: saving, investment, life cycle, pension schemes, risk sharing, commitment problem, discontinuity risk

JEL Code(s): D91, G11, G23

*Tilburg University and Netspar, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: A.L.Bovenberg@tilburguniversity.edu

†Tilburg University and Netspar, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: R.J.Mehlkopf@tilburguniversity.edu

1 Introduction

This paper reviews the role of pension fund regulation in addressing concerns with respect to intergenerational fairness and sustainability. Pension funds that facilitate intergenerational risk sharing can be found in many countries. Examples include the Social Security Trust Funds in the United States, the Japan Government Pension Investment Fund, the Canada Pension Plan, the Central Provident Fund in Singapore, the Employee Provident Fund in Malaysia, and occupational pension funds in Iceland and the Netherlands. Sovereign wealth funds (for example in Norway) face similar issues of intergenerational risk sharing in deciding on their dividend policies. Most of these funds are diversified with respect to asset class as well as internationally. Some funds, such as the US Social Security Trust Fund, are intended to act as a buffer against demographic shocks and are expected to be depleted in the coming decades. Others, such as the Canada Pension Plan, are permanent in nature and are expected to grow in size in the coming decades. Risk sharing between non-overlapping generations is not possible in countries in which individuals save and invest for retirement in individual retirement accounts. Several countries (for example, Australia and various Latin American and Eastern European countries) have established a mandatory funded social security tier with individual retirement accounts.

The paper starts out from the recent academic literature on optimal financial planning of individuals over the life cycle. Individuals face two main decisions in their financial planning over the life cycle. Through the saving decision, they decide how to smooth consumption over time by setting the pension premium and the pension benefits. Through the investment decision, individuals decide how to invest the premia in the various financial assets so as to smooth consumption across various future contingencies. Under a set of "golden assumptions," the young should invest a larger fraction of their financial wealth in risky assets than the elderly do. Indeed, life-cycle funds and target-date funds, which are based on these principles, are popular investment vehicles in many countries. We investigate how robust their recommendations are with respect to various extensions of our benchmark model. In particular, we explore alternative specifications of preferences, financial risks and labor income.

The second part of this paper explores how intergenerational risk sharing between non-overlapping generations can improve the situation in comparison to a laissez-faire economy. The literature on intergenerational risk sharing suggests that a long-lived entity such as a pension fund or the government can yield ex-ante welfare gains by allowing non-overlapping generations to trade risk. In particular, in the absence of commitment problems, it is optimal to diversify risk over as many generations and as many time periods as possible. In this way, shocks are smoothed out as broadly as possible across both individuals and periods so that each separate individual in each period is affected as little as possible. We show that intergenerational risk sharing can be conducted with alternative criteria for intergenerational fairness and associated solvency requirements for the buffers that must be reserved for future generations.

We investigate also how the ability to commit all generations to the risk-sharing contract affects the scope for intergenerational risk sharing. Whereas intergenerational risk sharing can be welfare improving for all generations from an ex-ante perspective, some generations are worse off from an ex-interim perspective if these generations collect a negative buffer from previous generations as a result of an adverse shock in the past. If young workers face substantial negative buffers when they enter the pension scheme, they may "*vote with their voice*" by exerting political pressure to change the rules that force them to pay for the risks that are shifted onto them by previous generations. These political-economy considerations render the risk-sharing solution vulnerable to discontinuity risk. We show how solvency rules can limit discontinuity risks by ensuring that negative wealth transfers from older to younger generations do not become too large.

Young generations can also "*vote with their feet*" by seeking employment elsewhere when workers in a particular sector or country are forced to pay for risks that materialized in the past. These behavioral responses introduce a trade-off between optimal intertemporal consumption smoothing and minimizing labor-market distortions. Labor supply responses can be substantial if risk sharing takes place in employer-based or industry-wide pension funds because workers can move to other employers or sectors rather easily. Associated concerns about discontinuity risk may result in tight solvency requirements that limit intergenerational risk sharing. Nonetheless, employer-based or sectoral pension funds may have some scope for committing workers to the risk-sharing contract if human capital is specific to the firm or the industry.

Our paper builds on three strands of the academic literature: namely, the literature on optimal consumption and portfolio choice over an individual's life cycle, on intergenerational risk sharing, and on asset pricing.

The rest of this paper is structured as follows. Technical derivations are contained in the appendix. Chapter 2 focuses on optimal individual saving and investment behavior in the absence of risk sharing between non-overlapping generations. Section 2.1 sets up our benchmark framework for analyzing life-cycle financial planning. Within this stylized framework, Section 2.2 explores the optimal saving and investment behavior from the perspective of the individual. Sections 2.3, 2.4 and 2.5 investigate the consequences of a more elaborate modelling of preferences, risk factors and human capital for optimal life-cycle behavior.

Chapter 3 explores the possibilities and limitations of risk sharing between non-overlapping generations. Section 3.1 sets up a stylized model for analyzing this intergenerational risk sharing. Section 3.2 derives the optimal risk-sharing solution in the absence of commitment problems, and shows how intergenerational risk sharing yields an ex-ante welfare gain as a result of improved diversification of risk across generations. Section 3.3 characterizes the set of ex-ante Pareto-efficient risk-sharing solutions and explores two fairness criteria that determine unique solvency requirements within this set. Section 3.4 explores how commitment problems impose constraints on the scope for intergenerational risk sharing and how these problems provide a rationale for tight solvency requirements on the buffers that must be established.

2 Saving and investing over the life cycle

2.1 Benchmark model

This section lays out our benchmark framework for exploring optimal lifetime saving and investment (see Merton (1971) and Merton and Samuelson (1974)). It describes our benchmark assumptions on financial markets, labor markets and preferences and discusses the parameter values employed in our numerical simulations.

Preferences Individuals aim to maximize lifetime utility, which is the weighted sum over time of expected utility at each point in time. Bequest motives are absent. In particular, the investor solves

$$\max E_t \left(\int_t^D \frac{e^{-\rho(s-t)}}{1-\theta} C_s^{1-\theta} ds \right), \tag{1}$$

where $\theta > 0$ represents the coefficient of relative risk aversion and $\rho > 0$ stands for rate of time preference. Utility at a point in time thus depends only on consumption at that time. The weights of future expected utilities decline exponentially at the rate of time preference. Preferences feature positive and constant relative risk aversion (CRRA). A positive parameter θ implies that individuals have a taste for moderation across time and across contingencies. More general models (than expected utility) distinguish the taste for moderation across *contingencies* from the taste for moderation across *time*. Risk aversion measures the taste for moderation across random outcomes while the preference for moderation across time is inversely related to the intertemporal elasticity of substitution. An individual exhibiting a low intertemporal elasticity of substitution prefers a stable level of consumption over time.

Financial markets The investor can invest in two assets, namely risk-free bonds B_t and stocks P_t. These assets develop over time as follows

$$\frac{dB_t}{B_t} = rdt, \tag{2}$$

$$\frac{dP_t}{P_t} = \mu dt + \sigma dZ_t, \tag{3}$$

where r denotes the risk-free (real) interest rate, $\mu > r$ represents the expected return on stocks and σ is the standard deviation of the log stock returns. dZ_t stands for a Brownion motion or Wiener process with zero drift and unit volatility. Log stock returns are thus identically and independently distributed according to a normal distribution. Furthermore, the interest rate, the volatility of equity, and the equity risk premium are constant over time so that mean reversion and stochastic volatility are absent. These prices are not affected by the decisions of individuals or financial intermediaries. We thus take a partial-equilibrium perspective of a small open economy, which takes prices as given on the world market. Financial markets are dynamically complete because agents can continuously trade stocks and bonds without constraints and equity-market risk is the only aggregate risk factor.

Labor earnings The investor starts working at time $t = 0$, works up to time T and passes away at time D. The investor collects constant labor income during the period $t \in [0, T]$ when it is normalized to $Y_t = 1$, while

$Y_t = 0$ for $t \in (T, D]$. The wage during the working career is constant and riskless with a fixed, exogenous retirement age. Moreover, labor-market risks are absent and labor supply is fixed. Death is predictable or perfect insurance of individual longevity risk is available. Aggregate longevity risk is thus absent.

The investor's financial wealth at time t is denoted by F_t, consumption at time t is given by C_t, labor income at time t is given by Y_t, and the share of financial wealth invested in risk-bearing assets is denoted by f_t^*. Financial wealth then develops as follows

$$dF_t = F_t(r + f_t^* \sigma \lambda)dt + F_t f_t^* \sigma dZ_t + (Y_t - C_t)dt, \tag{4}$$

where $\bar{\mu} \equiv \mu - r$ denotes the expected excess return and $\lambda \equiv \frac{\bar{\mu}}{\sigma}$ stands for the expected excess return per unit of risk, which is known as the Sharpe ratio.

Benchmark parameters in numerical calculations We illustrate our results with numerical simulations. Following Teulings and de Vries (2006), these simulations are based on a constant coefficient of risk aversion of 5, a working life of 45 years, an expected retirement period of 15 years, a rate of time preference and a risk-free interest rate of 2% per year, an equity risk premium of 4% a year, and an annual standard deviation of stock returns of 20%.

2.2 The first-best solution

This section derives the first-best solution to the financial planning problem of the individual under the assumptions outlined above. We analyze this case in depth because it generates intuition and a benchmark for the findings in more elaborate models. Subsection 2.2.1 explores consumption smoothing (i.e. saving decisions) in the absence of risk. Subsection 2.2.2 investigates optimal asset allocation in the presence of risky investment opportunities. Subsection 2.2.3 returns to optimal consumption smoothing but considers the case with risky assets.

2.2.1 Intertemporal consumption smoothing without risk

Together with the depreciation of human capital due to aging, the preference for a smooth consumption stream over time gives rise to a demand for pension saving. Capital markets allow individuals to engage in intertemporal trade. In particular, through capital markets, one can exchange resources in the active periods of life, when labor resources are relatively abundant but consumption is not so valuable at the margin, to the inactive periods of life when consumption is relatively more valuable at the margin but labor resources have already been depreciated.

Perfect consumption smoothing without impatience

The simplest case to consider is when the real interest rate is zero, individuals are not impatient (i.e. the rate of time preference is zero), and wage income is constant during the active life. In that case, complete consumption smoothing is optimal: consumption should remain constant during the life cycle. The life-time budget constraint limits the level of consumption. The individual begins the working life without any financial wealth so that he relies on human wealth only. As the discounted value of life-time wage income, human wealth at the beginning of the working life is simply the number of active years times the annual wage income. A constant consumption flow equal to the share of the active life in the remaining life time times the labor income flow during the active life exactly exhausts human capital at the end of life. The savings rate is thus equal to the share of retired life in the overall adult life.

Perfect consumption smoothing with impatience if the real rate equals the time preference

Individuals are typically impatient. This implies that individuals should be rewarded for saving through positive interest rates. If the real interest rate equals the rate of time preference, the individual still finds a constant consumption stream optimal because the reward of waiting (i.e. the interest rate) exactly balances the cost of waiting (i.e. the rate of time preference). Compared to the case with a zero interest rate and a zero rate of time preferences, the consumption level can be higher *ceteris paribus* because the pension saver benefits from positive net interest income on his accumulated savings.

Consumption smoothing in the general case, excluding risky investments

In the general case, the sign of the gap between the interest rate and the rate of time preference determines whether consumption is increasing or decreasing over time. With a positive gap, the benefits from waiting offered in the capital market exceed the subjective cost of waiting. Hence, the net reward for waiting is positive. This positive reward makes a rising path for consumption c^i of individual i optimal (see equation (12) in the appendix)

$$\frac{dC_t/dt}{C_t} = \frac{(r-\rho)}{\phi}. \tag{5}$$

The growth rate of consumption $\frac{dc/dt}{c}$ is not very sensitive to the gap between the interest rate r and the rate of time preference ρ if a small intertemporal substitution elasticity $(1/\phi)$ indicates that the individual exhibits a large preference for consumption smoothing over time. Intuitively, behavior is not very sensitive to intertemporal prices (i.e. the net reward for waiting) if agents dislike large differences between consumption levels at different points in time.

2.2.2 Risk taking

Financial markets allow agents to shift consumption not only across time, but also across various future contingencies when agents face uncertainty about which contingency will actually materialize. In particular, risk-averse individuals can buy resources in bad states by giving up resources in good states. Just as in the case of optimal intertemporal consumption smoothing, we can distinguish between the preferences for consumption smoothing (across *contingencies* rather than *time* so that risk aversion rather than the intertemporal elasticity of substitution measures this preference) and the prices for consumption smoothing (in this case, instead of the net interest rate $(r-\rho)$, the Sharpe ratio (i.e. the expected excess return (over the risk-free return) per standard deviation of the excess return) is relevant).

How much risk the individual optimally chooses to absorb depends on both risk aversion and the risk premium (for each unit of risk as measured by the Sharpe ratio). The investor chooses the optimal amount of risk in such a way that his coefficient of relative risk aversion (i.e. the negative elasticity of marginal utility with respect to consumption), θ, equals the marginal rate of substitution between the expected excess return of risk taking $f\overline{\mu}$ and variance $f^2\sigma^2$ of the investment portfolio

$$\theta = \frac{f\overline{\mu}}{f^2\sigma^2}, \tag{6}$$

where f represents the optimal share of total wealth invested in equity. This implies that

$$f\sigma = \frac{\lambda}{\theta}, \tag{7}$$

so that the optimal expected excess return due to risk taking amounts to $f\overline{\mu} = \lambda^2/\theta$. Risk taking (i.e. volatility of consumption across various contingencies) as measured by $f\sigma$ increases with the reward to risk taking λ and decreases with the preference for consumption smoothing as measured by relative risk aversion θ. With our benchmark parameters, the optimal investment share f is 20%.

The expression for optimal consumption smoothing across contingencies (7) is similar to that for optimal consumption smoothing across time (5). In both cases, the right-hand side involves the net price for smooth consumption (i.e. the Sharpe ratio λ in the case of (7) and the net interest interest rate $(r - \rho)$ in the case of (5)) and the preference for consumption smoothing (i.e. relative risk aversion θ in the case of (7) and the reciprocal inverse intertemporal substitution elasticity $(1/\phi)$ in the case of (5)). Prices and preferences together determine the optimal inequality in consumption (across *contingencies* as measured by $f\sigma$ or across *time* as measured by the growth rate of consumption) at the left-hand side of these expressions.

Optimal investment share age invariant

The share f of total wealth invested in the risk factor does not depend on age, or more generally, the investment horizon. Suppose that we invest for t years rather than one year. In that case, both the expected excess return $f\overline{\mu}t$ and the variance of the excess return $f^2\sigma^2 t$ in (6) vary proportionally with the length of the investment period so that the optimal investment share is not affected by the time horizon. In other words, both the marginal benefits and the marginal costs of investing more in equity rise linearly with time. This reasoning shows that the familiar argument that time diversification allows young people to take more risk is fallacious and relies on a wrong interpretation of the law of large numbers. The sum of n independent and identically distributed random variables exhibits a variance that is n times larger than the variance of each of the separate risks. The law of large numbers, in contrast, states that the variance of the *average* (rather than the *sum*) of n independent and identically distributed random variables goes to zero if n becomes very large.

Mossin (1968), Merton (1969) and Samuelson (1969) first independently derived the result that the investment share in the risky asset is independent of age. As we noted above on consumption smoothing over time (see (10)), the elasticity of consumption with respect to wealth is unity and thus independent of age. Hence, a young person is equally vulnerable (in terms of the *relative* change in the consumption flow, which is relevant in case *relative* risk aversion is constant) to the same *relative* change in wealth and should thus *ceteris paribus* hold the same wealth *share* (as opposed to the absolute amount of wealth) in risk-bearing assets.

Optimal investment share of tradable financial wealth

Human wealth is non tradable. Hence, financial rather than human wealth should be adjusted to achieve the right exposure to risk factors. If human capital is riskless, it acts like a risk-free asset and all the exposure to risk should come from financial wealth. As the wealth share of financial wealth increases from zero to one during the working life, the share of *financial* wealth invested in risk-bearing assets, f^*, falls from infinity at the beginning of the working life to f at retirement:

$$f_t^* = \frac{S_t}{F_t} = f\frac{W_t}{F_t} = f(1 + \frac{H_t}{F_t}) = \frac{1}{\theta}\frac{\lambda}{\sigma}(1 + \frac{H_t}{F_t}), \tag{8}$$

where S_t, F_t and H_t denote, respectively, wealth invested in equity, financial wealth, and human wealth at time t so that $W_t = H_t + F_t$. This financial wealth share tends to fall for two reasons. First, the absolute amount of wealth invested in equity tends to fall with time as an individual consumes part of human wealth during the working life (so that W_t tends to decline with time). Second, the stock of financial wealth F_t tends to increase as the individual saves part of their human capital.

2.2.3 Intertemporal consumption smoothing with risk

Subsection 2.2.1 explored optimal saving without risk. This subsection revisits optimal saving behavior in the presence of risky investment opportunities. The introduction of risk affects optimal saving through two channels. First, risk taking enhances welfare because investors can now capture the equity premium. This positive income effect raises optimal consumption and reduces saving out of labor income. Second, risk taking introduces a precautionary saving motive. The sign of this motive depends on whether marginal utility is convex (i.e. on the sign of the third derivative of the utility function, see Leland (1968)). In particular, risk implies an additional precautionary saving motive if marginal utility is convex so that the expected marginal utility of consumption (which determines the savings motive if the future is uncertain) exceeds the marginal utility of expected consumption (which determines the savings motive if the future is certain).

With constant relative risk aversion, marginal utility is indeed convex. As a direct consequence, the introduction of risk unambiguously increases the growth rate of expected consumption. In particular, with a variance $\psi^2 = f^2\sigma^2$ of aggregate wealth, we derive for a constant-relative-risk-aversion (CRRA) utility function (with $\phi = \theta$) and the associated optimal investment share $f = \frac{\lambda}{\sigma\theta}$ that the expected growth rate of consumption is (see equation (12)

$$E\left[\frac{dC_t/dt}{C_t}\right] = \frac{(r - \rho + \frac{(1+\theta)}{2}\theta\psi^2)}{\theta} = \frac{(r - \rho + \frac{(1+\theta)}{2}\frac{\lambda^2}{\theta})}{\theta}. \tag{9}$$

Shocks in wealth and consumption smoothing

The rate at which overall wealth is consumed depends on age. Since older people feature a shorter planning horizon, they consume a larger share of their overall wealth. If both a young and an old person get one additional euro, the old person will consume the euro more rapidly. However, if both agents obtain x % more wealth, both agents will increase their consumption by x % during the rest of their lives.Rather than spending it in a few periods, they choose to enjoy the wealth boost in equal relative consumption increases in each period during the rest of their lives (see equation (15) in the Appendix):

$$\frac{d\log C_s}{d\log W_t} = 1, \tag{10}$$

where $s \geq t$ and where W_t stands for total wealth (i.e. the sum of human and financial wealth) at time t and C_s denotes consumption at time s. The unitary elasticity of the consumption flow (in the rest of the life time) with respect to wealth implies that both pension contributions should decline and pension benefits increase following a positive wealth shock. Hence, rather than a pension system that keeps the premium fixed (a *defined-contribution* system) or a pension system that fixes the benefits (a *defined-benefit* system), a hybrid system that adjusts both premia and benefits in response to income and wealth shocks appears to be optimal.

2.3 Preferences

This subsection explores the implications of a coefficient of relative risk aversion that is not constant. Subsection 2.3.1 first of all considers the implications of minimum consumption needs. Subsection 2.3.2 then considers external and internal habits, respectively. Finally, subsection 2.3.3 investigates loss aversion and endogenous reference points.

2.3.1 Required minimum consumption

Preferences may imply that people require a minimum amount of consumption in order to derive any utility at all:

$$U = E\left[\int_t^D \exp(-\rho(s-t))\frac{1}{1-\theta}\left(C_s - \bar{C}_s\right)^{1-\theta} ds\right], \tag{11}$$

where $\bar{C}_s \geq 0$ represents the *minimum* (or *necessary* or *required*) consumption level at time s.

As regards investment behavior, the optimal wealth share invested in equity is then given by

$$f_t^* = f\frac{(W_t - X_t)}{W_t} = \frac{\lambda}{\theta\sigma}(1 - \frac{X_t}{W_t}), \tag{12}$$

where $X_t = \int_t^D \exp(-rs)\bar{C}_s ds$ stands for the discounted value (using the risk-free rate as a discount rate) of the minimum consumption levels over the rest of the life cycle. Intuitively, the investor first invests in risk-free assets X_t to guarantee the minimum consumption level and then invests a fixed fraction f (given by (7)) of remaining wealth $(W_t - X_t)$ in risk-bearing assets. In the terminology of asset-liability management, one thus matches the fixed liability X_t with risk-free assets.

2.3.2 External and internal habits

Campbell and Cochrane (1999) use the preferences (11) in which the consumption requirements $\bar{C}_t$ vary with the history of aggregate consumption. They employ these preferences to explain a number of asset pricing phenomena, including a large equity premium that varies countercyclically. Investors demand a high return on stocks because stocks perform poorly in recessions when consumption levels C_t fall close to the minimum levels $\bar{C}_t$.

External habit formation contrasts with internal habit formation where consumption requirements $\bar{C}_t$ depend on the history of *individual* consumption rather than *aggregate* consumption of the economy as a whole. Constantinides (1990), for example, assumes that habits of an individual $\bar{C}_t$ depend on actual consumption levels of that same individual as follows:

$$d\bar{C}_t/dt = \kappa C_t - \xi \bar{C}_t.$$

Internal habit formation implies more complicated behavior if agents anticipate how their current consumption levels C_t affect future consumption needs $\bar{C}_{t+s}$. For example, consumption at two neighbouring points in time become complements: a high consumption level at time t raises, by its effect on habits $\bar{C}_{t+s}$, the marginal utility of consumption in the immediate future $t+s$. This explains why, following an unexpected shock in wealth, consumption does not adjust immediately to its long-run level but adjusts gradually over time. Innovations in consumption will thus be correlated and the optimal consumption no longer follows a random walk.

Habit formation (i.e. $\kappa, \xi > 0$) implies also that the ability to absorb an unexpected shock increases if the length of the planning horizon increases. This implies that younger agents with longer horizons can adjust more easily to shocks and thus should invest a larger share of their wealth in risk-bearing assets. Intuitively, they have more time to adjust their habits to unexpected shocks.

2.3.3 Loss aversion

Standard utility functions, which are twice differentiable, imply that the costs of risk (see the denominator of the fraction at the right-hand side of (6)) are proportional to the variance of a portfolio, $f^2\sigma^2$.[1] Recent experimental work, however, suggests that human behavior is not always well described by twice differentiable concave utility functions. Tversky and Kahneman (1992) therefore developed prospect theory to better describe human behavior with respect to risk. In particular, behavior seems better captured by utility function with a kink at the so-called reference point: a gain from this reference point yields only a small increase in utility, while a loss yields a large decrease.[2] This kink implies that risk aversion is particularly large at consumption levels close to the reference point. In other words, relative risk aversion is not constant but depends on the actual consumption level and

[1] Mathematically, this result relies on a second-order Taylor expansion. The cost of risk taking rises exactly quadratically with the magnitude of the deviations if the coefficient of relative risk aversion is constant and the Taylor approximation is exact.

[2] On the basis of experimental work, Tversky and Kahneman (1992) find that losses tend to weigh about 2 1/4 times more heavily than gains.

the particular losses and gains considered. A second[3] element of prospect theory is that agents are risk averse in gains (twice as large gains yield less than twice as much increase in utility) and risk loving in losses (twice as big losses yield less than twice as much decrease in utility). Accordingly, agents exhibit a large preference for consumption levels close to the target level of the reference point. If their current wealth is insufficient to finance the consumption level at the reference point, they will take risks to get close to it. If their wealth level exceeds the target level, in contrast, they will shy away from risks in order to prevent the risk of falling below the target level.

Prospect theory, and the kink at the reference point in particular, has important implications for risk taking over time. We can best explain this by the piece-wise linear utility function with a kink at the average consumption level. For this utility function, the cost of risk taking increases proportionally in the deviation rather than quadratically as with a smooth concave utility function. This implies that the cost of risk taking (i.e. the risk premium) increases with the square root of the time horizon because the expected standard deviations rise with the time horizon in this way if the draws are independent of time. This is in contrast to the case of a smooth concave utility function when the cost of risk taking increases quadratically with the deviations and thus linearly with the time horizon. With kinked utility, the marginal benefit of risk taking (i.e. the expected excess return $\overline{\mu}T$) increases linearly in the investment horizon while the associated marginal cost increases only with the square root of the time horizon $\sigma\sqrt{T}$. As a direct consequence, optimal risk taking increases with the time horizon. This yields another reason why young agents should bear more risk than old agents: loss aversion as described by prospect theory implies that young agents with long horizons should invest more in equity than old agents.

The reference point in prospect literature plays a similar role as the habit level in the habit-formation literature. Hence, the literature has explored what happens if the reference point adjusts over time in response to current consumption levels and rational investors anticipate this. Berkelaar, Kouwenberg and Post (2004) find that updating the reference point implies more risk taking. Intuitively, preferences are more flexible: with adverse shocks the point from which gains and losses are being evaluated and to which you want to be close is adjusted downwards.

[3]A third element is that decision-makers employ subjective decision weights that differ from the true probabilities.

Barberis, Huang and Santos (2001) explore a setting in which investors derive direct utility from not only consumption but also from fluctuations in their financial wealth. After a run-up in stock prices, the agent gets further away from this reference point and thus becomes less risk averse. History thus determines risk aversion. Just as in Campbell and Cochrane (1999), this variation in risk aversion allows the equilibrium returns to be much more volatile than the underlying fundamentals.

2.4 Financial risk factors and longevity risk

Apart from human capital risk, the main risk factors that individuals face are interest rate risk, inflation risk, equity risk and longevity risk. The benchmark model laid out in Section 2.1 included only equity risk and imposed that stock returns were not predictable. Subsection 2.4.1 introduces interest-rate and inflation risk. Subsection 2.4.2 discusses the evidence on stock return predictability and its implications for optimal consumption smoothing and risk sharing over the life cycle. Subsection 2.4.3 considers micro as well as macro longevity risk and the value of annuity markets and longevity bonds.

2.4.1 Interest-rate and inflation risk

So far we have ignored inflation risk and assumed a constant interest rate. If (real) interest rates fluctuate over time, the investor is faced with an additional risk factor. Fluctuations in interest rates affect the value of the human capital as well as the value of the bond portfolio. Early in the life cycle, portfolios are to be held with long durations to hedge consumption during retirement against interest-rate risk. In this way, the duration of the assets is more closely aligned with the duration of the consumption pattern. Indeed, in the absence of inflation risk, long-term nominal bonds are riskless for long-term investors if the maturity equals the horizon of the investor. If inflation risk is taken into account, however, Campbell and Viceira (2005) report that the annualized standard deviation of real returns on these bonds is as large as 8% and exceeds that of stocks for maturities of over 35 years. Risk-averse long-term investors therefore exhibit a strong preference for index-linked bonds. Campbell and Viceira (2001) report that the certainty-equivalent wealth effect of access to inflation-indexed bonds can be as large as 10-30%.

In reality, inflation-linked assets are often not available.[4] In that case, investors can exploit the differences between bonds with different maturities with respect to the sensitivities to the real rate and expected inflation in order to hedge away substantial parts of the inflation risk using sophisticated dynamic trading strategies. Unexpected inflation, however, can not be hedged without access to inflation-linked assets. Morever, many investors will be constrained to hedge against the real interest and expected inflation factors or will not be equipped to implement dynamic trading strategies. In any case, these strategies suffer from model risk.

If investors cannot pursue these strategies and are forced to bear inflation risk, they will shorten the maturity of the bond portfolio compared to the case in which inflation risk is no concern (see Campbell and Viceira (2001, 2005)).This implies that the optimal duration of the fixed income portfolio is smaller than that of the consumption profile in retirement. Solvency requirements that compute the value of nominal liabilities on the basis of nominal interest rates can force pension funds to extend the maturity of their assets to that of their nominal liabilities even though this may not be optimal from the point of view of risk-averse individuals who care about the purchasing power of their pensions.

2.4.2 Equity risk reconsidered

There is a lively debate in the literature as to whether the expected stock returns are time-varying.[5] A typical finding in the literature is that, at the low frequencies that are relevant for financial planning over the life cycle, stock returns are negatively correlated, implying that the stock market is mean reverting. As discussed in subsection 2.2.2, both the mean and the variance of the excess return vary in the same way with the investment horizon (namely proportionally) if expected returns are constant over time. If stock returns are mean reverting, in contrast, the variance of the excess returns rises less rapidly with the time horizon than the mean of the excess return. As a direct consequence, stocks are less risky for long-term investors

[4]In the US and the UK inflation-linked bonds are traded and a number of other countries introduced them recently. The markets, however, are typically neither deep nor liquid. Moreover, the inflation index that is traded (e.g. national price inflation) may differ from the index that is relevant to the investor (e.g. regional wage inflation).

[5]See e.g. Ang and Bekaert (2007), Campbell and Yogo (2006), Cochrane (2008), Goyal and Welch (2008), and Lettau and van Nieuwerburgh (2008).

than for short-term investors so that young investors should invest an even larger over their financial wealth in equities. Campbell and Viceira (2005) estimate that the annualized standard deviation of real returns of about 17% of a diversified stock portfolio is reduced to an annualized standard deviation of only 8% on a 20-year horizon.

2.4.3 Longevity risk

Two types of longevity risk can be distinguished. Macro longevity risk refers to uncertain development in future mortality rates due to improved health care and changes in nutrition and life style. Macro longevity thus affects the average mortality of a pool of individuals. Micro longevity risk, in contrast, refers to the uncertainty for an individual on whether or not he will survive, irrespective of changes in mortality rates.

In the absence of bequest motives and liquidity constraints, annuities that are priced in an actuarially fair fashion are welfare improving (Davidoff, Brown and Diamond (2005) and Poterba (2006)). If real annuities or variable annuities that offer exposure to stock markets are not available, however, annuities become less attractive. Also adverse selection is a concern. Older individuals typically have superior information on their survival probabilities compared to that of the issuers of the annuities. As the asymmetry in information becomes more important for older agents with poor health conditions, annuities are typically irreversible. This gives rise to a fundamental trade-off between liquidity and longevity insurance.

For pension funds and other issuers of annuity contracts macro longevity risk is a relevant concern. Unexpected increases in life expectancy raise the market value of their liabilities and harm their solvency position. From a macro-economic point of view, the natural issuers of longevity protection are the young and the future generations. By issuing longevity bonds, the government can offer instruments to insure the longevity risk of older people by shifting it to younger cohorts – for example, by conditioning the retirement age of the public pension system on longevity. By this conditioning of the retirement age, the government reduces the impact of longevity risk on its balance sheet.

2.5 Human capital

This section considers human capital more closely. Subsection 2.5.1 explores how borrowing constraints due to the limited liability of human capital affect risk taking and intertemporal consumption smoothing. Subsection 2.5.2 considers the implications of non-insurable idiosyncratic human capital risk while systematic human-capital risks are investigated in 2.5.3. Subsection 2.5.4 analyzes what happens if labor supply becomes endogenous.

2.5.1 Borrowing constraints

The first-best asset allocation (8) implies that one should borrow at the beginning of one's career and invest the proceeds in the stock market to acquire sufficient exposure to the equity market. Adverse selection, moral hazard and limited liability of human capital, however, typically preclude borrowing against future labor income. Financial institutions cannot use human capital as a collateral to ensure that the loan is paid back. With borrowing constraints, agents must get all their risk exposure from positive financial capital. Hence, in contrast to older workers who can accumulate substantial financial capital through saving, young workers cannot get enough exposure to equity risk (see also Constantinides, Donaldson and Mehra (2002)). The restricted access of younger workers to the equity market harms their welfare. The associated negative income effects reduce consumption and thus boost saving initially. Another reason why borrowing constraints raise saving is that additional saving allows the young worker to acquire more access to equity markets so that they can take more advantage of the equity premium.

With our benchmark parameters, the presence of the borrowing constraints implies an additional welfare loss of approximately 2.8% of certainty-equivalent consumption (see Bovenberg et al. (2007)). To put these welfare loss of borrowing constraints into perspective, we compute the welfare costs of taking no risks at all during the life cycle. This is the most extreme form of limits on risk sharing. Individuals in effect do not have any risk-bearing capital at all. Teulings and de Vries (2006) compute the following approximation for the welfare loss (in terms of the relative decline in certainty-equivalent consumption) associated with this extreme limitation on risk taking:

$$\frac{1}{4}\frac{\lambda^2}{\theta}D. \tag{13}$$

The intuition behind this expression is that the average duration of risky investments is half the adult lifetime D and that the welfare loss for a year of riskless investment is given by half of the potential reward for risk[6] $\frac{1}{2} f\overline{\mu} = \frac{1}{2}\frac{\lambda^2}{\theta}$. The welfare loss thus rises with the reward to risk taking (i.e. the Sharpe ratio λ) and declines with the willingness to do so (i.e. relative risk aversion θ). With our benchmark parameters, this approximation yields a welfare loss of 12% (in terms of certainty equivalent consumption).

Borrowing constraints hamper not only risk taking but also intertemporal consumption smoothing. This is especially relevant for young high-skilled agents facing a steep career pattern. These agents would like to borrow against the future value of their human capital for the purpose of consumption smoothing. Rather than assuming a constant wage path over the life cycle, several recent papers consider the impact of actual career patterns over the life cycle in the presence of borrowing constraints (see e.g. Cocco, Gomes and Maenhout (2005) and Davis, Kubler and Willen (2005)).

2.5.2 Idiosyncratic risks

The idiosyncratic component of human capital risk can in principle be pooled. However, this would give rise to serious moral hazard problems; individuals would reduce their effort and claim that the resulting low wage income has been due to misfortune instead of shirking. Indeed, human capital is special in that it is not really tradable in financial markets and introduces market incompleteness. Whereas social insurances (such as disability and unemployment insurances) do provide partial insurance against idiosyncratic human capital risk, most insurances leave a substantial amount of non-insured human-capital risk with individuals.

Combined with borrowing constraints (and steep career patterns) limiting intertemporal consumption smoothing, the individual is not able to optimally diversify labor income risks over time. As a direct consequence of the reduced ability to smooth risks intertemporally, the individual will take less risk in the investment portfolio. Gollier (2005) shows that this may cause young households with small financial reserves to invest less in equity than older households who are less constrained by borrowing constraints in

[6]Only half of the reward to risk is actually a welfare gain. The reason is that the individual also bears additional risk, which harms welfare. In fact, one can interpret the welfare loss as a Harberger triangle due to a tax on risk taking. The tax on risk taking is given by $\bar{\mu}$ (since the reward for taking risk is taken away), while the behavioral response (as a share of overall wealth) is given by f. The Harberger triangle is thus given by $\frac{1}{2}f\bar{\mu}$.

diversifying risks over a longer period. In a similar vein, Cocco, Gomes and Maenhout (2005) find that young investors choose portfolios that are less tilted towards equity than the portfolios of middle-aged investors. Thus, the combination of labor-market risks and a steep career pattern preventing optimal intertemporal consumption smoothing (or the potential of large drops in labor income) turn around the result from unconstrained models that young households should hold substantial amounts of equity that exceed the amounts (both in absolute values and as a share of financial wealth) of older households. Indeed, the desire to hold a safe, liquid stock of precautionary savings has become one of the explanations for the equity premium puzzle and for why young households do not participate much in the stock market (see Constantinides and Duffie (1996) and Brav, Constantinides, and Geczy (2002)).

Liquidity constraints not only make investment behavior more conservative but also strengthen precautionary saving motives. Indeed, by investing conservatively and setting aside resources through saving, individuals ensure the presence of a financial buffer that helps them to optimally time diversify temporary risks.[7] Precautionary motives rather than saving for retirement tend to be the main reason why young households save (see Cocco, Gomes and Maenhout (2005)).

The behavior of individuals facing borrowing constraints resembles that of a pension fund facing solvency constraints on minimum buffer requirements. Both borrowing constraints and these solvency constraints harm the ability to time-diversify risk. They thus reduce risk-bearing capacities and encourage precautionary saving to escape these constraints.

2.5.3 Aggregate risks

An aggregate component to wage risk introduces other issues. First of all, to the extent that aggregate human-capital risk is correlated with equity risk, human capital becomes more like equity. As a direct consequence, the wealth share invested in equity declines compared the case in which human capital does not carry any macro-economic risk. This is especially so at the beginning of the working life when human capital constitutes the bulk of overall wealth (see e.g., Benzoni et al. (2007)).

[7] In the case of permanent shocks, optimal consumption smoothing dictates that consumption is reduced in line with income. Habit formation, however, can then still give rise to a desire to reduce consumption gradually and thus a need for a financial buffer.

Another issue arises if wage risk is not perfectly correlated with equity risk. In that case, a second macro-economic risk factor emerges. This risk factor can in principle be traded through wage-indexed bonds. In particular, young agents who are long on this risk factor can trade this risk with older agents who do not bear (much of) this risk. In practice, however, we do not observe wage-indexed bonds because of limited liability of human capital. A wage-indexed defined-benefit pension may, however, act as a substitute for tradable wage-indexed bonds. Through such a pension system, the younger generations share wage risk with the retired agents (see Beetsma and Bovenberg (2009)). Hence, compulsory pension schemes allow not only young workers with ample human capital to capture the equity risk premium but also older agents to capture part of the human-capital risk premium.

Recent research explores how idiosyncratic risk interacts with systematic macro-economic risks. In particular, non-insurable idiosyncratic shocks in human capital seem to feature countercyclical left-skewness: in recessions, large downward drops in labor earnings become more likely while large upward movements become less likely (see Govenen, Ozkan and Song (2012)). This systematic pattern of idiosyncratic risk deters households from holding stocks, which tend to decline in recessions (see also Constantinides and Duffie (1996) for the impact of countercyclical idiosyncratic wage on stock holdings).

2.5.4 Endogenous labor supply

Bodie et al. (1992) show that the flexibility to adjust labor supply may raise the willingness of individuals to take in stock-market risk. In particular, with endogenous labor supply, human capital amounts to the discounted value of *potential* wage income (i.e. wage income if no leisure would be consumed) and is thus larger than the discounted value of *actual* wage income. Hence, the stock of wealth is larger than without the flexibility to raise labor supply. As the share f_t in (7) applies now to a larger stock of wealth, the overall demand for risk-bearing capital rises. Indeed, shocks can be absorbed by adjusting consumption of both produced commodities and leisure.

If labor supply becomes more flexible over the entire life cycle, the share of financial wealth invested in risk-bearing assets, f_t^*, now varies even more with age. This changes if labor supply is especially flexible around retirement because agents can adjust the time and speed with which they retire. In that case, also older workers can afford to invest in risk-bearing assets. This points to the importance of a flexible labor market for older workers to sustain the supply of risk-bearing capital in an aging economy. A formal analysis of the impact of flexibility of the retirement date on the optimal risk taking is provided by Farhi and Panageas (2007).

2.6 Pre-labor-market-entry participation in the stock markets

Teulings and de Vries (2006) show that the individual first-best outcome explored in subsection 2.2 can be improved further if we allow for the possibility that agents assume financial risks already before they start working.

The welfare gain (in terms of certainty equivalent consumption) from optimally participating in stock-market risk B years before entering the work force can be approximated by[8] (see Teulings and de Vries (2006))

$$\frac{1}{2}\frac{\lambda^2}{\theta}B. \tag{14}$$

The intuition behind this expression is that the duration of the risky investments is B and that the welfare gain for a year of optimal participation in the stock market is given by half of the potential reward for risk[9] $\frac{1}{2}\,f\overline{\mu} = \frac{1}{2}\frac{\lambda^2}{\theta}$. The welfare gain thus rises with the reward to risk taking (i.e. the Sharpe ratio λ) and declines with the willingness to do so (i.e. relative risk aversion θ). For the benchmark parameters, an investor starting to participate in the stock-market 30 years before entering the work force would enjoy a welfare gain in terms of certainty equivalent consumption of 6.7%. This back-on-the-envelope calculation shows that starting to take equity risk before entering the labor market yields substantial welfare-gains.

The welfare gain in (14) can be realized only if an investor is able to borrow against his future labor earnings and invest the proceeds of the loan in the stock market when the investor has not even started his or her career. As explained in section 2.5.1, adverse selection, moral hazard and limited liability of human capital typically prevent investors from being able to engage in such financial transactions. Investors are therefore not able to take equity risk before they start to work in a laissez-faire economy. The associated

[8]This equation is closely related to equation (13). Note that the average length of the investment period in (14) is different, namely B rather than $\frac{1}{2}D$.

[9]Only half of the reward to risk is actually a welfare gain because the individual also bears additional risk, which harms welfare.

restriction $B = 0$ thus implies that the welfare gain in (14) cannot be realized in such an economy. Long-lived institutions such as a pension fund or a government, however, may be able to transfer financial buffers between non-overlapping generations, thereby effectively relaxing the constraint $B = 0$. Chapter 3 examines the possibilities and limitations of risk sharing between non-overlapping generations through these institutions.

3 Intergenerational risk sharing

The theory of general equilibrium of Arrow and Debreu (1954) states that the allocation of risk in financial markets is Pareto-efficient if financial markets are complete. The literature on intergenerational risk sharing considers a particular form of market incompleteness, namely that current generations are not able to trade with unborn generations because these generations are not present in financial markets at the same time. Financial markets allow trade only between generations that are present on financial markets at the same time because they are born close to each other. This 'biological trading constraint' can cause the allocation of risk across generations in financial markets to be inefficient. Indeed, subsection 2.6 pointed out that future generations can potentially reap welfare gains if they would be able to share in current risks. Unfortunately, these gains cannot be realized in a laissez-faire economy, because future generations do not participate in current financial markets. A long-lived entity (e.g. a government or a pension fund), however, can enhance welfare compared to a laissez-faire economy by implementing trade in risk between non-overlapping generations by transferring resources between these generations. Early contributions by Diamond (1977), Merton (1983) and Gordon and Varian (1988) already made this point. More recent contributions include Allen and Gale (1997), Shiller (1999), Bohn (2006, 2009, 2012), Smetters (2006), Ball and Mankiw (2007), Campbell and Nosbusch (2007), Gollier (2008) and Cui, de Jong and Ponds (2011).

3.1 Benchmark model

The section formulates a stylized model for exploring risk sharing between non-overlapping generations. A long-lived entity such as a pension fund allows currently-living generations to transfer (positive or negative) wealth to future generations. Generations live during a single period and only a single generation is alive at each point in time. Hence, the model abstracts from the life-cycle patterns discussed in chapter 2. Each period in the model stands for the average length of the investment period over the life-cycle of a single generation, say 30 years or so.[10] Generation t ($t \geq 1$) is referred to as the generation that lives during period t.

Preferences As in the benchmark model in chapter 2, generations maximize expected utility from consumption C_t:

$$U_t = \mathbf{E}_{t=0}\left[\frac{1}{1-\theta}C_t^{1-\theta}\right], \tag{15}$$

and θ denotes the coefficient of relative risk aversion with respect to consumption.[11]

Financial markets Just as the benchmark model of chapter 2, the present model includes a single source of risk. Furthermore, financial markets offer two investment opportunities: a riskfree asset and a risky asset. The riskfree asset yields a deterministic and time-invariant return r in each period t. The return on the risky asset in excess of the riskfree rate is denoted by $\tilde{x}_t$ ($t \geq 1$). It is i.i.d. distributed with a time-invariant mean $\bar{\mu}$ and standard deviation σ. The Sharpe ratio $\lambda \equiv \frac{\bar{\mu}}{\sigma}$ represents excess return per unit of risk. A tilde above a variable indicates its stochastic nature (the uncertainty is resolved during the period indicated by the subscript).

Labor earnings In line with the benchmark model of chapter 2, labor earnings Y_t are deterministic. $\eta_t \equiv \frac{1}{Y_t}\frac{Y_{t+1}}{1+r}$ stands for the discounted value of

[10]The difference between the moments when saving and dissaving occurs determines the duration of life-cycle investments. In the life-cycle model of chapter 2, saving occurs on average in the middle of the working period, $\frac{D-R}{2}$, while dissaving happens on average in the middle of the retirement period, $(D-R)+\frac{R}{2}$. Hence, the average duration of investments is thus approximated by $(\ (D-R)+\frac{R}{2}\ )-(\ \frac{D-R}{2}\ )=\frac{D}{2}$. With $D=60$, the average duration is thus 30 years.

[11]The choice for constant relative risk aversion (CRRA) in equation (15) can also be viewed as notational convenience, because it is not necessary for obtaining the results in this section. Under the Arrow-Pratt approximations that are applied in this section, all analytical expressions remain valid for any utility function $U_t = E_{t=0}[u(C_i)]$, with θ representing $-\frac{u''(C_i)}{u'(C_i)}$.

the labor earnings of generation $t+1$ relative to the income of the currently-living generation, and is assumed constant over time: $\eta_t = \eta$. The economy is dynamically efficient so that $\eta < 1$.

The consumption level C_t of agent t is given by labor income Y_t plus the proceeds of savings and investments, minus the risk-sharing transfer $\tilde{\tau}_{t+1}$ that is passed onto the next generation $t+1$, plus (only for generations $t \geq 2$) the risk transfer $\tilde{\tau}_t$ that is received from the previous generation $t-1$:

$$C_t = \begin{cases} \{(1+r) + f_t \tilde{x}_t\} Y_t - \tilde{\tau}_t & \text{for } t = 1 \\ \{(1+r) + f_t \tilde{x}_t\} Y_t - \tilde{\tau}_t + (1+r)\tilde{\tau}_{t-1} & \text{for } t > 1 \end{cases}, \tag{16}$$

where f_t denotes the fraction of gross (before intergenerational transfers) wealth of generation t that is invested in the risky asset. The remaining fraction $1 - f_t$ is invested in the riskfree rate.

Intergenerational transfers $\tilde{\tau}_t$ ($t \geq 1$) represents the transfer that is passed onto generation $t+1$ by generation t. Solvency requirements determine this buffer. In general, the transfer $\tilde{\tau}_t$ can be conditioned upon all shocks that occurred up to period t. Our benchmark model is restricted to the case in which the intergenerational transfer $\tilde{\tau}_t$ depends on period-t risk only. Linear transfers between generations then take the form:[12]

$$\tilde{\tau}_t = \alpha_t + \beta_t \tilde{x}_t, \tag{17}$$

where α_t represents the deterministic component of the transfer and where β_t represents the exposure of the transfer to period-t risk. The tilde in the notation of $\tilde{\tau}_t$ indicates that the transfer represents a stochastic variable that can be conditional on the state of the economy that materializes during period t.

Benchmark parameters in numerical calculations The benchmark parameters are taken from chapter 2 so that $\theta = 5$. We interpret each period t ($t \geq 1$) in the model as the average duration of investments over the life-cycle of a single generation, i.e. $N = \frac{D}{2} = 30$. Hence, with a lognormal distribution of i.i.d. stock returns, this implies an equity risk premium of $\bar{\mu} = N \times 3\,\% = 90\,\%$ and a standard deviation of stock returns of $\sigma = \sqrt{N} \times 20\,\%$.

[12] Subsection 3.4.3 explores also non-linear intergenerational transfers.

3.2 The first-best risk-sharing solution

This section explores the first-best intergenerational risk-sharing solution. The analytical expressions in this section are based on Arrow-Pratt approximations that use the first two moments (the mean and variance) of the return distribution. Subsection 3.2.1 explores optimal consumption smoothing across generations. Subsection 3.2.2 explores optimal risk taking, while subsection 3.2.3 analyzes the welfare gain from risk sharing compared to a laissez-faire economy.

3.2.1 Consumption smoothing across generations

In the absence of commitment problems, diversification of risk over generations is optimal because it exploits the possibilities for intertemporal consumption smoothing. By smoothing shocks across both individuals and periods, each separate individual in each period is affected as little as possible. The scope for intergenerational risk sharing is assymetric. Only generations who consume after a shock occurs can possibly share in that shock: generations who are deceased when a shock occurs ten no longer have any risk-bearing capacity. Future generations who have not yet started consuming, in contrast, can share in current risks. Hence, current shocks can in principle be smoothed across not only the currently-living generations but also future generations. In our benchmark setting, we limit risk sharing to two periods only. Hence, optimal consumption smoothing with CRRA utility across generations implies that that period-t risk is smoothed proportionally equally across the consumption of generations t and generation $t+1$: (see appendix)[13]

$$\frac{\frac{dC_{t+1}}{d\tilde{x}_t}}{C_{t+1}} = \frac{\frac{dC_t}{d\tilde{x}_t}}{C_t} \tag{18}$$

for all $t \geq 1$. The optimal diversification of risk thus has two dimensions, namely consumption smoothing of risk across the remaining life cycle of individuals (see (10)) and consumption smoothing across generations (see (18)).

Bohn (2009) points out that the observed public pension provisions around the world allocate risk inefficiently, because real-world pension schemes typ-

[13] In a more general setting, optimal consumption smoothing occurs across the current generation t and *all* future generations $s > t$.

ically provide relatively safe transfers to retirees and shift risk onto workers and future generations. Hence, (18) is violated. Safe pensions for retirees can be rationalized as efficient if preferences are not CRRA but instead display habit formation. As explained in subsection 2.3.2, intertemporal consumption smoothing after a shock is no longer optimal (i.e. $\frac{\frac{dC_{t+1}}{d\tilde{x}_t}}{C_{t+1}} > \frac{\frac{dC_t}{d\tilde{x}_t}}{C_t}$) in the case of habit formation.

3.2.2 The demand for transferrable risk

The improved diversification of risk across generations results in a higher risk-bearing capacity compared to a laissez faire economy. The human wealth of future generations who share in current risks determines the potential increase in risk-bearing capacity and the associated additional demand for risk. In our benchmark model, the relative value of the human wealth η of the next unborn generation $t+1$ determines the increase in the demand for the transferrable risk $\tilde{x}_t$ (see appendix):

$$f_t - \{f_t\}_{\tilde{\tau}_t=0} = \eta \frac{\bar{\mu}}{\gamma\sigma^2}, \tag{19}$$

where f_t denotes risk taking in the optimal risk-sharing solution, and where $\{f_t\}_{\tilde{\tau}_t=0} = \frac{\bar{\mu}}{\gamma\sigma^2}$ (see (2)) denotes optimal risk taking in the absence of intergenerational risk sharing. Intergenerational risk sharing thus yields a relative increase of $\frac{f_t - \{f_t\}_{\tilde{\tau}_t=0}}{\{f_t\}_{\tilde{\tau}_t=0}} = \eta_t$ in the demand for risk.

In our partial equilibrium setting, the risk premium is not affected by the additional demand for the risky asset. Hence, intergenerational risk sharing does not reduce risk but rather raises the risk taken by future generations (see Gollier (2008) and Cui, de Jong and Ponds (2011)). In a general equilibrium context in which the supply of risk-bearing assets would be less than infinitely elastic, however, the additional risk-bearing capacity of younger generations would reduce the reward for risk, thereby inducing current generations to bear less risk (see Campbell and Nosbush (2007)).

The introduction of an aggregate risk component to wage risk (see 2.5.3) introduces other considerations section. In particular, the willingness of future generations to share in current financial market risk is reduced when the human wealth of future generations is correlated with current financial shocks. This reduces the risk appetite of future generations and hence reduces the attractiveness of risk sharing. Bohn (2009) analyzes risk sharing in an overlapping generations model in which stock and labor markets are subject to a common aggregate risk factor, and estimates 30-year correlations between wages and share values.

3.2.3 Welfare gain from risk sharing

The welfare gain associated with optimal intergenerational risk sharing (expressed as the increase of the discounted sum of certainty equivalent consumption of the current generation t and the unborn generation $t+1$ as a fraction of the present value of the wealth of the unborn generation $t+1$) amounts to

$$\frac{1}{2}\frac{\lambda^2}{\theta}. \tag{20}$$

This welfare gain is entirely due to the possibility of generation $t+1$ to be able to share in period-t risk that materializes before this generation is born. Accordingly, (20) matches (14) with $B = 1$ because risk is shifted one period (i.e. 30 years) into the future in the benchmark model. Equation (20) yields a welfare gain of 6.7% for our benchmark parameters (just as was obtained with respect to (14) with $B = 30$ and the parameters measured in annual frequencies).

Cui, de Jong and Ponds (2011) explore risk sharing in a setting that mimics the working of real-world funded pension systems. They obtain a 2.3% increase in certainty equivalent consumption as a result of risk sharing.

3.2.4 Implementation of the risk-sharing solution

Various institutional arrangements can implement the optimal risk-sharing solution: namely, not only pension funds but also governments and families. In particular, public debt policy and public investments implement intergenerational transfers. Smetters (2006) points out that an appropriate chosen tax on capital can play an important role in implementing intergenerational risk sharing.

Risk sharing between non-overlapping generations can also occur within families (see Barro (1989)). Older cohorts leaving intentional bequests to their children can help to share risk between them. In the face of limited liability, this channel is full operative only in rich dynasties that leaves positive bequests to next generations. Although intergenerational risk sharing may thus not be very relevant in many families, it still may apply to a substantial share of financial wealth that is concentrated in the hands of the very wealthy.

3.3 Intergenerational fairness

Section 3.3.1 explores the set of ex-ante Pareto-efficient risk-sharing solutions. The Pareto-efficient solution is not unique: the welfare gain from risk sharing can be distributed across generations in different ways. Section 3.3.2 explores two alternative fairness criteria for allocating the welfare gain across generations. These two alternative fairness criteria result in distinct solvency requirements for the buffers that must be transferred to future generations. Section 3.3.3 explores value-based generational accounting as an instrument for assessing the intergenerational fairness of pension contracts.

3.3.1 Ex-ante Pareto-efficient risk-sharing solutions

With the help of aggregate certainty equivalent consumption, we can assess whether intergenerational risk sharing is potentially Pareto-improving in comparison to autarky. This approach is similar to the lump-sum redistribution authority in Auerbach and Kotlikoff (1987), but now applied in a stochastic environment. Aggregate certainty equivalent consumption is an attractive welfare measure in the context of risk sharing, because it is unaffected by deterministic redistributive transfers.[14]

The set of Pareto-efficient solutions is found by maximizing the discounted sum of the certainty equivalent consumption under the restriction that certainty equivalent consumption does not decline for any generation compared to the lassez faire economy without intergenerational risk sharing. The ex-ante Pareto-efficient solution is not unique, because the welfare gain from intergenerational risk sharing can be distributed across generations in various alternative ways.

The optimization of aggregate certainty equivalent consumption in our benchmark model yields the optimal risk-sharing conditions (18) and (19). The restriction that certainty equivalent equivalent consumption cannot de-

[14]The appendix shows that the maximization of aggregate equivalent variation in our benchmark model is unaffected by the deterministic component α_t of intergenerational transfers.

cline for any generations puts limits on α_t. The next subsection examines two criteria for intergenerational fairness yielding a unique solution for α_t.

3.3.2 Criteria for intergenerational fairness

This subsection examines two alternative criteria for intergenerational fairness that yield a unique risk-sharing solution within the set of of Pareto-efficient solutions. The two criteria differ in the way in which the welfare gain from risk sharing is distributed across generations and thus yield different solvency requirements for the capital buffers that must be reserved for future generations. Whereas the first criterion is utility-based, the second criterion is based on market value. The reason why criteria based on utility and market value yield different solutions has to do with the biological trading constraint. The associated market incompleteness causes a gap between the market prices of risk and unborn generations' marginal rate of substitution between various contingencies.

The first, utility-based, fairness criterion, which is applied by Gollier (2008), requires that all generations experience the same percentage increase in certainty equivalent consumption as a result of intergenerational risk sharing. In our benchmark model, this criterion implies that the welfare gain from sharing risk between generation t and $t+1$ (through the stochastic transfer $\tilde{\tau}_t$) fully accrues to the older generation t. The younger generation $t+1$ gains nothing from sharing risk with the older generation but enjoys the full benefit from sharing risk (through the stochastic transfer $\tilde{\tau}_{t+1}$) with generation $t+2$. This solution ensures that the first-born generation $t = 1$, which is alive when the intergenerational risk-sharing scheme is instituted, benefits just as much from this scheme in terms of relative certainty equivalent consumption as younger generations $t > 1$, even though it does not share in risks that materialized before it was born. The percentage increase in certainty equivalent consumption equals $\frac{1}{2}\frac{\lambda^2}{\theta}\eta$ for all generations $t \geq 1$.[15]

The second fairness criterion, which is adopted by Teulings and de Vries (2006), is based on ex-ante market value. It requires that no generation loses from risk sharing in terms of market value. Since intergenerational risk sharing is a zero-sum game in terms of market value, this criterion implies that the market value of risk sharing equals zero for all generations. Hence, intergenerational redistribution is ruled out in terms of market value. This

[15]This is the welfare gain in (20) adjusted with η because it is expressed in terms of the wealth of the old instead of the young generation involved in the risk sharing.

equilibrium would occur if future generations would be able to participate in capital markets at current prices. This criterion implies that $\alpha_t = 0$ for all t so that intergenerational transfers $t(\tilde{x}_t)$ take the form of a multiple of the excess return on risky assets.

With the second fairness criterion, each intergenerational risk transfer includes the full risk compensation that is provided by the market. This is attractive for the young generation, because they receive the full risk compensation from being exposed to current risk (see (14) and (20)). The first-born generation $t = 1$, in contrast, has already started working at the time when the risk-sharing contract is initiated. The risk-sharing contract does not increase its trading opportunities, and thus does not improve welfare compared to a laissez-faire economy.

In our benchmark model, the second criterion implies that certainty equivalent consumption for future generations $t > 1$ rises (in relative terms) by $\frac{1}{2}\frac{\lambda^2}{\theta}$ (see equation (20)), while certainty equivalent consumption for the first-born generation $t = 1$ is unaffected. Whereas the first, utility-based, criterion transfers the welfare gain associated with the stochastic transfer $\tilde{\tau}_t$ $(t \geq 1)$ between generation t and $t + 1$ to the older generation t, the second criterion transfers this welfare gain to the younger generation $t + 1$. The welfare gain for future generations in the second fairness criterion is larger than under the first criterion (i.e. $\frac{1}{2}\frac{\lambda^2}{\theta} > \frac{1}{2}\frac{\lambda^2}{\theta}\eta$) because it provides them with the full benefit from risk sharing. In the first fairness criterion, in contrast, they have to share the welfare gain with the generation that is alive when the risk-sharing scheme is introduced. Compared to the first criterion, the second criterion implies tighter requirements for minimum buffers $\tilde{\tau}_t$ that must be transferred to future generations.

3.3.3 Value-based generational accounting

Value-based generational accounting measures intergenerational redistribution in terms of ex-ante market value (see Ponds (2003)). The value of participation of a particular generation is calculated as the difference between the discounted ex-ante market value of all cash flows that the generation receives from the pension scheme and the discounted ex-ante market values of all resources that the generation pays into the scheme.

Value-based generational accounting addresses concerns with respect to intergenerational fairness. A pension fund is a zero-sum game in terms of ex-ante market value. A positive ex-ante market value for one generation thus necessarily leads to negative market value for one or more other generations. Generational accounting is particularly useful in the context of the fairness criterion based on market value discussed in the previous subsection 3.3.2. Koijen and Nijman (2006), Kortleve and Ponds (2006), Bikker and Vlaar (2007), Hoevenaars and Ponds (2008) and Broeders (2010) apply value-based generational accounting to assess the intergenerational distributional effects of Dutch occupational pension schemes.

Value-based generational accounting applies asset pricing theory (see e.g. Cochrane (2001)) to determine the ex-ante market value of stochastic cashflows in pension contracts. Asset pricing theory is based on the replication principle. When all risk factors are traded in financial markets, then any stochastic cashflow can be replicated by a portfolio strategy based on assets that are traded in financial markets. The value of any stochastic cashflow is then given by the price of its replicating portfolio strategy. The law of one price implies that this price is unique in a financial market that satisfies the no-arbitrage condition.

A pricing technique that is commonly used in the context of pension contracts is *risk-neutral pricing*, in which an adjusted probability measure is used to determine the value of stochastic payoffs. The adjusted probability measure is specified such that the value of any cashflow is equal to its discounted expected value when the expectation is taken under the adjusted probability probability measure. The adjusted probability measure is referred to as the *risk-neutral probability measure* and is different from the physical risk measure if financial markets are risk-averse. In particular, the risk-neutral probability measure attaches a higher weight to "bad" or "unpleasant" states of the world in which a payoff is more expensive than an equal payoff in a "good" or "pleasant" state of the world (with an equal physical probability of occurrence).

An alternative pricing technique is the *stochastic discount factor*. A stochastic discount factor can be interpreted as a "change of measure": a transformation from physical probabilities to risk-neutral probabilities, see e.g. Cochrane (2001). Pricing on the basis of stochastic discount factors is therefore equivalent to risk-neutral pricing and yields, by definition, the same pricing results.

A fundamental implication of the no-arbitrage condition in financial markets is that excess returns of risky assets (i.e. the return on risky assets in excess of the riskfree rate) have zero ex-ante market value. The economic intuition here is that any multiple of the excess return on the risky asset can be obtained in the financial market at zero costs by buying risky assets with money that is borrowed against the riskfree rate. Another way to see this is that, from an ex-ante perspective, the payoff from a one-dollar investment in the risky asset has the same market value as the one-dollar investment in the riskfree asset. The same euro cannot have different market values under the no-arbitrage condition. Put in terms of risk-neutral pricing, this implies that the expectation of the excess return on risky assets is equal to zero when calculated under the risk-neutral probability measure (because it has zero market value).

In our benchmark model, this implies

$$E^Q_{t=0}[\tilde{x}_t] = 0 \tag{21}$$

for all t, where Q denotes the risk-neutral probability measure and where $E^Q_{t=0}$ indicates that the expected value is taken under risk-neutral probabilities. The generational account GA_t of generation t is given by the ex-ante market value of the transfer $\tilde{\tau}_t$ that is passed onto generation $t+1$ discounted back to the beginning of period t, minus the ex-ante market value of the transfer $\tilde{\tau}_{t-1}$ that is received from generation $t-1$ (only for generations $t > 1$):

$$GA_t = \begin{cases} -E^Q_{t=0}[\tilde{\tau}_t]/(1+r) & \text{for } t = 1 \\ -E^Q_{t=0}[\tilde{\tau}_t]/(1+r) + E^Q_{t=0}[\tilde{\tau}_{t-1}] & \text{for } t > 1 \end{cases}. \tag{22}$$

The no-arbitrage condition in financial markets implies that the discounted sum of generational accounts amounts to zero: (see appendix)

$$\sum_{t=1}^{\infty} \frac{GA_t}{(1+r)^{t-1}} = 0. \tag{23}$$

In the context of the fairness criterion based on market value discussed in the previous subsection 3.3.2, intergenerational redistribution in market value is ruled out and all generational accounts are equal to zero. In contrast, the utility-based fairness criterion discussed in the previous subsection 3.3.2 is more beneficial to the first-born generation $t = 1$ in comparison to the criterion based on market value. Hence, the utility-based fairness criterion provides the first-born generation, which is alive at the time when the risk-sharing scheme is introduced, with a positive generational account (i.e. $GA_1 > 0$) at the expense of later-born generations (i.e. $GA_t < 0$ for all $t > 1$), see Appendix. This situation can be interpreted as a pay-as-you-go element inside a pre-funded pension scheme in terms of market value. In each period $t \geq 1$, the earlier-born generation t passes a transfer with a negative market value onto the later-born generation $t+1$, i.e. $E^Q[\tilde{\tau}_t] < 0$.

Market value can be an attractive criterion to assess the intergenerational fairness of pension schemes because the technique for the valuation of cashflows is based on observed prices in financial markets. No subjective assumptions with regards to the preferences of individuals are required. Such subjective assumptions can be difficult to determine objectively, which can be problematic in the situation where generations are involved in a divisive battle over pension scheme design. Unfortunately, market-consistent valuation may also require subjective assumptions as a result of market incompleteness. To illustrate, it may be hard to determine the value of cashflows with very long maturities because trade in assets with maturities beyond 30 years is limited.[16] Furthermore, subjective assumptions giving rise to model risk are required to value pension entitlements that depend on untraded risk factors. An example of such a risk factor is provided by aggregate wages since some pension promises are linked to aggregate wages. The paradox is thus that pension schemes that complete markets by trading untraded risk factors are difficult to value objectively and may thus give rise to intergenerational conflicts and political risks. A number of studies have explored valuation techniques for non-traded risk factors in pension schemes. De Jong (2008) examines the valuation of wage-linked cashflows in an incomplete market setting in which the wage index cannot be hedged perfectly with financial market instruments. He discusses several methods to find a value in such incomplete markets and advocates utility-based valuation. Geanakoplos and Zeldes (2010) derive the value of wage-linked cashflows on the basis of an assumed theoretical long-run relationship between wages and stocks.

Applying market value as a criterion for intergenerational fairness in a *partial* equilibrium framework raises questions. Arguably, it would be better to apply a *general* equilibrium framework in which current generations can trade with future generations and market prices would adjust. Such 'fictional' trading between non-overlapping generations would change the market prices of risk, thereby redistributing resources between agents (see Ball and Mankiw (2007)).

[16]This gives rise to the discussions about the so-called ultimate forward interest rate.

3.4 Commitment problems

3.4.1 Discontinuity risk

The previous analysis has shown that properly designed intergenerational risk-sharing contracts enhance welfare of all generations from an *ex-ante* perspective (i.e. *before* the economic shocks materialize that determine the size and direction of risk-sharing transfers between generations). The ex-ante perspective is similar to the “original position” introduced by John Rawls. It takes the point-of-view of the “original position” when the risk-sharing contract is initiated, i.e. before future cohorts start to work and pay into the system.

Some generations, however, may be worse off from the *ex-interim* perspective in which welfare is evaluated when these generations enter the labor force (i.e. *after* the economic shocks materialize that determine the size and direction of risk-sharing transfers between them and the earlier-born generations). Hence, the first-best risk-sharing solution can be implemented only if workers that join the labor force can be forced to participate in the risk-sharing contract even though it is no longer in their interest to join the contract when they start to work (i.e. from an ex-interim perspective).

The government is probably best equipped to commit future generations to a risk-sharing contract through its power to tax. To illustrate, by issuing indexed bonds and longevity bonds, the government protects current generations against inflation and longevity risks and shifts these risks in part to future generations of taxpayers through debt policies and generation-specific taxes and transfers.

Even the national state, however, is unable to perfectly commit future generations to a risk-sharing contract. This section examines the discontinuity risk that results from imperfect commitment.

3.4.2 Discontinuity risk resulting from ”voting by feet”

If young workers face substantial negative buffers on entry into a pension scheme, they may *vote with their feet* by seeking employment elsewhere (for example, abroad) or by reducing labor supply. These behavioral responses introduce a trade-off between facilitating intergenerational consumption smoothing and minimizing distortions in the labor market (see Bonenkamp and Westerhout (2013)). On the one hand, labor market distortions are avoided without intergenerational transfers so that intergenerational risk sharing is absent altogether. On the other hand, perfect intergenerational consumption smoothing involves large transfers in some contingencies, thereby distorting labor supply substantially.[17] The optimal solution lies in between these two extremes.

The trade-off between consumption smoothing and minimizing distortions can be modeled by introducing elastic labor supply. In our benchmark model, the introduction of elastic labor supply modifies the expression for optimal intergenerational consumption smoothing (18) into: (see appendix)

$$\frac{\frac{dC_{t+1}}{d\tilde{x}_t}}{C_{t+1}} = \frac{\gamma}{\gamma+\epsilon}\frac{\frac{dC_t}{d\tilde{x}_t}}{C_t}, \tag{24}$$

where ϵ denotes the compensated wage elasticity of labor supply. (24) shows that perfect consumption smoothing across generations is no longer optimal if endogenous labor supply is endogenous (i.e. $\frac{\frac{dC_{t+1}}{d\tilde{x}_t}}{C_{t+1}} < \frac{\frac{dC_t}{d\tilde{x}_t}}{C_t}$ if $\epsilon > 0$). In this case, solvency requirements limit intergenerational risk sharing. As labor supply becomes more elastic, the share of current risks that is shifted onto future generations becomes smaller so that the additional demand for risky assets from future generations is smaller than in the presence of inelastic labor supply. In the extreme case in which labor supply is infinitely elastic ($\epsilon \to \infty$), current generations bear all current risks and the laissez-faire economy is in fact optimal.

Labor supply flexiblity (giving rise to more elastic labor supply) may increase the demand for the risky asset in an individual system on account of income effects; in case of negative (positive) financial shocks individuals can self-insure themselves by working more so that employment behaves in a counter-cyclical fashion (see Bodie, Merton and Samuelson (1992)). With intergenerational risk sharing, in contrast, substitution effects in labor supply work in the opposite direction by calling for tighter solvency requirements. Hence, flexibility in labor supply reduces the risk-bearing capacity of an economy and depresses the demand for risky assets.

The welfare losses from labor supply distortions and the associated discontinuity risk erode the potential welfare gains on account of integenerational risk sharing. Introducing elastic labor supply in our benchmark model

[17]Distortions in the labor market that are due to intergenerational risk sharing become more costly if they add to already existing labor-market distortions originating in other public policies (e.g. existing labor-income taxes).

modifies the welfare gain (20) into: (see appendix)

$$\frac{1}{2}\frac{\lambda^2}{\gamma}\frac{\gamma}{\gamma+\epsilon}, \tag{25}$$

so that the welfare gain from risk sharing declines with the wage elasticity of labor supply ϵ.

Employer-based or industry-wide occupational pension schemes face substantially larger commitment problems than public schemes that control the entire labor market of a country. Whereas workers can escape contributions to a nation-wide pension fund only by working less or by moving abroad, they can escape occupational pension contributions by moving to another sector or by becoming self-employed. Accordingly, it terms of (25), an occupational pension pension scheme faces a higher wage elasticity ϵ in comparison to a public, nationwide scheme. With perfectly mobile labor across sectors, employers with workers who have to participate in sectoral pension schemes may have to pay compensating wage differentials to attract workers to their sector if these workers are forced to service implicit debt extended by the sectoral pension scheme. In a perfectly-competitive labor market without industry-specific human capital, industry-wide pension schemes cannot extract any rents so that the owners of the firm rather than the workers are on the hook for a shortfall in the pension fund. This calls for tighter solvency regulations on buffers that must be accumulated so that pension schemes can still pay their pension commitments even though the sponsoring companies or sectors decline or fail altogether.

Bohn (2012) shows how firm-specific or industry-specific human capital allows occupational pension schemes to engage in some intergenerational risk sharing by limiting the sensitivity of firm-specific or sector-specific labor supply to wage differentials with others firms or industries. Also other factors can make it unattractive for workers to switch employers, such as limited portability of pension rights or implicit labor contracts involving deferred wages. Whereas private pension schemes may face more discontinuity risk and thus need tighter solvency regulations than public national schemes because they pool a smaller subset of workers, these schemes may enjoy some strengths compared to public schemes. First of all, they are typically less vulnerable to political risk than public schemes. Indeed, a trade-off exists between alleviating political risks (in private schemes) and reducing discontinuity risks (in public schemes). Second, investments in the private firms by private occupational pension schemes are less controversial and politically charged than such investments by public schemes. The latter give rise to concerns that the government effectively nationalizes part of the economy. Indeed, investment decisions by public schemes may be driven by short-term political considerations. However, concerns about legitimate and effective governance may beset also occupational schemes. In particular, whether the governing board of an occupational scheme represents the interests of all stakeholders (including future generations) may be questioned, especially if labor unions appoint trustees but lack the legitimacy to act on behalf of non-union workers.

3.4.3 Discontinuity risk resulting from "voting by voice"

If young workers face substantial negative buffers on entry into a pension scheme, they may not only *vote with their feet* but also *vote with their voice* by exerting political pressure to change the rules of the game that commits them to pick up the risks that occurred in the past. In particular, they may encourage the pension scheme to default on the pension obligations to older workers and retirees, for example by no longer indexing their pensions to inflation. Similarly, in the presence of large buffers due to positive shocks in the past, current generations, who govern the scheme as future generations are not on the scene, can change the existing contract and consume buffers that were assigned by the contract to future generations. Hence, whereas after large positive shocks older generations may break the intergenerational contract, younger generations may stop participating if shocks are rather negative. These political-economy considerations challenge the enforcement of credible contracts between generations that are not all present at the table when the contract is designed.

Commitment problems that relate to voting by voice can be modelled in terms of a discrete choice variable (i.e. a 0/1 variable) that represents the political decision making of young workers. Transfers from young to old cannot be too big – otherwise the young will abolish the current contract and commence a new one.[18] the willingness of the young to participate depends on their belief about the willingness of the future young to do so as well. If exerting abolishing a pension contract involves exit costs for young generations, then a risk-sharing contract can be sustainable if the absolute size of the negative buffers that are shifted onto new workers is limited to

[18] Vice-versa, the old will want to renegotiate if transfers to the young are too big.

the size of these exit costs.[19][20]

Managing discontinuity risk therefore requires non-linear option-like contracts that put a lower bound on the absolute size of transfers.[21] Indeed, these non-linear contracts closely correspond to those implied by solvency rules that supervisors impose on pension schemes in order to limit discontinuity risk. Gollier (2008) explored intergenerational risk sharing under the constraint that, at any point in time, the fund must be large enough to repay the contributions made in the past. This solvency constraint limits the absolute size of the negative transfer to future generations. Gollier (2008) shows that the ex-ante welfare gains from intergenerational risk sharing decline compared to the case in which no solvency constraints are imposed. Bovenberg, Koijen, Nijman and Teulings (2007) calculate the decline in the ex-ante welfare gains from intergenerational risk sharing if exit costs are modelled in terms of a number of months of salary. They find that the welfare gain from risk sharing declines by more than half if new workers are willing to participate in the contract only if their initial funding deficit is limited to six months of salary. The probability that this constraint is reached is 58%. If new participants are willing to put two years of salary at risk, the majority of the welfare gain from risk sharing is preserved, while the probability that the maximum of two years of salary will be lost is 14%.

Solvency constraints yield important consequences also for optimal saving and investment. As regards saving, the solvency constraint strengthens the incentives to save for precautionary reasons if a pension scheme is close to the solvency constraint. Indeed, the scheme would like to put aside additional solvency buffers to escape the constraint on intergenerational risk sharing; the financial reserves of the fund in fact act like a buffer stock. As

[19]The exit costs for young workers associated with abolishing a pension contract may involve losing access to the scale advantages of occupational schemes or the trading of risk factors within occupation schemes that are not traded on financial markets. To illustrate, Beetsma, Romp and Vos (2012) assume that individuals can benefit from risk-free promises only if they participate in occupational schemes. Hence, exit costs become large if individuals are rather risk averse.

[20]Paradoxally, new workers are more willing to participate if the gains from intergenerational risk sharing go to older generations (conform the utility-based fairness criterion). The reason is that workers who enter the labor force lose these welfare gains if they opt to exit.

[21]Other types of dynamic portfolio strategies that reduce downward risk include portfolio strategies under a value-at-risk constraint (Basak and Shapiro (2001)) or under a shortfall constraint (Gundel and Weber (2008)).

regards investment, the scheme takes less investment risk if it is close to the solvency constraints. Intuitively, the pension scheme constructs a put option by trading dynamically and reducing its portfolio share of risk-bearing assets if solvency buffers become small. The fund in fact behaves as if it exhibits decreasing absolute risk aversion. This behavior of selling equity if the stock market declines contrasts with the rebalancing behavior of an individual with CRRA preferences and without borrowing constraints.

4 Conclusions

Pension policies conduct two tasks, namely implementing optimal individual life-cycle behavior and arranging optimal risk sharing between generations. This paper discusses these two tasks in turn. As regards optimal life-cycle behavior, chapter 2 explores the academic literature on optimal saving and investment over the life cycle. We started out with a simple benchmark model in which human capital is riskless, standard preferences apply and equity risk is the only risk factor. Subsequently, we explored a number of extensions, namely more elaborate models of human capital, additional risk factors in financial markets as well as non-standard specifications of preferences. As regards intergenerational risk sharing, Chapter 3 explored models of optimal intergenerational risk sharing in order to discusses the benefits and costs of sharing risks across non-overlapping generations. We discussed alternative intergenerational fairness critera for sharing the gains from intergenerational risk sharing as well as tools for measuring the benefits and costs of stochastic intergenerational transfers for each generation. Moreover, the rationale for solvency requirements in terms of not only enhancing intergenerational fairness but also reducing discontinuity risks were explored.

5 References

Allen F, Gale D. 1997. Financial markets, intermediaries, and intertemporal smoothing. *Journal of Political Economy*.105: 523-545.

Ang A, Bekaert G. 2007. Stock return predictability: is it there? *The Review of Financial Studies*. 20: 651-707.

Arrow KJ, Debreu G. 1954. Existence of an equilibrium for a competitive economy. *Econometrica*. 22: 265-290

Auerbach AJ, Kotlikoff LJ. 1987. *Dynamic Fiscal Policy.*Cambridge University Press.

Ball L, Mankiw N. 2007.Intergenerational Risk Sharing in the Spirit of Arrow, Debreu and Rawls, with Applications to Social Security Design. *Journal of Political Economy.* 115(4): 523-547.

Barbaris N, Huang M, Santos T. 2001. Prospect Theory and Asset Prices. *Quarterly Journal of Economics.* 116: 1-53.

Barro RJ.1989.The Ricardian Approach to Budget Deficits. *Journal of Economic Perspectives.* 3(2): 37-54.

Basak S, Shapiro A. 2001. Value-at-risk-based risk management: optimal policies and asset prices. *Review of Financial Studies.* 14(2): 371-405.

Beetsma RMWJ, Bovenberg AL. 2009. Pensions, intergenerational risk sharing and inflation. *Economica.* 76: 364-386.

Beetsma RMWJ, Romp WE, Vos SJ. 2012. Voluntary participation and intergenerational risk sharing in a funded pension system. *European Economic Review.* 56(6): 1310-1324.

Benzoni L, Collin-Dufresne P, Goldstein RS. 2007. Portfolio Choice over the Life-Cycle when the Stock and Labor Markets are Cointegrated. *Journal of Finance.* 62(5): 2123-2167.

Berkelaar A, Kouwenberg R, Post T. 2004. Optimal Portfolio Choice under Loss Aversion. *Review of Economics and Statistics.* 86: 973-986.

Bikker JA, Vlaar PJG. 2007. Conditional Indexation in Defined Benefit Pension Plans in the Netherlands. *The Geneva Papers on Risk and Insurance, Issues and Practice.* 32: 494-515.

Bodie Z, Merton RC, Samuelson WF.1992. Labor Supply Flexibility and Portfolio Choice in a Life-Cycle Model. *Journal of Economic Dynamics and Control.* 16: 427-449.

Bohn H. 2006. Who Bears What Risk? An Intergenerational Perspective. In *Restructuring Retirement Risks*, ed D Blitzstein, O Mitchell, SP Utkus, 2: 10-37. Oxford: University Press.

Bohn H. 2009. Intergenerational risk sharing and fiscal policy. *Journal of Monetary Economics.* 56(6): 805-816.

Bohn H. 2012. Private versus public risk sharing: should governments provide reinsurance?. In *The future of multi-pillar pensions*, ed AL Bovenberg, C van Ewijk, E Westerhout. 6: 187-224. Cambridge: University Press. 421 pp.

Bonenkamp J, E Westerhout. 2013. Intergenerational risk sharing and endogenous labour supply within funded pension schemes. *Economica.* Forthcoming.

Bovenberg AL, Koijen RSJ, Nijman TE, Teulings CN. 2007. *Saving and investing over the life cycle and the role of collective pension funds.* Netspar Panel paper No. 1.

Brav A, Constantinides GM, Geczy CC. 2002. Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence. *Journal of Political Economy.* 110: 793-824.

Campbell JY, Cochrane JH. 1999. By force of habit: a consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy.* 107: 205-251.

Campbell J, Viceira L. 2001. Who should buy :Long term Bonds?. *American Economic Review.* 91: 99-127.

Campbell J, Viceira L. 2005. The Term Structure of the Risk-Return Trade-off. *Financial Analysts Journal.* 61(1): 34-44.

Campbell J, Yogo M. 2006. Efficient tests of Stock Return Predictability. *Journal of Financial Economics.* 81: 27-60.

Campbell JY, Nosbusch Y. 2007. Intergenerational risksharing and equilibrium asset prices. *Journal of Monetary Economics.* 54: 2251-2268.

Cocco JF, Gomes FJ, Maenhout PJ. 2005. Consumption and Portfolio Choice over the Life Cycle. *The Review of Financial Studies* 18: 492-533.

Cochrane J. 2001. *Asset Pricing.* Princeton, University Press.

Cochrane J. 2008. The Dog that did not bark: a Defense of Return Predictability. *The Review of Financial Studies.* 21: 1533-1575.

Constantinides GM. 1990. Habit formation: A resolution of the equity-premium puzzle. *Journal of Political Economy.* 48: 519-543.

Constantinides GM, Duffie D. 1996. Asset Pricing with Heterogeneous Consumers. *Journal of Political Economy.* 105: 219-240.

Constantinides GM, Donaldson JB, Mehra R. 2002. Junior can't borrow; A new perspective on the equity premium puzzle. *Quarterly Journal of Economics.* 117: 269-296.

Cui J, de Jong F, Ponds E. 2011. Intergenerational risk sharing within funded pension schemes. *Journal of Pension Economics and Finance.* 10: 1-29.

Davidoff T, Brown JA, Diamond P. 2005. Annuities and Individual Welfare. *American Economic Review.* 95: 1573-1590.

Davis SJ, Kubler F, Willen P. 2005. *Borrowing Costs and the Demand for Equity over the Life Cycle.* Work. Pap. 05-7, Federal Reserve Bank of Boston.

Diamond P. 1977. A framework for social security analysis. *Journal of Public Economics*. 8(3): 275-298.

Farhi E, Panageas S. 2007. Saving and Investing for Early Retirement: A Theoretical Analysis. *Journal of Financial Economics*. 83: 87-121.

Geanakoplos J, Zeldes SP. 2010. Market Valuation of Accrued Social Security Benefits. In *Measuring and Managing Federal Financial Risk* ed. D Lucas, 213-233. Chicago: University of Chicago Press.

Gollier C. 2001. *The Economics of Risk and Time*. MIT Press.

Gollier C. 2005. *Understanding Saving and Portfolio Choices with Predictable Changes in Asset Returns*. Work. Pap. University of Toulouse.

Gollier C. 2008. Intergenerational risk sharing and risk taking in a pension fund. *Journal of Public Economics*. 92: 1463-1485.

Gordon RH, Varian HR. 1988. Intergenerational risk sharing. *Journal of Public Economics*. 37(2): 185-202.

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

Govenen F, Ozkan S, Song J. 2012. *The nature of countercycical income risk* Work. Pap. W18035., NBER.

Greenwood J. Hercowitz Z, and Huffman G. 1988. Investment, capacity utilitzation, and the real business cycle. *American Economic Review*, 78: 402-417.

Gundal A, Weber S. 2008. Utility maximization under a shortfall constraint. *Journal of Mathematical Economics*. 44(11): 1126-1151.

Hoevenaars RPMM, Ponds EHM. 2008. Valuation of intergenerational transfers in collective funded pension schemes. *Insurance: Mathematics and Economics*. 42(2): 578-593.

Jong, de FCJM. 2008. Valuation of pension liabilities in incomplete markets. *Journal of Pension Economics and Finance*. 7(3): 277-294.

Koijen RSJ, Nijman TE. 2006. Valuation and Risk Management of Inflation-Sensitive Pension Rights. In *Fair Value and Pension Fund Management*, ed N Kortleve, TE Nijman, E Ponds. Elsevier Publishers.

Kortleve N, Ponds EHM. 2006. Pension Deals and Value-Based ALM, chapter, In *Fair Value and Pension Fund Management*, ed N Kortleve, TE Nijman, E Ponds. Elsevier Publishers.

Leland HE. 1968. Savings and uncertainty: The precautionary demand for savings. *Quarterly Journal of Economics*. 45: 621-636.

Lettau M, van Nieuwerburgh S. 2008. Reconciling the return predictability evidence. *Review of Financial Studies*. 21: 607-1652.

Merton RC. 1969. Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics*. 51: 247-257.

Merton RC. 1971. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*. 3: 373-413.

Merton RC. 1983. On the role of Social Security as a means for efficient risk sharing in an economy where human capital is not tradable. In *Financial Aspects of the United States Pension System*, ed. Z. Bodie J. Shoven, Chicago: University of Chicago Press for the NBER.

Merton RC, Samuelson PA. 1974. Fallacy of the log-normal approximation to optimal portfolio decision-making over many periods. *Journal of Financial Economics*. 1: 67-94.

Mossin J. 1968. Optimal multiperiod portfolio policies. *Journal of Business*. 42(2): 215-229.

Ponds E. 2003. Pension funds and value-based generational accounting. 2(3): 295-325.

Prescott EC, Rios-Rull JV. 2005. On the equilibrium concept for overlapping generations organizations. *International Economic Review*. 46(4): 1065-1080.

Poterba J. 2006. Annuity markets. In *The Oxford Handbook of Pensions and Retirement Income*, ed. G. Clark, A.Munnel and M. Orszag, Oxford: University Press.

Samuelson PA. 1969. Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*. 51: 239-246.

Shiller RJ. 1999. Social Security and Institutions for Intergenerational, Intragenerational, and International Risk-Sharing, *Carnegie-Rochester Conf. Ser. Public Policy*. 50: 165204.

Smetters K. 2006. Risk Sharing across generations without publicly owned equity. *Journal of Monetary Economics*. 53: 1493-1508.

Teulings C, de Vries C. 2006. Generational Accounting, Solidarity, and Pension Losses. *De Economist*. 146: 63-83.

Tversky A, Kahnemann D. 1992. Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*. 5: 297-323.

A Appendix

A.1 Technical derivation of expressions in chapter 2

Optimization of individual decision making

In order to come to the optimal investment and consumption strategies in the life-cycle model of chapter 2, we proceed backwards. The Hamilton-Jacobi-Bellman equation for the problem is given by, with $J(W,t)$ denoting the value function:

$$\rho J = \max_{x,C}\left(\frac{1}{1-\theta}C^{1-\theta} + J_F(F(f^*\sigma\lambda + r) + 1 - C) + \frac{1}{2}J_{FF}(Ff^*\sigma)^2 + J_t\right). \tag{1}$$

where the subscripts of the value function J denote first- and second order derivatives with respect to financial wealth F_t. We know (see e.g. Bodie, Merton and Samuelson (1992)) that the value function is of the form:

$$J(F,t) = \frac{(F_t + H_t)^{1-\theta}}{1-\theta}g(D-t)^{\theta}, \tag{2}$$

where $g(\cdot)$ is a funtion of time and where H_t denotes the human capital of the investor at time t. We have for $t \in [0,T]$:

$$H_t = \frac{1}{r}\left(1 - e^{-r(T-t)}\right) \tag{3}$$

and $H_t = 0$ for $t \in (T,D]$. Substituting of the value function and solving for the resulting first order conditions in (1), we find for the optimal consumption and portfolio policy:

$$\begin{aligned} C_t &= (F_t + H_t)g(D-t)^{-1}, & (4)\\ f_t^* &= \frac{1}{\theta}\frac{\lambda}{\sigma}\left(1 + \frac{H_t}{F_t}\right). & (5) \end{aligned}$$

Let $W_t \equiv F_t + H_t$ denote the total wealth of the investor while f_t denotes the fraction of total wealth invested in risk bearing assets. From this it follows that $f_t^* = f_t(1+\frac{H_t^i}{F_t^i})$. The results in equations (6) and (7) now directly follow from equation (5). Substitution of the optimal policies (4) and (5) into (1) provides a linear ordinary differential equation for the function $g(D-t)$ which is solved for:

$$g(D-t) = \frac{1}{A}(e^{A(D-t)} - 1), \tag{6}$$

where

$$A = \frac{(1-\theta)r - \rho}{\theta} + \frac{1}{2}\frac{1-\theta}{\theta^2}\lambda^2. \tag{7}$$

Derivation of consumption dynamics

Applying Ito's lemma to (4), we find that the consumption dynamics are given by:

$$dC_t = g(D-t)^{-1}dW_t - \frac{\dot{g}(D-t)(W_t)}{g(D-t)^2}dt. \tag{8}$$

where $\dot{g}(\cdot)$ denotes the first-order derivative of function $g(\cdot)$. Employing the optimal portfolio rule given in (5), we find for the wealth dynamics:

$$\frac{dW_t}{W_t} = \left(\frac{\lambda^2}{\theta} + r - g(D-t)^{-1}\right)dt + \frac{\lambda}{\theta}dZ_t. \tag{9}$$

Using the expression for $g(D-t)$ in (7), we have:

$$-g(D-t)^{-1} - \frac{\dot{g}(D-t)}{g(D-t)} = A. \tag{10}$$

Substituting this result into of (10) and (9) into (8), we establish:

$$\frac{dC_t}{C_t} = \left(\frac{\lambda^2}{\theta} + r + A\right)dt + \frac{\lambda}{\theta}dZ_t. \tag{11}$$

Substitution of (7) and taking expectations on both sides yields

$$E_t\frac{dC_t}{C_t} = \frac{\left(r - \rho + \frac{(1+\theta)}{2}\frac{\lambda^2}{\theta^i}\right)}{\theta}dt, \tag{12}$$

which corresponds to equation (9). Excluding risky asset investments, i.e. $\lambda = 0$, yields equation (5), in which $\phi = \theta$, such that the intertemporal elasticity of substitution of the individual is given by $\frac{1}{\theta}$.

Consumption smoothing

In this section we derive in which way a shock in wealth (dW_t) at time t affects the consumption level (dC_s) at any time $s \geq t$. From equation (4) it follows that:

$$\frac{d\log C_t}{d\log W_t} = 1 \tag{13}$$

since $g(D-t)$ is a deterministic function of time. From equation (11) it follows that consumption growth is independent of wealth such that:

$$\frac{d\log C_s}{d\log C_t} = 1 \tag{14}$$

Combining those two results we obtain equation (10) :

$$\frac{d\log C_s}{d\log W_t} = \frac{d\log C_s}{d\log C_t}\frac{d\log C_t}{d\log W_t} = 1 \tag{15}$$

A.2 Technical derivation of expressions in chapter 3

Optimal solution in absence of risk sharing

In absence of risk sharing (i.e. $\tilde{\tau}_s = 0$ for all s) the utility of generation t is given by (substitute (16) in (15) to eliminate C_t)

$$\{U_t\}_{\forall s:\tilde{\tau}_s=0} = E_{t=0}\left[\frac{1}{1-\theta}\left((1+r)Y_t + f_t Y_t \tilde{x}_t\right)\right] \tag{16}$$

The Arrow-Pratt approximation (see Gollier (2001, 2008))

$$\begin{aligned} & E_{t=0}\left[\frac{1}{1-\theta}\left((1+r)Y_t + f_t\tilde{x}_t Y_t\right)^{1-\theta}\right] \\ \approx\ & \mathbf{E}_{t=0}\left[\frac{1}{1-\theta}\left((1+r)Y_t + f_t Y_t\bar{\mu} - \frac{1}{2}\frac{\theta}{Y_t}f_t^2 Y_t^2\sigma^2\right)^{1-\theta}\right] \end{aligned} \tag{17}$$

is applied to obtain (substitute (17) into (16)):

$$\{CEC_t\}_{\forall s:\tilde{\tau}_s=0} \approx (1+r)Y_t + f_t Y_t\bar{\mu} - \frac{1}{2}\frac{\theta}{Y_t}f_t^2 Y_t^2\sigma^2. \tag{18}$$

in which CEC_t denotes the certainty equivalent consumption level of generation t and is defined as the non-stochastic consumption level that yields U_t:

$$\mathbf{E}_{t=0}\left[\frac{1}{1-\theta}\left(CEC_t\right)^{1-\theta}\right] \equiv U_t \tag{19}$$

In (18), the term $f_t \bar{\mu} Y_t$ represents the excess return on risky investments, and the term $\frac{1}{2}\frac{\theta}{Y_t} f_t^2 Y_t^2 \sigma^2$ represents the welfare loss associated with the volatility in consumption that is due to risk taking. The first-order derivative of (18) with respect to f_t solves the optimal risk exposure:

$$\{f_t\}_{\forall s:\tilde{\tau}_s=0} = \frac{\lambda}{\theta\sigma}. \tag{20}$$

Substitution of (20) in (18) to eliminate f_t yields:

$$\{CEQ_t\}_{\forall s:\tilde{\tau}_s=0} = (1+r)Y_t + \frac{1}{2}\frac{\lambda^2}{\theta}Y_t. \tag{21}$$

in which the second term $\frac{1}{2}\frac{\lambda^2}{\theta}Y_t$ represents the welfare gain associated with the exposure to period-t risk.

Optimal risk sharing solution

In the presence of risk sharing transfers, utility U_t is given by (substitute (16) in (15) to eliminate C_t)

$$U_t = \begin{cases} \mathbf{E}_{t=0}\left[\frac{1}{1-\theta}\left((1+r)Y_t + f_t Y_t \tilde{x}_t - \tilde{\tau}_t\right)^{1-\theta}\right] & \text{for } t=1 \\ \mathbf{E}_{t=0}\left[\frac{1}{1-\theta}\left((1+r)Y_t + f_t Y_t \tilde{x}_t - \tilde{\tau}_t + (1+r)\tilde{\tau}_{t-1}\right)^{1-\theta}\right] & \text{for } t>1 \end{cases}. \tag{22}$$

with $\tilde{\tau}_t$ specified in (17). Applying an Arrow-Pratt approximation (along the lines of (17)) yields the expression for the certainty equivalent consumption CEC_t (defined in (19)):

$$CEC_t = \begin{cases} (1+r)Y_t + (1-\hat{\beta}_t) f_t Y_t \bar{\mu} - \frac{1}{2}\frac{\theta}{Y_t}(1-\hat{\beta}_t)^2 f_t^2 Y_t^2 \sigma^2 - \alpha_t & \text{for } t=1 \\ (1+r)Y_t + (1-\hat{\beta}_t) f_t Y_t \bar{\mu} - \frac{1}{2}\frac{\theta}{Y_t}(1-\hat{\beta}_t)^2 f_t^2 Y_t^2 \sigma^2 - \alpha_t & \\ \quad +(1+r)\left\{\hat{\beta}_{t-1} f_{t-1} Y_{t-1} \bar{\mu} - \frac{1}{2}\frac{\theta}{Y_t}\hat{\beta}_{t-1}^2 f_{t-1}^2 Y_{t-1}^2 \sigma^2 + \alpha_{t-1}\right\} & \text{for } t>1 \end{cases}. \tag{23}$$

in which $\hat{\beta}_t \equiv \frac{\beta_t}{f_t Y_t}$ represents the fraction of period-t risk that is transferred to generation $t+1$. The remaining fraction $1-\hat{\beta}_t$ is born by generation t.

The optimal risk sharing solution solves from the maximization of the discounted sum of certainty equivalent consumption:

$$\max_{f_t, \alpha_t, \hat{\beta}_t} \sum_{i=1}^{\infty} \frac{CEC_i}{(1+r)^{i-1}}. \tag{24}$$

It follows immediately that the optimal risk sharing policy is unaffected by deterministic transfers α_t (substitution of (23) into (24) to eliminate CEC_t yields an expression that does not depend on α_t).

The optimal allocation of risk across generations solves from substitution of (23) into (24) and taking the first-order derivative with respect to $\hat{\beta}_t$:

$$\hat{\beta}_t = \frac{\eta}{1+\eta} \tag{25}$$

where $\eta \equiv \frac{1}{Y_t}\frac{Y_{t+1}}{1+r}$. Equation (25) states that period-t risk is allocated across generations t and $t+1$ proportional to relative wealth. This leads to consumption smoothing equation (18) in the text (substitute (17) and (4) into (16) to eliminate τ_t and β_t and take the partial derivative with respect to $\tilde{x}_t$).

The optimal risk exposure solves from from substitution of (23) and (25) into (24) and taking the first-order derivative with respect to f_t:

$$f_t = \frac{\lambda}{\theta\sigma}(1+\eta) \tag{26}$$

Equation (19) then follows from subtracting (20) from (26).

Substitution of (25) and (26) into (23) yields:

$$CEC_t = \begin{cases} (1+r)Y_t + \frac{1}{2}\frac{\lambda^2}{\theta}Y_t - \alpha_t & \text{for } t=1 \\ (1+r)Y_t + \frac{1}{2}\frac{\lambda^2}{\theta}Y_t - \alpha_t + \frac{1}{2}\frac{\lambda^2}{\theta}Y_t + (1+r)\alpha_{t-1} & \text{for } t>1 \end{cases}. \tag{27}$$

in which the second term $\frac{1}{2}\frac{\lambda^2}{\theta}Y_t$ in (27) represents the welfare gain associated with the exposure to period-t risk that also appears in (21), and in which the fourth term $\frac{1}{2}\frac{\lambda^2}{\theta}Y_t$ in (27) represents the welfare gain associated with the ability of generation t to share in the risk from period $t-1$ (cf (20)).

Utility-based fairness criterion

All generations experience the same relative gain from risk sharing in terms of certainty equivalent consumption if:

$$\alpha_t = -\eta \frac{1}{2} \frac{\lambda^2}{\theta} Y_t \tag{28}$$

for all $t \geq 1$, in which case the gain from risk sharing equals (substitute (28) into (27) to eliminate α_t and subsequently subtract (21))

$$CEC_t - \{CEQ_t\}_{\forall s: \tilde{\tau}_s = 0} = \eta \frac{1}{2} \frac{\lambda^2}{\theta} Y_t \tag{29}$$

for all $t \geq 1$.

Market-based fairness criterion

In our benchmark model, the ex-ante market value of the intergenerational transfer $\tilde{\tau}_t$ ($t \geq 1$) is equal to its deterministic component α_t because the loading β_t on the excess return on risky assets has no market value (substitute (21) into (22)):

$$GA_t = \begin{cases} -\alpha_t/(1+r) & \text{for } t = 1 \\ -\alpha_t/(1+r) + \alpha_{t-1} & \text{for } t > 1 \end{cases}. \tag{30}$$

which leads to the zero-sum property in (23). Redistribution in terms of ex-ante market value is ruled out by setting

$$\alpha_t = 0 \tag{31}$$

for all t, in which case $GA_t = 0$ for all t (substitute (31) in (30)). The welfare gain from risk sharing under this criterion is given by (substitute (31) in (27) to eliminate α_t and substract (21))

$$CEC_t - \{CEQ_t\}_{\forall s: \tilde{\tau}_s = 0} = \begin{cases} 0 & \text{for } t = 1 \\ \frac{1}{2} \frac{\lambda^2}{\theta} Y_t & \text{for } t > 1 \end{cases}. \tag{32}$$

Elastic labor supply

Expressions in section 3.4.2 are derived on the basis of a setting in the labor supply of generation $t+1$ (for all $t \geq 1$) is elastic with respect to the size of the transfer $\tilde{\tau}_t$ that is levied upon this generation by the previous generation t. The benchmark model is extended as follows. The labor earnings of generation t are redefined as the product of the wage rate w_t and the labor supply level h_t of generation t, i.e. $Y_t \equiv w_t h_t$. The preferences of generation t, previously given by (15), are now specified over consumption C_t and labor h_t:

$$U_t = \mathbf{E}\left[\frac{1}{1-\theta}\left(C_t - \frac{\epsilon}{\epsilon+1}(h_t)^{\frac{\epsilon+1}{\epsilon}} + \frac{\epsilon}{\epsilon+1}(h_t^*)^{\frac{\epsilon+1}{\epsilon}}\right)^{1-\theta}\right], \tag{33}$$

in which ϵ represents the elasticity of labor supply with respect to the marginal wage rate. Accordingly, a drop in the wage rate by one percent results in a decline in the labor supply level of ϵ percent. Originating from Greenwood (1988), the specification in (33) features no income effects in labor supply. If labor supply is undistorted, labor supply choices are given by:

$$h_t^* \equiv w_t^{\epsilon}. \tag{34}$$

and (33) reduces into (15).

Assume that risk-sharing transfer $\tilde{\tau}_{t-1}$ are levied upon the labor earnings of generation t in the form of *proportional* taxes and subsidies.[22] The marginal tax/subsidy on labor earnings then equals the average tax/subsidy and is given by $\frac{\tilde{\tau}_{t-1}}{w_t h_t}$. As a result, the effective marginal wage rate against which labor is supplied by generation t equals $w_t(1+\frac{\tilde{\tau}_{t-1}}{w_t h_t})$ and it follows from (34) that the labor-supply choice h_t of generation t is given by:

$$h_t = \left(w_t\left(1+\frac{\tilde{\tau}_{t-1}}{w_t h_t}\right)\right)^{\epsilon} = h_t^*\left(1+\frac{\tilde{\tau}_{t-1}}{w_t h_t}\right)^{\epsilon} \approx h_t^*\left(1+\frac{\tilde{\tau}_{t-1}}{w_t h_t^*}\right)^{\epsilon}, \tag{35}$$

where the approximation is applied because the labor supply choice h_t appears on both sides of the first equality of (35) and does not adopt an explicit solution.

[22]Proportional taxes and subsidies are, by approximation, optimal if preferences over consumption obey constant relative risk aversion. Under constant relative risk aversion, it is optimal for agents to share proportionally equally in a shocks. This result holds exactly in absence of distortions, and holds by approximation in the model with elastic labor supply if distortions are not too large.

Under elastic labor supply, the deterministic component α_t matters for the social surplus from risk sharing because it distorts labor supply. We therefore derive expressions for the welfare gain from risk sharing under the condition $\alpha_t = 0$ as in (31). Substitution of (35) into (33) and applying an Arrow-Pratt approximation along the lines of (18) and (23), the expression for certainty equivalent utility now becomes:

$$CEC_t = \begin{cases} (1+r)Y_t^* + (1-\hat{\beta}_t)f_t Y_t^* \bar{\mu} - \frac{1}{2}\frac{\theta}{Y_t^*}(1-\hat{\beta}_t)^2 f_t^2 (Y_t^*)^2\sigma^2 - \alpha_t & \text{for } t = 1 \\ (1+r)Y_t^* + (1-\hat{\beta}_t)f_t Y_t^* \bar{\mu} - \frac{1}{2}\frac{\theta}{Y_t^*}(1-\hat{\beta}_t)^2 f_t^2 (Y_t^*)^2\sigma^2 - \alpha_t & \\ \quad +(1+r)\left\{\hat{\beta}_{t-1}f_{t-1}Y_{t-1}^*\bar{\mu} - \frac{1}{2}\frac{\theta+\epsilon}{\theta}\frac{\theta}{Y_t^*}\hat{\beta}_{t-1}^2 f_{t-1}^2 (Y_{t-1}^*)^2\sigma^2 + \alpha_{t-1}\right\} & \text{for } t > 1 \end{cases}. \tag{36}$$

in which $Y_t^* \equiv w_t h_t^*$. The term $\frac{1}{2}\frac{\theta+\epsilon}{Y_t^*}\hat{\beta}_{t-1}^2 f_{t-1}^2 (Y_{t-1}^*)^2\sigma^2$ now includes not only the welfare loss from volatility in consumption associated with the risk transfer (as measured by θ) but now also includes the welfare loss from from labor-supply distortions associated with the risk transfer (as measured by ϵ). In the special case where labor supply is inelastic, i.e. $\epsilon = 0$, certainty equivalent consumption in (36) reduces into (23).

Substituting (36) into (24) and taking the first-order conditions with respect to $\hat{\beta}_t$ and f_t solves the optimal risk sharing solution under elastic labor supply (cf (25) and (26)):

$$\hat{\beta}_t = \frac{\eta}{1+\eta}\frac{\theta}{\theta+\epsilon}. \tag{37}$$

and

$$f_t = \frac{\lambda}{\theta\sigma}\left(1+\eta\frac{\theta}{\theta+\epsilon}\right), \tag{38}$$

where now $\eta \equiv \frac{1}{Y_t^*}\frac{Y_{t+1}^*}{1+r}$. (37) leads to (18) in the same way as (37) leads to (18).

Substitution of (37) and (38) into (36) yields: (cf (27))

$$CEC_t = \begin{cases} (1+r)Y_t + \frac{1}{2}\frac{\lambda^2}{\theta}Y_t - \alpha_t & \text{for } t = 1 \\ (1+r)Y_t + \frac{1}{2}\frac{\lambda^2}{\theta}Y_t - \alpha_t + \frac{1}{2}\frac{\lambda^2}{\theta}\frac{\theta}{\theta+\epsilon}Y_t + (1+r)\alpha_{t-1} & \text{for } t > 1 \end{cases}. \tag{39}$$

in which the term $\frac{1}{2}\frac{\lambda^2}{\theta}\frac{\theta}{\theta+\epsilon}Y_t$ represents the welfare gain associated with the ability of generation t to share in the risk from period $t-1$ (cf (25)).