

A Multivariate Model of Strategic Asset Allocation with Longevity Risk

Emilio Bisetti, Carlo Favero, Giacomo Nocera and Claudio
Tebaldi

A MULTIVARIATE MODEL OF STRATEGIC ASSET ALLOCATION WITH LONGEVITY RISK[†]

Abstract

Generalized unexpected raise in life expectancy is a source of aggregate risk. Longevity-linked securities are a natural instrument to reallocate these risks by making them tradable in the financial market. This paper extends the Campbell and Viceira (2005) strategic asset allocation model including a longevity-linked investment possibility in addition to equity and fixed income securities. Estimation of the model, based on prices for standardized annuities publicly offered by US insurance companies, shows that aggregate shocks to survival probabilities are predictors for long term returns of the longevity linked securities, and reveals an unexpected predictability pattern. The empirical valuation of the market price of longevity risk confirms that longevity linked securities offer cheap funding opportunities to asset managers willing to leverage their investment portfolio.

JEL Classification: G11, G12 and G22

Keywords: longevity risk and strategic asset allocation

Emilio Bisetti bisetti@cmu.edu
Carnegie Mellon University

Carlo A. Favero carlo.favero@unibocconi.it
Università L. Bocconi, IGIER and CEPR

Giacomo Nocera gnocera@audencia.com
Audencia and CAREFIN

Claudio Tebaldi claudio.tebaldi@unibocconi.it
Bocconi University

[†] We would like to thank B. Hollifield, E. Biffis, S. Paci for valuable comments, the participants to the International Pension Workshop Netspar 2015 Amsterdam and in particular B. Melenberg (the discussant) for useful insights, the usual disclaimer applies.

1. Introduction

Actuarial risk securitization is already an industry practice¹. In particular, given demographic trends and progress in health care, the market for longevity risk trading is knowing a fast development. Insurance business is potentially a source of cheap leverage well known to asset managers. For example Frazzini et al. (2013) estimate that 36% of Berkshire's liabilities consist of insurance float, with an estimated cost which is only 2.2%, more than 3 percentage points below the average T-bill rate. Hence, Buffett's low-cost insurance and reinsurance businesses have given him a significant advantage in terms of unique access to cheap, term leverage. Clearly this reduction in the cost of leverage comes at a risk: in December 2010 Berkshire Hathaway opened a dispute with the reinsurer Swiss Re. Financial press reported that management felt that the losses it had suffered on the book of life reinsurance business were greater than expected. It recorded a pre-tax loss of \$642m in 2011 due to mortality rates exceeding expectations under the quota share type deal.²

Although neglecting or accounting only partially for insurance risk exposure (especially their non-diversifiable component) might result in a wrong or biased representation of the true risk-return profile of a financial portfolio, actuarial risks that may affect liabilities' value and modify optimal allocations are usually not included within the set of stochastic driving factors.

This paper takes a step toward an integrated evaluation of actuarial and market risks. The market risk-return tradeoff is redetermined by including a longevity-linked security in the set of tradable investment opportunities. Empirically, returns on this synthetic security are derived from the prices of publicly available standardized annuities contracts, and are used to quantify the impact of the longevity risk reinsurance business on the performance of optimal portfolio allocations.

Longevity risk is the risk that an annuitant lives more than forecasted by the annuity provider, so that the company has to pay the annuity for a longer-than-expected period

¹for a recent review on the market practice see Blake et al. (2013)

²As of March 28 2013, the reinsurer Swiss Re has announced that its dispute with Warren Buffett's Berkshire Hathaway regarding a 2010 life retrocession deal has been settled.

after her retirement. Longevity risk can be decomposed in two underlying components: an idiosyncratic random variation risk and a common trend risk. Random variation risk is the risk that mortality rates differ from their expected outcome as a result of chance or individual-specific characteristics. Trend risk, on the other hand, is the risk that unanticipated changes in life-style behavior or medical advances significantly improve longevity for the population as a whole. Idiosyncratic risk is dealt with by pooling a large number of different individuals. Trend risk, similarly to any macroeconomic risk, is on the other hand an “aggregate risk” that cannot be diversified away by pooling. Longevity linked securities are instruments designed to reduce the concentration of this aggregate longevity risk on public and private institutions balance sheets offering life insurance services. One path toward this reduction is the creation of a liquid market for these securities, in order to enhance risk sharing among different categories of financial investors and insurance sellers and to produce an efficient valuation of the cost of longevity risk.³ The present paper assumes that individual specific risk is optimally diversified within the insurance sector by risk-pooling and reinsurance business is carried out by financial investors who are available to trade aggregate longevity risk in order to optimize the risk-return tradeoff of their portfolio. In this respect, the present approach is complementary to the analysis of Kojien and Yogo (2013) who focus on the valuation of risks and benefits arising from concentration of shadow, off balance-sheet, reinsurance activities carried out by insurance companies.

The determination of the risk-return term structure is the major innovation of our approach to (aggregate) longevity risk security valuation models, and represents our first contribution to the literature. A risk term structure describes the dependence of the risk-return tradeoff on the investor’s holding period. As longevity shocks are known to be driven by low frequency trends on time scales comparable to business cycles, a proper description of the risk-return tradeoff term structure is therefore crucial to promote an effective integration of longevity risk management in the insurance industry with asset management business in the financial sector. To produce a quantitative assessment of the impact of longevity risks

³An early proposal to design of financial instruments for hedging longevity risk is done in Blake and Burrows (2001).

on the term structure of risk return trade-offs of optimal investment allocations we extend the Vector AutoRegression (VAR) framework originally proposed by Campbell and Viceira (2005) to estimate the term structure of the Markowitz (1952) risk-return efficient frontier generated by US stocks, Treasury bonds and Treasury bills. In their model, Campbell and Viceira (2005) estimate a VAR model including returns on US stocks, Treasury bonds and bills and a set of associated predictors, namely the dividend-price ratio, the spread between long-term and short-term bonds and the nominal T-bill yield.

We extend this framework by assuming that the set of investment opportunities also includes a synthetic financial security indexed to aggregate longevity risk and that the set of predictors includes a mortality predictor. Although the return of the longevity linked security is not observable, it is proved that it can be estimated using the observations of log-price changes for standardized immediate annuities offered and publicly reported by North American insurance companies. The relation between annuity log-price changes, the returns of the tradable synthetic security and aggregate longevity risk predictor is derived applying a novel procedure similar to the Campbell and Shiller (1988) price-dividend log-linearization.

The mortality predictor, on the other hand, is delivered by the Lee and Carter (1992) stochastic mortality model, a common “market standard” used in actuarial science and in the insurance industry to describe the aggregate actuarial uncertainty driving the evolution of annuity prices.

At first sight the use of the Lee Carter mortality modeling approach could appear an arbitrary choice that underestimates more accurate models recently proposed in the literature, but the derivation of the valuation model will clearly show that this statistical framework applied to the analysis of mortality data turns to be the natural extension of the log-normal approximation which is conventionally applied to model prices and predictors within the VAR approach.

The estimation of the extended VAR highlights many interesting features: first, as expected from the theoretical derivation, the mortality predictor appears to predict annuity price changes; second, shocks to aggregate mortality show a mean reverting behavior which implies some degree of predictability.

The second contribution of the paper is a quantitative assessment of the longevity risk compensation. Its computation shows that for short and medium term investment horizons, the optimal (zero investment) longevity risk hedging portfolio is formed by a long position in traded financial assets financed by a short position in the synthetic longevity risk security. Evaluation of the Sharpe ratios for portfolios corresponding to different hedging policies produce reasonable risk compensation estimates. When hedging instruments are restricted to conventional fixed income securities, then risk compensation is in line with those used in the insurance industry practice, when allocations are left unconstrained and hedges may include equity, then large Sharpe ratio estimates confirm that longevity linked securities offer a good source of cheap leverage to risk seeking asset managers. Quite remarkably the optimal portfolios finance equity investment shorting longevity-linked liabilities, thus suggesting strategies that are surprisingly similar to those that Berkshire Hathaway has offered to investors in the past years.

Along this direction, our precise description of the term structure of longevity risk-return trade-offs illustrates that the creation of liquid longevity-linked securities with a stable demand critically hinges on an efficient maturity transformation activity. In fact, only with the support of the latter it is possible to diversify longevity risks among investors with shorter holding periods and averse to liquidity and credit risks of long-duration bonds.

The paper is organized as follows: Section 2 analyzes the contribution in relation to the existing actuarial and financial literature. Section 3 describes the construction of the aggregate longevity risk state variable and the estimation of an extended VAR that includes the aggregate longevity risk shock and annuity price changes. Section 4 describes the optimal allocation for investors who have the opportunity to invest in a synthetic longevity-linked security with short duration and discusses the normative implications for the design of an efficient market for longevity risk transfer. Section 5 defines an hedging portfolio for aggregate longevity risk and quantifies longevity risk compensation as measured by the Sharpe Ratio of this hedging portfolio. Section 6 concludes. We report details about the estimation of the Lee-Carter model and about the derivation of the VAR specification in the Appendix.

2. Related Literature

Long horizon mean-variance allocations share many properties with the strategic asset allocations chosen by intertemporal utility maximizing investors, but they are easier to compute (Campbell and Viceira (2005)). In this respect, our VAR is very close in spirit to the Campbell et al. (2003) strategic asset allocation model but allows to extend the analysis of Cocco and Gomes (2012) to analyze allocation choices for an investor who can choose between a larger set of securities, namely equity and long term bonds.

From a portfolio allocation point of view, we build on Cocco and Gomes (2012), who analyze the portfolio choice problem of an agent investing in financial assets whose returns are correlated with the shocks to survival probabilities and which can be used to buy insurance against aggregate longevity risk. In particular, the authors study both the portfolio allocation between these bonds and risk-free assets and how their demand changes over the life-cycle depending on individual characteristics. Differently from these authors, we assume a point of view closer to the one of a reinsurance (representative) investor assuming that the idiosyncratic, life cycle component is optimally managed within the insurance sector and does not play any role in the analysis. Allocation in longevity-linked securities is freely determined by an unconstrained reinsurance investor that optimizes his allocation conditionally on the risk-return performance of tradable investment opportunities.

In this respect it is important to remark that standardized retail insurance annuity contracts which are used to estimation of the VAR are significantly different from tradable financial securities: in fact they are individual-specific and their purchase is irreversible. The approach pursued in this paper overcomes this problem proving (see Section 3) that annuity log-price changes can be considered as the returns of a tradable synthetic security with a payoff which only depends on aggregate longevity risk - as measured by the (ideally publicly available) index of our construction - and thus offers a stylized example of a longevity-linked security.

The distinction between annuities and this synthetic longevity-linked securities is particularly important as the valuations of financial and actuarial contracts differ in a significant

way, as previously highlighted in the actuarial literature. From an empirical point of view Mitchell et al. (1999), Poterba (2001), Horneff et al. (2009) among others, study the welfare benefits from purchasing annuities and discuss the well-known under-annuitization puzzle. On the theoretical side, on the other hand, the actuarially fair pricing of annuities is a well known, investigated problem (see e.g. Pitacco et al. (2009) and Milevsky (2006) and references therein). In addition to idiosyncratic random variation risk and a common trend risk, Biffis et al. (2010) point out a third source of risk affecting insurance securities: basis risk, i.e. the risk that the population from which the survival probabilities were estimated differs from the insurer's cohort. Our paper analyzes only the aggregate mortality risk component, the one which is not removable through conventional actuarial policy actions (e.g. pooling, screening of subscribers) and rational market agents are accepting to hold in exchange for a reward.

The point of contact between the valuation of an annuity as from the above literature and that of the longevity-linked security we employ in this paper is a “fair pricing” argument of widespread use in actuarial science (Milevsky (2006)): rational agents decide whether to annuitize or to defer the purchase of the annuity for a given period of time by comparing the return offered by the annuity with the potential return from self-annuitization during the deferral period. As the efficiency of self-annuitization depends on the alternative financial investments available in the market, it is reasonable to expect that an annuity contract will offer a return that is both competitive as compared to that of similar financial securities and includes a specific mortality credit component.

In fact, the main goal for our modelling procedure is a precise quantification of the potential benefits for investors and insurance providers generated by the creation of an integrated market for (aggregate) longevity risk sharing. In this context, it is important to remark that a transparent and accessible evaluation of the price of aggregate risk is necessary not only for financial longevity risk sharing schemes but also for actuarial ones.

An approach similar to ours for the pricing of longevity risk is followed by Lin and Cox (2005) and Lin and Cox (2008), who apply the 1-factor and 2-factor Wang transform to estimate longevity premia from annuity prices. While we retain their main actuarial valuation

principles, we avoid the use of the Wang transform which, as pointed out by Pelsser (2008), “does not lead to a price which is consistent with the arbitrage-free price” and, therefore, “cannot be a universal framework for pricing financial and insurance risks” (Wang (2002)). Milevsky et al. (2006) and Bayraktar et al. (2009) develop a theory for pricing undiversifiable mortality risk in an incomplete market. They postulate that an issuer of a life contingency requires compensation for this risk according to a pre-specified instantaneous Sharpe ratio (see also Bayraktar and Young (2008)). Within our model, the incompleteness generated by demographic uncertainty is accounted for by including an additional state variable which is extracted from the Lee and Carter (1992) model for stochastic mortality. Previous attempts to quantify the impact of longevity risk on market prices, like Friedberg and Webb (2007), who apply the Capital Asset Pricing Model (CAPM) and the Consumption Capital Asset Pricing Model (CCAPM) to quantify risk premia for potential investors in longevity bonds, produce very low estimates of such a premium. The authors acknowledge that there is likely to exist a “mortality premium puzzle” similar to the well-known “equity premium puzzle” (Mehra and Prescott (1985)) driving higher mortality risk premia in the data than those economic models would suggest.

3. Risk and returns in a VAR model for financial securities and annuity prices

Our empirical strategy follows the approach to the optimal portfolio choice problem under return predictability proposed by Campbell et al. (2003) and Campbell and Viceira (2005). However, we extend the traditional investment opportunity set made by stocks, bonds and T-bills with a longevity-linked security and derive the optimal portfolio allocation at different horizons based on a Vector Autoregressive (VAR) specification for returns and their predictors. As we include a longevity-linked security in the investment opportunity set, we build an appropriate associated predictor from the estimation of a stochastic mortality model. We therefore first illustrate how annuity valuation implies that the unexpected generalized mortality innovation as from the popular Lee and Carter (1992) model can be used as a predictor of the return of a longevity-linked security. We then propose a VAR model of the joint dynamics of the returns on stocks, bills, bonds, longevity-linked securities

and their predictors. The VAR specification is intended to be a reduced form model of the optimal dynamic allocation choice as set by a rational agent. This reduced form model can be derived as an approximation of the solution of a standard utility maximization problem extending the arguments discussed in Campbell et al. (2003) along the lines suggested by Cocco and Gomes (2012). In the context of the VAR dynamics, we build the optimal intertemporal hedging portfolio and address the issue of pricing longevity risk.

3.1. A reduced form model for annuity valuation

Our estimate of the longevity risk-return trade-off is based on the historical time series of observed prices for standardized annuity contracts offered by insurance companies to voluntary individual annuitants. By “standardized” annuities we mean single premium (involving a one-time investment) immediate (commencing regular income payments one period after the premium has been paid) single life (guaranteeing to make payments only to a single beneficiary until her death) fixed (providing fixed payments) annuities.

It is important to remark that annuity contracts significantly differ from tradable financial contracts, as they are individual-specific and their purchase is irreversible. Moreover, insurance companies cannot liquidate the subscribers and annuities cannot be replicated or sold short.⁴ Finally, informational asymmetry between the subscriber and the insurance company is known to affect annuity pricing: it is documented that voluntary subscribers of life annuities live longer than average population.⁵

Despite these differences, it is possible to set bounds on the returns from annuities using the information on the returns offered from alternative financial investment opportunities using a simple argument. Assume that a rational agent of age x at time t faces the alternative between immediate annuitization at price $P_{x,t}$ or a deferral of the entry decision by one year,

⁴An important element distinguishing insurance companies from several other financial intermediaries is the lack of a secondary market where the contracts written by insurance companies may be traded. The holder cannot sell her insurance policy to a third final investor, albeit, in recent years, secondary markets for some insurance contracts have developed.

⁵The adverse selection problem in annuity pricing has been discussed, among others, by Mitchell et al. (1999) and Finkelstein and Poterba (2004). More recently the role of informational asymmetries in longevity markets has been analyzed in Biffis and Blake (2010)

buying the annuity at time $t + 1$ and age $x + 1$ at a price $P_{x+1,t+1}$. In order for the agent to opt for immediate annuitization, the expected return provided by the annuity must be at least as large as the one she would expect from a financial investment with a similar risk-return profile *plus a mortality credit*, the extra return required by the annuitant as a compensation for the exposure to the (actuarial) risk of a premature death between x and age $x + 1$. If this is not the case, the agent would prefer self-annuitization, i.e. deferral of the annuitization anticipated by a short term financial investment of the pension savings. To precisely quantify such risk, let $q_{x,t}$ denote the mortality rate for individuals of age x in year t , i.e. the probability that a person aged x and alive at the beginning of year t dies within the end of the year. We define by $s_{\bar{x}+i,t+i}$ the probability to be alive in year $t + i$, of age $\bar{x} + i$, conditional on being alive at age \bar{x} , so that ⁶

$$\begin{aligned} s_{\bar{x},t} &= 1, \\ s_{\bar{x}+i,t+i} &= s_{\bar{x}+i-1,t+i}[1 - q_{\bar{x}+i,t+i}] \quad i = 1, \dots, \infty. \end{aligned}$$

Life expectancy for a person aged x at time t is defined as $e_{x,t} = \sum_{i=1}^{\infty} s_{x+i,t+i}$. Survival probabilities tend to zero as time increases, given that mortality rates $q_{x,t}$ increase with age x , and the probability of a premature death between x and age $x + 1$ is then quantified by $1 - s_{x+1,t+1}/s_{x,t}$.

Assuming that annuities are offered to rational agents in a competitive market, prices set by insurance companies should correspond to the lowest return making the investor indifferent between immediate annuitization or deferral. Given this premise and the previous definitions, we derive an approximate accounting identity providing an explicit expression of these contributions.

As a starting point consider the definition of the one-period holding return for an annuity paying a coupon C in each period (year) to a person aged x at time t as follows:

⁶The common actuarial notation for the survival probability $s_{\bar{x}+i,t+i}$ would be ${}_i p_{\bar{x},t}$. It is modified in order to keep using the common financial convention where p indicates the logarithmic price of a risky security, e.g. $p - d$ will indicate the price-dividend ratio.

$$R_{t+1}^A = \frac{(P_{x+1,t+1} + C)s_{x+1,t+1}}{P_{x,t}s_{x,t}} - 1. \quad (1)$$

Dividing both sides of (1) by $(1 + R_{t,t+1}^A)$ and multiplying both sides by $\frac{P_{x,t}}{C}$, we have:

$$\frac{P_{x,t}}{C} = \frac{1}{(1 + R_{t,t+1}^A)} \left(\frac{s_{x+1,t+1}}{s_{x,t}} \right) \left(1 + \frac{P_{x+1,t+1}}{C} \right).$$

Denoting with lowercase letters the natural logarithms of uppercase letters we have:

$$p_{x,t} - c = -r_{t+1}^A + \ln \left(\frac{s_{x+1,t+1}}{s_{x,t}} \right) + \ln \left(1 + e^{p_{x+1,t+1} - c} \right).$$

Finally, taking a Taylor expansion of the last term about the average log price-coupon ratio, $\frac{P}{C} = e^{\overline{p-c}}$, we have

$$\begin{aligned} p_{x,t} - c &\simeq -r_{t+1}^A + \ln \left(\frac{s_{x+1,t+1}}{s_{x,t}} \right) + \ln \left(1 + \frac{P}{C} \right) + \\ &\quad + \frac{P/C}{1 + P/C} (p_{x+1,t+1} - c - \overline{p-c}) \\ &= -r_{t+1}^A + \ln \left(\frac{s_{x+1,t+1}}{s_{x,t}} \right) + k^s + \rho (p_{x+1,t+1} - c), \end{aligned}$$

where $\rho \equiv \frac{e^{\overline{p-c}}}{1 + e^{\overline{p-c}}}$. Therefore annuity returns can be written as

$$r_{t+1}^A = k + \rho (p_{x+1,t+1}) - p_{x,t} + (1 - \rho) c + \ln \left(\frac{s_{x+1,t+1}}{s_{x,t}} \right),$$

and recalling that

$$\ln \left(\frac{s_{x+1,t+1}}{s_{x,t}} \right) = \ln (1 - q_{x,t}),$$

we get by rearranging that

$$p_{x,t} = k + (1 - \rho) c + \rho (p_{x+1,t+1}) - r_{t+1}^A + \ln (1 - q_{x,t}).$$

Taking the term-by-term difference of the valuation equation between time $t + 1$ and t , the value of the coupon (nominal payment is fixed) disappears from the valuation equation, so that

$$\Delta p_{x,t} = \rho(p_{x+1,t+1} - p_{x+1,t}) - \Delta r_{t+1}^A + \Delta \ln(1 - q_{x,t}), \quad (2)$$

with $\Delta p_{x,t} \equiv p_{x,t} - p_{x,t-1}$.

Consider now the popular Lee and Carter (1992) model for stochastic mortality. This model has both strong within-sample fitting properties and remarkable out-of-sample predictive power. Together with the relative ease of its computation, these characteristics have made it the standard mortality forecasting model among practitioners and academics. The model consists of a system of equations for the logarithm of the central death rate $m(x, t)$ of each age cohort x at time t , and a time-series equation for an unobservable time-varying mortality index k_t , common among all age cohorts. In particular, in terms of the mortality rate $q_{x,t}$ we have⁷

$$\ln(q_{x,t}) \simeq -(a_x + b_x k_t + \epsilon_{x,t}), \quad (3)$$

$$k_t = c_0 + c_1 k_{t-1} + e_t, \quad (4)$$

$$\epsilon_{x,t} \sim NID(0, \sigma_\epsilon^2),$$

$$e_t \sim \text{ZeroMean} - \text{Stationary},$$

where a_x and b_x are age-specific constants. The error term $\epsilon_{x,t}$ captures cross-sectional errors in the model-based prediction of mortality for different cohorts, while the error term e_t captures random fluctuations in the time series of the common factor k_t driving mortality at all ages; e_t is assumed to be a zero mean stationary process. This common factor k_t evolves over time as an auto-regressive process and the favorite Lee-Carter specification

⁷In the original formulation Lee and Carter assume an exponential affine expression of the central death rates, while in our framework the exponential affine expression is applied directly to the mortality rates. The two coincide within the log-linear approximation of the mortality rates evolution that is considered in the VAR specification. Unreported estimation tests do show that results are essentially unaffected if the original specification is adopted.

makes it a unit-root process by setting $c_1 = 1$. Identification is achieved by imposing the restrictions $\sum_t k_t = 0$ and $\sum_x b_x = 1$, so that the unobserved mortality index k_t is estimated through Singular Value Decomposition.⁸

Under the Lee-Carter specification we have that the revision of the mortality credit contribution is linear in the innovation to the unobserved common factor component k_t . In fact,

$$\begin{aligned} 1 - q_{x,t} &= \exp - [a_x + b_x k_t + \varepsilon_{x,t}], \\ k_t &= c_0 + c_1 k_{t-1} + e_t, \end{aligned}$$

so that for small variations in mortality rates

$$\begin{aligned} \ln(1 - q_{x,t}) - \ln(1 - q_{x,t-1}) \\ \simeq - [b_x e_t + \varepsilon_{x,t} - \varepsilon_{x,t-1}]. \end{aligned}$$

A similar measure has already been introduced and discussed in Friedberg and Webb (2007). Assuming that agents will compensate only aggregate risk, the priced contribution to mortality credit is given by e_t , and we can re-write (2) as

$$\Delta p_{x,t} = \rho (\Delta p_{x+1,t+1}) - \Delta r_{t+1}^A + b_x (-e_t). \quad (5)$$

Solving this relation forward up to period $t + m$ and taking expectations given the information set available at time t , we have:

$$\Delta p_{x,t} \simeq - \sum_{j=0}^m \rho^j E_t [\Delta r_{t+1+j}^A] + \sum_{j=0}^m \rho^j b_{x+j} E_t [-e_{t+j}],$$

which shows that the annuity log price change is determined by future expectations on changes Δr_{t+k}^A in the annuitant holding period returns and on revisions in trend mortality factor e_{t+k} .

⁸See Appendix A for a full description of the adopted identification and estimation strategy.

It turns useful at this point to observe that the annuity log price change can be interpreted as the return on a longevity-linked security whose payoff is indexed to the log price variation $\Delta p_{x,t}$ of annuity prices sold each year t to annuitants in the cohort x . By convention we define Annuity Linked Security (ALS hereafter) a reference synthetic security indexed on annuity prices based on the cohort of 65-years-old US annuitants. A long position in the ALS corresponds to an (nontraded) tontine insurance in which contracts are terminated and then possibly renegotiated every year (for an actuarial discussion of such synthetic contracts see Milevsky (2006), p. 224). A short position in the ALS allows the investor to sell protection on the longevity risk of the cohort of 65-years-old US annuitants. Then the last equation identifies a relation between the return on ALS, the innovations in the common mortality factor in the Lee-Carter model and the annuitant holding period returns.

3.2. A model for stochastic mortality and its performance on the US data

We apply the Lee and Carter (1992) model to estimate shocks to mortality for cohorts in the age interval between 65 and 110. We restrict the estimation to the cohorts of retired population for several reasons. First, the active rebalancing of the contributions is not feasible for these cohorts, as they typically consist of people who left the accumulation phase and entered the decumulation phase. Hence, reallocation via securitization or reinsurance is the only viable strategy that insurance companies can pursue to hedge the associated longevity risk. Second, the largest publicly available empirical data sets on annuity prices apply to annuitants belonging to these cohorts. Third, limiting the specification to retired cohorts alleviates some well-known limitations of the Lee-Carter model when applied to the entire population (see Lee (2000)). On the other hand, the approach we propose is fully flexible and we do not see specific obstacles to extend it to any other (log-linearized) version of stochastic mortality models.

In Figure 1 we report evidence on the performance of the Lee-Carter model in fitting US mortality rates. Our data come from the Human Mortality Database of the University of Berkeley.⁹ In Figure 1.1 we plot realized mortality at age 65 throughout the period

⁹The data are publicly available at <http://www.mortality.org/>

1952-2010 (red dashed line) against its Lee-Carter fitted value (blue continuous line). The model is estimated on cohorts aged 65 to 110. In Figure 1.2, we report the cross-sectional R^2 of the estimate for all age cohorts in the same period. The model performs very well in fitting mortality rates at all ages but those greater than 95, where the volatility of mortality is high: for more than fifty percent of ages, the R^2 is above 95%, and for more than seventy-five percent of ages it is above 80%. Figure 1.3 reports the estimated unobservable common mortality index k from Equation (4), which clearly features a negative trend. The autoregressive coefficient c_1 from Equation (4) is equal to 0.977 with an associated standard deviation of 0.015, and this persistence allows to make strong predictions about future mortality. Figure 1.4 reports the innovations in the unobservable mortality index, which has a persistent autoregressive structure: this variable will be our measure of the relevant uncertainty on mortality and the predictor of returns of longevity-linked securities.

Insert Figure D.1 about here

To derive an observable counterpart of the priced contribution to mortality credit we consider an estimation of the Lee-Carter model restricted to the post-retirement cohorts of individuals aged between 65 and 110. This measure, called qk_t , is very similar to the one discussed in Friedberg and Webb (2007), and coincides with the unexpected variation in the survival rate pooled over all retired cohorts:¹⁰

$$qk_t \simeq \frac{1}{45} \sum_{x=65}^{110} \{ \ln(1 - q_{x,t}) - E_{t-1} [\ln(1 - q_{x,t})] \}.$$

¹⁰We estimate the model on retired cohorts, only.

Using the Lee Carter specification, the observable index qk_{t+1} can be approximated by:

$$\begin{aligned}
qk_t &\simeq -\frac{1}{45} \sum_{x=65}^{110} (\alpha_x + b_x k_t + \varepsilon_{x,t}) - (\alpha_x + b_x E_{t-1} [k_t]) \\
&= -\frac{1}{45} \sum_{x=65}^{110} (\alpha_x + b_x (k_{t-1} + e_t) + \varepsilon_{x,t}) - (\alpha_x + b_x k_{t-1}) \\
&= -\frac{1}{45} \sum_{x=65}^{110} (b_x e_t + \varepsilon_{x,t})
\end{aligned}$$

and, taking into account the normalization condition $\sum_{x=65}^{110} b_x = 1$ we have

$$qk_t \simeq -\frac{1}{45} \sum_{x=65}^{110} (b_x e_t + \varepsilon_{x,t}) = -\frac{e_t}{45} - \frac{1}{45} \sum_{x=65}^{110} \varepsilon_{x,t}. \quad (6)$$

The longevity shock qk_t describes the time evolution of the unexpected variation in mortality rates which has a uniform impact across cohorts, and is estimated by applying the Lee and Carter model only to retired cohorts. Notice that this shock includes two contributions, the first contribution is proportional to the opposite of the Lee Carter aggregate mortality shock e_t , the second one is the sum of the cohort specific innovations. The filtered innovation qk_t is included as a predictor in the vector of autoregressive variables to account for unexpected trend variation in mortality rates. This variable offers a publicly available, cohort-independent, information which investors can observe and use to quantify the impact on prices of the variability of aggregate longevity.

Notice that the second contribution accounts for longevity fluctuations that are non-uniform across cohorts, e.g. a higher reduction of mortality on younger cohorts followed by an increase in mortality at higher ages. It is interesting to analyze the dynamic properties of the second contribution. By construction it has zero expectation but, as originally highlighted by Lee and Carter, $\varepsilon_{x,t}$ are generally correlated across cohorts and across time. On the other hand we retain only the testable assumption that this second contribution is stationary or equivalently, to the assumption that the only non-stationary component is common across cohorts and is accounted by the evolution of k_t .

3.3. A reduced VAR dynamic model for the Annuity Linked Security returns

In this subsection we show that, using the above reduced valuation approach, it is possible to model the stochastic evolution of ALS returns (log annuity price changes) $\Delta p_{x,t}$ using a VAR specification which extends that of Campbell and Viceira (2005) (hereinafter CV-VAR(1)). Following Barberis (2000), Campbell and Viceira (2002, 2005), we describe dynamics of asset returns and relevant predictors using a VAR(1) model:

$$\mathbf{z}_t^{Mkt} = \Phi_0^{Mkt} + \Phi_1^{Mkt} \mathbf{z}_{t-1}^{Mkt} + \nu_t^{Mkt} \quad (7)$$

where

$$\mathbf{z}_t^{Mkt} = \begin{bmatrix} r_{0t} \\ \mathbf{x}_t^{Mkt} \\ \mathbf{s}_t^{Mkt} \end{bmatrix}$$

is a $m \times 1$ vector, with r_{0t} being the log real return on the asset used as a benchmark to compute excess returns on all other asset classes, \mathbf{x}_t being the $n \times 1$ vector of log excess returns on all other asset classes with respect to the benchmark, and \mathbf{s}_t is the $(m - n - 1) \times 1$ vector of returns predictors. The exact specification and its estimation results are reviewed in Appendix B.

As described in the previous subsections, although an annuity is not a financial security and cannot be priced accordingly, rationality of the annuitant forces the (log) holding period return r_t^A to be comparable (but for the mortality credit) to the compensation one would get by investing in a portfolio of traded financial securities with similar risk and return characteristics while deferring by one year the annuitization. Hence we claim, and later show empirically, that the financial component of the return r_t^A can be replicated using a portfolio of securities whose evolution is described by the CV-VAR(1) model. Moreover, assuming a stationary evolution for r_t^A , the VAR(1) specification implies that also $-\sum_{j=0}^m \rho^j E_t \Delta r_{t+1+j}^A$ is a linear function of the state variables \mathbf{z}_t^{Mkt} . This is equivalent to assuming:

$$-\sum_{j=0}^m \rho^j E_t \Delta r_{t+1+j}^A = \phi_0^A + \phi_1^{A,Mkt} \mathbf{z}_t^{Mkt}.$$

Similarly, the mortality credit component affecting the evolution of ALS price is proxied by the observed innovations in the longevity trend. For this reason an additional predictor state variable qk_t is introduced and it is assumed that

$$\sum_{j=0}^m \rho^j b_{x+j} E_t [-e_{t+j}] = \phi_3 qk_t.$$

In conclusion, the standard CV-VAR(1) model can be augmented to include the evolution of the (65 cohort) annuities' (log) price growth in excess to the return of the safe asset, $x\Delta p_t := \Delta p_{65,t} - r_{f,t}$, following the specification:

$$x\Delta p_{t+1} = \phi_0^A + \phi_1^{A,Mkt} \mathbf{z}_t^{Mkt} + \phi_2 x\Delta p_t + \phi_3 qk_t + \nu_{t+1}^A,$$

where ν_{t+1}^A is the combination of all shocks in the state variables and idiosyncratic mortality shocks. Excess returns of the ALS are determined by a combination of market returns, market return predictors and the aggregate longevity predictor. We now analyze the effect of extending the traditional portfolio to include excess annuity prices by considering the following augmented VAR specification:

$$\begin{aligned} \mathbf{z}_t &= \Phi_0 + \Phi_1 \mathbf{z}_{t-1} + \nu_t \\ \nu_t &\sim \mathcal{N}(0, \Sigma_\nu), \end{aligned} \tag{8}$$

where

$$\mathbf{z}_t = \begin{bmatrix} r_{0t} \\ \mathbf{x}_t^{Mkt} \\ x\Delta p_t \\ \mathbf{s}_t^{Mkt} \\ qk_t \end{bmatrix}$$

and Σ_ν is the $(m+2) \times (m+2)$ variance-covariance matrix of the returns on financial assets, the annuity prices and their associated predictors.

3.4. The dynamics of returns of US bonds, bills, stocks and annuities

To evaluate how the inclusion of the ALS excess returns and of a predictor for the change in their prices modifies the optimal portfolio allocation at different horizons, we compare the results obtained from the CV-VAR(1) estimation over the yearly sample 1953-2010 (the most recently update of the human mortality database) to those obtained from our extended VAR. The first model includes six variables: the ex-post real T-bill rate, the annual excess returns on stocks, the annual excess returns on long-term (20-year) bonds, the log yield on a 90-day T-bill, the log dividend-price ratio and the yield spread (defined as the difference between the 20-year zero-coupon bond yield from the CRSP Fama-Bliss data file - the longest maturity yield available in the file - and the T-bill rate). The second model is an eight-variables VAR obtained by adding to the standard CV-VAR(1) the log difference in the annuity premium minus the risk-free rate (which extends the set of excess returns) and the aggregate mortality shock (which extends the set of predictors). Table D.1 shows sample statistics for all variables. Our sample which includes observations up to the most recent update of mortality data, compares well with the annual sample used in previous studies. Only the statistics on long term bond indicate a lower expected return, a result which is clearly driven by the recent trends in monetary policy.

Table D.2 and Table D.3 show the results for the original and the extended VAR including longevity predictor and ALS returns. Both estimated VAR's include constants in each equation. Detailed information on the data used in the estimation are reported in Appendix C. The results from the standard model are well in line with those reported in Campbell et al. (2003). When the extended VAR is estimated, no major changes take place in the coefficients attached to the six financial variables in the original model. For those variables whose explanatory power is significantly different from zero, the impulse response coefficients are qualitatively similar between the two tables and confirm all the stylized properties found in the original estimation: real T-bill, stock and bond returns are predicted by nominal short rate, dividend-price and term spread. The longevity shock is persistent and helps to predict the change in annuity prices (a positive shock to longevity increases the price of annuities). It also has some significance in predicting excess returns on bonds. Although it

is not possible to identify the structural origin of this predictive relation within our reduced partial equilibrium model, it is in line with some macroeconomic arguments supporting a positive relation between rise in longevity and price of long duration safe assets. The two new equations included in the extended VAR describe the evolution of the aggregate longevity shock and the logarithmic yearly change of the annuity price in excess to the nominal T-bill. The estimated aggregate longevity shock dynamics qk_t is substantially a univariate mean reverting process with a root of 0.75 and zero expectation (estimation provides a value $E_{t-1}[qk_t] = 0.0005$) confirming both that the information conveyed by the aggregate longevity shock is not spanned by other variables and that the spread between expected and realized aggregate longevity is mean reverting and thus may be used to forecast aggregate longevity growth.

Notice that the persistence of this shock does not contradict the standard Lee Carter specification, in fact the usual specification assumed for the trend component k_t is a unit root evolution process. This requires only the first difference of the process to be a stationary process, which is consistent with our findings. In addition, as shown in Figure **D.10**, a closer analysis of the contributions to qk_t highlights that mean reversion is almost entirely driven by the second contribution to qk_t in eq.(6), i.e. by the sum of cohort specific shocks. These fluctuations measure the time t deviation from zero of the cohort-specific shocks sample mean and the longer the persistence of this process, more slowly they return to the mean. Notice that from a demographic point of view these fluctuations measure also the deviations from the original expected distribution of the cohort specific mortality rates which, especially at old ages, are driven by health and life quality improvements. As it should be expected, we verify in sub-samples that this mean reversion is slowly increasing as the time span of the sample increases.

Insert Figure D.10 about here

The aggregate longevity shock is a significant predictor for $x\Delta p_{t+1}$ (the ALS return in excess to the T-Bill rate), which is also significantly predicted by past real and nominal T-bill

rates and the excess returns on long-term bonds. Annuity price growth in excess to the T-bill rate have a positive loading on the real rate and a negative one on the nominal T-bill rate, a negative loading on long term excess bond rate returns and a positive dependence on the aggregate longevity shock. In Figure D.2 the time series of historical (real) logarithmic price changes is compared to the replication as operated by the VAR dynamic model aggregating the information of financial securities returns and of the forecasting variables, including the aggregate longevity shock.

Insert Figure D.2 about here

The good fit indicates that the VAR estimation produces a realistic “reduced form” pricing model for the annuity contract offered by insurance companies to annuitants. We underline once more that this approach is tailored to capture the main stylized features of the interaction between financial markets and annuity risk management in the insurance industry. This approach accounts only for aggregate longevity risk and does not account for the actuarial components of insurance premia which must be included to hedge basis risks or adverse selection effects, which in turn require a discussion of the specific characteristics of annuitants.

4. The impact of longevity securitization on optimal allocations

The possibility to trade longevity-linked securities extends the set of investment opportunities and offers a new diversification dimension. Following Campbell and Viceira (2005), we compute the optimal portfolios for investors adopting a buy-and-hold strategy with holding period between 1 and 40 years. The set of investment opportunities is composed by T-bills, equity, a rolling strategy in a long term bond and the ALS. This security grants to its holder a yearly return equal to $\Delta pr_{65,t}$ the variation of the mean (logarithmic) price observed on the US insurance market for a standardized annuity contract. The estimation indicates that, as expected, the return from this contract will raise if aggregate longevity is raising and will decrease if aggregate mortality increases.

A term structure of conditional volatilities at different horizons can be naturally derived from the estimation of our VAR process for returns and predictors.

Insert Figure D.3 about here

In Figure D.3, we compare the term structures of the standard deviation of the ALS and of traditional financial securities for an horizon ranging from 1 to 40 years. Notice that the annualized volatility increases with the holding period, the clear sign of the long-term nature of the risks underlying annuity prices. For holding periods shorter than 10 years, ALS is less risky than equity and (rolling) bond investments, while on longer horizons its risk exceeds that of other securities: the irreversible nature of annuitization implies that, from a pure financial point of view, this contract has a risk-return profile similar to the one by a buy-and-hold strategy on a long term bond. ALS price fluctuations reflect changes in the long-run expectations about inflation and longevity trends, and a small persistent change to future expectations can have a relevant impact on the current evaluation of the annuity contract, making prices fluctuate accordingly. Note that an ALS is not the contract that an insurance company would like to use to reinsure aggregate longevity risk. In fact, an insurance company with a portfolio of annuities under management would be willing to reinsure only aggregate longevity risk but would prefer to retain the remunerative core business of the insurance industry the diversification of the residual risk by pooling together annuitants of all ages and different lifestyles.

Figures D.4, D.5 plot the correlations among the extended set of securities as a function of the holding period.

Insert Figure D.4, D.5 about here

The correlations between the ALS and financial securities have a sharp decline with the holding period. These correlations, and more in general the term structure of assets' risk and returns, determine the weight that each asset receives in the portfolio allocation of an investor with mean-variance preferences for any given horizon. To have a sense of how this portfolio allocation changes by including the possibility to invest in the ALS we again follow

Campbell and Viceira (2005) and first consider the generalized absolute minimum variance portfolio (henceforth GMV), the portfolio with the lowest variance on the mean-variance efficient frontier. For each holding period this portfolio is described in Figure D.6.

Insert Figure D.6 about here

Figure D.6 shows that the investor overweights the allocation in the T-bill to buy a combination of ALS and long term bond independently of the investment horizon. This combination is a long position in the bond and a short one in the ALS when the holding period is below 10 years, while for longer horizons the two positions are switched. Hence, over periods of time smaller than a decade, risk exposure is minimized by selling protection to longevity while the same investment becomes speculative over longer holding periods. As a consequence, it is expected that demand for longevity exposure and the liquidity of longevity-linked securities can be considerably increased by offering products with short durations.

Figure D.7 compares the term structure of risk of the extended GMV with that of the “Campbell-Viceira” GMV and that of a T-bill. Inclusion of the ALS reduces risk for all holding periods with the exception of the interval between 10 and 15 years, where the optimal allocation in the ALS shifts from negative to positive values and the rolling position in long term bonds from positive to negative.

Insert Figure D.7 about here

Roughly speaking, the allocation and the risk profile of the GMV confirm that a long position in aggregate longevity is financially appealing, with low risk and good diversification properties, only for an investor with an horizon longer than 12-13 years. This result is consistent with the hypothesis that annuity prices offer a return which is competitive with the alternative investments available to the investor: annuities for a 65 year old investor have an effective duration around 12 years (see, for example, Loeys et al. (2007)). Note that the 12-year minimum variance portfolio corresponds to an allocation in the ALS which is essentially zero.

Based on these observations it is possible to conclude that the creation of short-duration (less than 10 years) longevity-linked securities is the key step for an efficient securitization of longevity risk. These securities offer a stochastic liability which can be efficiently used to finance investments with good diversification properties. The risk-return analysis of the ALS shows that securitization of longevity, which is a long run risk in the sense of carrying a small but persistent component, does not necessarily require the use of long-duration securities. On the contrary, upon a precise quantification of the term structure of longevity risk exposures, a more efficient management of maturity transformation can be realized using structured securities, like for example swaps. These findings indicate that the problems which affected early longevity-indexed security issuances were determined by long durations giving rise to liquidity and credit risk components so large as to overwhelm the effect of longevity risk both for pricing and hedging.

As a second illustration of the optimal mean-variance allocations including a position in the short-term ALS, in Figure D.8 we plot the optimal allocations for a portfolio with an expected return of 10% as a function of the holding period returns. As expected, the ALS short position is used to leverage a portfolio of T-bills, equity and long-term bonds.

Insert Figure D.8 about here

5. Longevity securitization and inter-temporal hedging of the aggregate longevity risk

The results from the estimation of the extended VAR provide evidence of a significant response of annuity prices to variations in aggregate longevity rates. Quoted annuity prices are expected to include a compensation for the insurance company to bear a risk exposure for the unexpected raise of the undiversifiable longevity risk component. Following the conventional Intertemporal CAPM (ICAPM) interpretation (Merton (1973)) expected utility maximizers try to hedge the stochastic changes of their investment opportunity created by unexpected aggregate longevity shocks. The hedging portfolio is determined by an allocation in traded securities whose return is maximally correlated with the longevity shock qk_t . This

portfolio is determined by the constrained minimization problem:

$$\begin{aligned}
& \min_{\mathbf{w}} \text{Var}_{t-1} \left[R_t^{qk}(\mathbf{w}_{t-1}) - qk_t \right] \\
s.t. \quad & R_t^{qk}(\mathbf{w}_{t-1}) = \mathbf{w}_{t-1} \cdot \mathbf{x}_t + W_{0,t-1} rtb_t \\
\mathbf{w}_{t-1} &= [w_{xr,t-1}, w_{xb,t-1}, w_{x\Delta p,t-1}], \\
\mathbf{x}_t &= [xr_t, xb_t, x\Delta p_t],
\end{aligned}$$

where \mathbf{x}_t includes the log excess returns of market securities plus the annuity log price growth in excess to the T-bill rate, R_t^{qk} is the return on the replication portfolio and $W_{0,t}$ is the investor's wealth at time t . We assume that the trading strategy is constrained by the condition that, in each period t ,

$$w_{xrtb,t} + w_{xr,t} + w_{xb,t} + w_{x\Delta p,t} = W_{0,t},$$

and that the real return of the portfolio at each period is given by

$$\begin{aligned}
R_t(\mathbf{w}_{t-1}, w_{rtb,t-1}) &= w_{xr,t-1} \cdot (xr_t + rtb_t) + w_{xb,t-1} \cdot (xb_t + rtb_t) \\
&\quad + w_{x\Delta p,t-1} \cdot (x\Delta p + rtb_t) + (W_{0,t} - \mathbf{w}_{t-1} \cdot \mathbf{1}) rtb_t \\
W_{0,t} &= w_{rtb,t-1} + \mathbf{w}_{t-1} \cdot \mathbf{1}.
\end{aligned}$$

Recall that the purchase of an annuity is irreversible and payments are done until the death of a single beneficiary, while $x\Delta p_t$ is the ALS return, the annuity price variation between time $t-1$ and t for the 65 year old male cohort in excess to the T-bill rate. Hence the Aggregate Longevity Hedging Portfolio (hereinafter ALHP) is not tradable unless an Annuity Linked Security paying off the return $\Delta p_{65,t}$ on a yearly basis is made available to investors. This is a benchmark example of the theoretical motivations underpinning the necessity of longevity securitization.

By construction, the ALHP tracks the aggregate longevity shock qk_t and is therefore the best available product to reinsure aggregate longevity risk. Note that the efficiency of

the replication increases with the number of investment opportunities exposed to aggregate longevity risk. In practice, the hedging portfolio is determined by performing the minimization over the set of unconstrained allocations $\mathbf{w}_t = [w_{xr,t}, w_{xb,t}, w_{x\Delta p,t}]$, while the position in the short rate is set equal to $w_{rtb,t} = W_0 - \mathbf{w}_t \cdot \mathbf{1}$. Since our VAR model generates a stationary dynamics, the minimum variance replication portfolio corresponds to a time-independent allocation $\mathbf{w}_{t-1}^* = \mathbf{w}$. The first order condition¹¹ therefore implies the solution

$$\begin{aligned}\mathbf{w}^T &= Var[\mathbf{x}_t]^{-1} \{Cov[\mathbf{x}_t, qk_t]\}, \\ w_{rtb} &= W_0 - \mathbf{w} \cdot \mathbf{1}\end{aligned}$$

We consider three alternative replication portfolios corresponding to zero initial investment ($W_0 = 0$) with an increasing set of restrictions on the allocations. Table D.4 reports three hedging portfolios. The first portfolio considers an unrestricted allocation, while the second an allocation where investment in equity is not allowed ($w_{xr} = 0$). In both cases the allocation strategy is a short position in T-bill and equity and a long position in long term bond and in the ALS.

The third hedging portfolio is further restricted by forcing a zero allocation in the T-bill, $w_{rtb} = 0$, thus making the long-term bond the only available financial security available to finance the annuity liability. In all the three cases the volatility induced by the aggregate longevity shock as measured by the volatility of the aggregate longevity replication portfolio is close to 60 basis points in annual terms, and this value remains almost constant for any holding period with an essentially flat term structure of volatility. This value is slightly

11

$$\begin{aligned}& \min_{\mathbf{w}} Var_{t-1} [R_t^{qk}(\mathbf{w}) - qk_t] \\ &= \min_{\mathbf{w}} Var_{t-1} [qk_t] + Var_{t-1} [R_t^{qk}(\mathbf{w})] - 2Cov_{t-1} [qk_t, R_t^{qk}(\mathbf{w}_{t-1})] \\ &= \min_{\mathbf{w}} Var_{t-1} [R_t^{qk}(\mathbf{w})] - 2Cov_{t-1} [qk_t, R_t^{qk}(\mathbf{w})] \\ &= \min_{\mathbf{w}} Var_{t-1} [\mathbf{w} \cdot \mathbf{x}_t + W_0 r_{tb,t}] - 2Cov_{t-1} [qk_t, \mathbf{w} \cdot \mathbf{x}_t + W_0 r_{tb,t}] \\ &= \min_{\mathbf{w}} \mathbf{w} Var_{t-1} [\mathbf{x}_t] \mathbf{w}^T + \mathbf{w} \cdot \{-2Cov_{t-1} [\mathbf{x}_t, qk_t]\}\end{aligned}$$

bigger but comparable to the 50 basis points which are usually considered as the market standard for longevity risk (see Loeys et al. (2007)).

The above measures of aggregate longevity risk and the values of its replication portfolios can be revised on a yearly basis, at the highest revision frequency of the mortality rates. By making ALHP tradable on a yearly basis, one could sell aggregate longevity protection without incurring in the liquidity problem of long duration securities.

According to Campbell (1996), an intertemporal utility-maximizing agent will optimally demand to invest or sell the hedging portfolio for aggregate longevity, if the state variable qk forecasts changes in financial or human capital. The empirical estimation of the extended VAR shows that the aggregate longevity shock indeed predicts price changes in annuity prices and in long term bonds, and supports the hypothesis of existence of non-zero potential demand for ALS. While a complete discussion of the demand for longevity-linked securities requires a structural equilibrium framework like, for example, the model of Cocco and Gomes (2012), in the next section we estimate the size of the compensation for bearing longevity risk assuming that the set of investment opportunities also includes the ALS.

5.1. Pricing longevity risk

Milevsky et al. (2005) propose to use the notion of Sharpe Ratio as an actuarial measure of aggregate longevity risk compensation. While the Sharpe Ratio of an investment is determined by the ratio between the expected return from the investment in excess to a benchmark security (usually the T-bill) and the expected volatility, in actuarial science the Sharpe Ratio determines the excess markup per unit of volatility that an aggregate longevity protection seller would charge to the protection buyer.

The discussion of the previous subsections suggests the possibility to use the information conveyed by the VAR dynamic model to estimate a longevity risk compensation. It is easy to understand that within our framework the Sharpe Ratio of the ALHP is a reliable measure

of such compensation. Note that by definition the ALHP is a zero investment portfolio, as

$$\begin{aligned}\mathbf{w}^{ALHP} &= [w_{xrtb,t}^{ALHP}, w_{xr,t}^{ALHP}, w_{xb,t}^{ALHP}, w_{x\Delta p,t}^{ALHP}], \\ \mathbf{w}^{ALHP} \cdot \boldsymbol{\iota} &= 0,\end{aligned}$$

and that the corresponding return can be split as the differential between the return of a long position in financial securities and a short position in the ALS liability. Hence, the ratio between the expected differential return and its risk provides a properly defined Sharpe Ratio. Moreover, differently from previous approaches, our estimation procedure identifies a dynamic compensation component whose evolution is maximally correlated with longevity shocks as opposed to other intermediation margins, which are instead not expected to be correlated with longevity shocks.

Conditional expectations of risks and returns may substantially differ from unconditional ones, and therefore the conditional Sharpe ratios depend on the holding period τ and generate a term structure. This term structure depends on the initial level of the VAR state variables corresponding to the current level of financial returns and the current level of predictors. Over an horizon τ the Sharpe Ratio is the ratio between the τ -period expected excess simple return and the τ -period standard deviation. Hence, recalling that the extended VAR(1) models logarithmic returns we have:

$$\begin{aligned}SR_{\tau, \mathbf{r}_t} &= \frac{Ex_{\tau, \mathbf{r}_t}}{Std_{\tau, \mathbf{r}_t}}, \\ Std_{\tau, \mathbf{r}_t}(\mathbf{w}) &= \sqrt{Var_{t, \mathbf{r}_t} \left[\sum_{k=1}^{\tau} \mathbf{w} \cdot (\mathbf{r}_{t+k} + r_{0,t+k} \boldsymbol{\iota}) \right] / \tau}, \\ Std_{\tau, \mathbf{r}_t}(w_{rtb}) &= \sqrt{Var_{t, \mathbf{r}_t} \left[\sum_{k=1}^{\tau} w_{rtb} r_{0,t+k} \right] / \tau}, \\ Ex_{\tau, \mathbf{r}_t} &= E_{t, \mathbf{r}_t} \left[\frac{1}{\tau} \sum_{k=1}^{\tau} \mathbf{w} \cdot \mathbf{r}_{t+k} \right] + \frac{Std_{\tau, \mathbf{r}_t}^2(\mathbf{w})}{2\tau}.\end{aligned}$$

Figure D.9 reports the term structure of equity Sharpe Ratios obtained by setting the initial

condition of the state variables to their long-term expected values.

Insert Figure D.9 about here

The above conditional model shows that the relative compensation of different securities as measured by the Sharpe Ratios depends on the horizon and on the state of the economy. When the holding period goes to infinity ($\tau \rightarrow \infty$) the Sharpe Ratio SR_{τ, \mathbf{r}_t} for a generic portfolio w converges to a limit SR_{∞} :¹²

$$\begin{aligned}
 SR_{\infty} &= \frac{Ex_{\infty}}{Std_{\infty}}, \\
 Ex_{\infty} &= \lim_{\tau \rightarrow +\infty} \left\{ E_{t, \mathbf{z}_t} \left[\frac{1}{\tau} \sum_{k=1}^{\tau} \mathbf{w} \cdot \mathbf{r}_{t+k} \right] + \frac{Std_{\tau, \mathbf{r}_t}^2(\mathbf{w})}{2\tau} \right\}, \\
 Std_{\infty} &= \lim_{\tau \rightarrow +\infty} \sqrt{Var_{t, \mathbf{r}_t} \left[\sum_{k=1}^{\tau} \mathbf{w} \cdot (\mathbf{r}_{t+k} + r_{0, t+k} \mathbf{l}) \right] / \tau},
 \end{aligned} \tag{9}$$

As shown in the Figure D.9, convergence of the SR_{τ, \mathbf{r}_t} to SR_{∞} does occur at very long horizons, but the limiting procedure is necessary in order to produce a bona fide unconditional measure of expected performance consistent with the predictability patterns we documented. Table D.5 reports the estimation of long term Sharpe Ratios for all the financial securities included in the extended VAR(1): an equity index, a rolling position in long-term bonds and the ALS. As expected, the Sharpe Ratio of the ALS security is negative, as its performance is lower than that of benchmark risk free security, the T-Bill. On the other hand, the low level of the period variance implies that the unconditional level of the ALS Sharpe Ratio is as high as $SR_{\infty}^{ALS} = (-)0.43$, indicating the potential usefulness of this synthetic security as a “liability” offering a good potential reward to investors seeking new diversification

¹²These quantities can be easily computed using the following expressions of the long run mean and covariance:

$$\begin{aligned}
 \mu_{\infty} &= (\mathbb{I} - \Phi_1)^{-1} \Phi_0 \\
 \Sigma_{\infty} &= (\mathbb{I} - \Phi_1)^{-1} \Phi_1 (\mathbb{I} - \Phi_1)^{-T}
 \end{aligned}$$

strategies. As this liability is financed by a short term T-Bill, and the maturity mismatch is well known to increase interest rate risk variations, the high value of the Sharpe Ratio can however be misleading.

A similar problem arises when measuring the Sharpe Ratio of the ALHP, the measure of aggregate longevity risk premium. The ALHP can be split in a short position in a ALS security and a long position in a portfolio of traded financial securities, thus making the Sharpe Ratio depend on the composition of the portfolio used to finance the short position.

Table D.6 reports the Sharpe measure of aggregate longevity risk compensation for the three alternative allocations defined in Table D.4: the unrestricted one $\mathbf{w}^{ALHP-U_{nr}}$, the one excluding allocation to equity, \mathbf{w}^{ALHP-1} , and the one where the ALS stochastic liability can be hedged using only long term bonds \mathbf{w}^{ALHP-2} . As the Table shows, Sharpe Ratios decrease with increasing restrictions. Notice that the opposite of the Unrestricted allocation can well be interpreted as a feasible investment policy for an asset management company which is using a leverage policy similar to the one adopted by Berkshire. In this case the optimal allocation corresponds to assume an insurance liability which is selling longevity risk protection (the opposite position with respect to the one originally accepted by Berkshire in the life retrocession deal with Swiss Re) to finance a long position in equity and bond markets.¹³ The Sharpe Ratio value of 0.50 for this strategy witnesses the benefits which can be achieved using such a form of optimal asset and liability management policy and confirms that this allocation strategy is one of the ingredients that contributed to raise Berkshire performance up to the 0.66 Sharpe ratio computed in Frazzini Kabiller and Pedersen (2013).

In fact this anomalous performance sharply disappears when the allocation used to hedge the stochastic liability induced by the short position in the ALS are constrained. The Sharpe Ratio is the lowest if the investor is allowed to hedge using long term bonds, the security which is typically held in the reserves of insurance companies. In fact, the hedging portfolio \mathbf{w}^{ALHP-2} using only long term bonds is safer (but less profitable) than \mathbf{w}^{ALHP-1} which also uses T-bills, as its duration matches that of the stochastic liability and thus has identical

¹³As on february 5th 2013, Berkshire Hathaway announced a deal to reinsure the risks associated with \$4 billion of future claims for two of health insurer Cigna Corp's run-off variable annuity businesses.

response to (small) interest rate fluctuations.

In conclusion, the actuarial longevity premium estimate consistent with a prudent hedging policy is given by $SR_{\infty}^{ALHP.2} = 0.33$. Its value is not far from the conventional level 0.25 used in the actuarial pricing of longevity products as discussed in Loeys et al. (2007). This estimation is expected to overestimate the potential Sharpe Ratio from longevity liability, as shorting costs are not explicitly accounted in this analysis. In addition the adverse selection effect is also expected to play a role here: the mortality rates of annuitants are known to be significantly smaller from those of average population (see Poterba (2001) and Mitchell et al. (1999)). Given the scarcity of data on prices of traded longevity-linked securities, the same problems affect virtually any empirical measure of longevity risk compensation.

It is possible to draw some conclusion on the robustness of the performance of a portfolio like $\mathbf{w}^{ALHP-U_{nr}}$ with respect to exogenous liquidity shocks thanks to the analysis of Kojien and Yogo (2014) of the dynamics of annuity prices during the recent financial crisis period 2008-2009. They prove that during the financial crisis, life insurers sold long-term policies at deep discounts relative to their actuarial value and this discount was supply driven. The average markup was as low as -19 percent for annuities despite interest rates were falling. In fact insurance companies were trying to sell more annuities in the attempt to increase statutory capital i.e., assets relative to accounting liabilities. By selling new policies life insurers were able to raise capital because their reserve valuation was more aggressive than mark to market during the financial crisis. In a way, life insurance policies sold directly to households were used by insurance companies as a source of short term funding when traded assets like equity and bonds were hit by financial markets illiquidity. On the other hand, longevity risk is poorly correlated with market risk and short term volatility of longevity shocks is extremely low, thus the liquidity was provided by annuities at the expense of a raise in longevity risk exposure having a negligible impact over the short term. In other terms, diversification benefits created by the possibility to use longevity liabilities to finance financial market investment seem to be robust to exogenous liquidity shocks.

From a normative point of view these considerations indicate a further indirect motivation to promote the integration between financial and actuarial markets: their development would

drive a more transparent and precise assessment of the price for aggregate longevity risk. Among other benefits it is worth mentioning that this assessment can certainly reduce the dangerous unawareness of the public and private costs deriving from generalized longevity increase. Unreported robustness checks show that our results are unaffected if the estimation is based on maximum and minimum annuity premia rather than on the mean ones. Similarly, the estimated values are robust to changes (reductions) of the sample used for the extended VAR(1) estimation.

6. Conclusions

Our analysis shows that a promising direction to improve the efficiency of longevity risk sharing is integration between insurance and financial markets. We believe that our results uncover some critical issues to improve longevity risk securitization. First, the long term nature of the longevity risk requires an accurate analysis of the term structure of the risk return trade-offs generated by including a longevity-linked security in the set of investments. Second, a potentially large number of short term investors would be willing to increase their exposure to longevity risk without increasing their investment horizon. This requires the organization of a maturity transformation activity by financial intermediaries that seems to be a crucial step to increase the interest of the market for longevity-linked securities, as well as their liquidity. Third, an integrated market for insurance and financial contracts with a publicly traded longevity index would also imply a more transparent and efficient pricing of life annuities with a direct benefit for annuity subscribers.

- Barberis, N., 2000. Investing for the long run when returns are predictable. *The Journal of Finance* 55 (1), 225–264.
- Bayraktar, E., Milevsky, M. A., Promislow, S. D., Young, V. R., 2009. Valuation of mortality risk via the instantaneous Sharpe ratio: applications to life annuities. *Journal of Economic Dynamics and Control* 33 (3), 676–691.
- Bayraktar, E., Young, V. R., 2008. Pricing options in incomplete equity markets via the instantaneous Sharpe ratio. *Annals of Finance* 4 (4), 399–429.
- Biffis, E., Blake, D., 2010. Securitizing and tranching longevity exposures. *Insurance: Mathematics and Economics* 46, 186–197.
- Biffis, E., Denuit, M., Devolder, P., 2010. Stochastic mortality under measure changes. *Scandinavian Actuarial Journal* 2010 (4), 284–311.
- Blake, D., Burrows, W., 2001. Survivor bonds: Helping to hedge mortality risk. *Journal of Risk and Insurance* 68, 339–348.
- Blake, D., Cairns, A., Coughlan, G., Dowd, K., 2013. The new life market. *Journal of Risk and Insurance* 80, 501–557.
- Brown, J. R., Mitchell, O. S., Poterba, J. M., 2002. Mortality risk, inflation risk and annuity products. In: Mitchell, O. S., Bodie, Zvi, H. P. B., Zeldes, S. (Eds.), *Innovations in Retirement Financing*. University of Philadelphia Press, Philadelphia, PA.
- Campbell, J. Y., 1996. Understanding risk and return. *Journal of Political Economy* 61 (1).
- Campbell, J. Y., Chan, Y. L., Viceira, L. M., 2003. A multivariate model of strategic asset allocation. *Journal of Financial Economics* 67 (1), 41–80.
- Campbell, J. Y., Lo, A. W., MacKinlay, A. C., Whitelaw, R. F., 1998. *The econometrics of financial markets*. Cambridge Univ Press.
- Campbell, J. Y., Shiller, R. J., 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of financial Studies* 1 (3), 195–228.
- Campbell, J. Y., Viceira, L. M., 2002. *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford University Press, New York, NY.
- Campbell, J. Y., Viceira, L. M., 2005. The term structure of the risk-return tradeoff. *Financial Analysts Journal* 61 (1), 34–44.
- Cocco, J. F., Gomes, F. J., 2012. Longevity risk, retirement savings, and financial innovation. *Journal of Financial Economics* 1034 (3), 507–529.
- Cox, S. H., Lin, Y., 2007. Natural hedging of life and annuity mortality risks. *North American Actuarial Journal* 11 (3), 1–15.
- Finkelstein, A., Poterba, J. M., 2004. Selection in insurance markets: policyholder evidence from the U.K.

- annuity market. *Journal of Political Economy* 112 (1), 183–208.
- Frazzini, A., Kabiller, D., Pedersen, L. H., 2013. Buffetts alpha. Tech. rep., National Bureau of Economic Research.
- Friedberg, L., Webb, A., 2007. Life is cheap: using mortality bonds to hedge aggregate mortality risk. *The B.E. Journal of Economic Analysis & Policy* 7 (1), 31.
- Friedman, B. M., Warshawsky, M., 1988. Annuity prices and saving behavior in the United States. In: Bodie, Z., Shoven, J. B., Wise, D. A. (Eds.), *Pensions in the U.S. Economy*. University of Chicago Press, Chicago, IL.
- Horneff, Wolfram J. and Maurer, R. H., Mitchell, O. S., Stamos, M. Z., 2009. Asset allocation and location over the life cycle with survival-contingent payouts. *Journal of Banking and Finance* 33 (9), 1688–1699.
- Koijen, R. S., Yogo, M., 2012. The cost of financial frictions for life insurers. Tech. rep., National Bureau of Economic Research.
- Koijen, R. S., Yogo, M., 2013. Shadow insurance. Tech. rep., National Bureau of Economic Research.
- Lee, R., 2000. The lee-carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal* 4 (1), 80–93.
- Lee, R. D., Carter, L. R., 1992. Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association* 87 (419), 659–671.
- Lin, Y., Cox, S. H., 2005. Securitization of mortality risks in life annuities. *The Journal of Risk and Insurance* 72 (2), 227–252.
- Lin, Y., Cox, S. H., 2008. Securitization of catastrophe mortality risks. *Insurance: Mathematics and Economics* 42 (2), 628–637.
- Loeys, J., Panigirtzoglou, N., Ribeiro, R. M., 2007. Longevity: a market in the making. Tech. rep., J.P.Morgan Global Market Strategy.
- Markowitz, H., 1952. Portfolio selection. *The Journal of Finance* 7 (1), 77–91.
- Mehra, R., Prescott, E. C., 1985. The equity premium: a puzzle. *Journal of monetary Economics* 15 (2), 145–161.
- Merton, R. C., 1973. An intertemporal capital asset pricing model. *Econometrica*, 867–887.
- Milevsky, M. A., 2006. *The calculus of retirement income*. Financial models for pension annuities and life insurance. Cambridge University Press Cambridge, UK.
- Milevsky, M. A., Promislow, S. D., Young, V. R., 2005. Financial valuation of mortality risk via the instantaneous Sharpe ratio. IFID Center, Toronto, Canada.
- Milevsky, M. A., Promislow, S. D., Young, V. R., 2006. Killing the law of large numbers: mortality risk premiums and the Sharpe ratio. *Journal of Risk and Insurance* 73 (4), 673–686.
- Mitchell, O. S., Poterba, J. M., Warshawsky, M. J., Brown, J. R., 1999. New evidence on the money’s worth

- of individual annuities. *The American Economic Review* 89 (5), 1299–1318.
- Pelsser, A., 2008. On the applicability of the wang transform for pricing financial risks. *Astin Bulletin* 38 (1), 171.
- Pitacco, E., Denuit, M., Haberman, S., Olivieri, A., 2009. *Modelling longevity dynamics for pensions and annuity business*. Oxford University Press.
- Poterba, J. M., 2001. The history of annuities in the united states. In: *The Role of Annuity Markets in Financing Retirement*. The MIT Press, Cambridge, MA.
- Wang, S., 2002. A universal framework for pricing financial and insurance risks. *Astin Bulletin* 32 (2), 213–234.
- Warshawsky, M., 1988. Private annuity markets in the united states: 1919-1984. *Journal of Risk and Insurance*, 518–528.

Appendix A. Identification and estimation of the Lee-Carter mortality model

The Lee and Carter (1992) model consists of a system of equations for logarithms of central death rates for age cohort x at time t , $m_{x,t}$, and a time-series equation for an unobservable time-varying mortality index k_t :

$$\begin{aligned} \ln(m_{x,t}) &= a_x + b_x k_t + \epsilon_{x,t} & (A.1) \\ k_t &= c_0 + c_1 k_{t-1} + e_t \\ \epsilon_{x,t} &\sim NID(0, \sigma_\epsilon^2) \\ e_t &\sim \text{MeanZero - Stationary Process} \end{aligned}$$

where a_x and b_x are age-specific constants. The error term $\epsilon_{x,t}$ captures cross-sectional errors in the model based prediction for mortality of different cohorts, while the error term e_t captures random fluctuations in the time series of the common factor k_t driving mortality at all ages. This common factor, usually known as the unobservable mortality index evolves over time as an autoregressive process and the favorite Carter-Lee specification makes is a unit-root process by setting $c_1 = 1$. Identification is achieved by imposing the restrictions $\sum_t k_t = 0$ and $\sum_x b_x = 1$, so that the unobserved mortality index k_t is estimated through Singular Value Decomposition (SVD). SVD is a technique based on a theorem of linear algebra stating that a $(m \times n)$ rectangular matrix M can be broken down into the product of three matrices - an $(m \times m)$ orthogonal matrix U , a diagonal $(m \times n)$ matrix S , and the transpose of an orthogonal $(n \times n)$ matrix V . The SVD of the matrix M will be therefore be given by $M = USV'$ where $U'U = I$ and $V'V = I$. The columns of U are orthonormal eigenvectors of AA' , the columns of V are orthonormal eigenvectors of $A'A$, and S is a diagonal matrix whose elements are the square roots of eigenvalues from U or V in descending order. The restriction $\sum_t k_t = 0$ implies that a_x is the average across time of

$q_{x,t}$, and Equation A.1 can be rewritten in terms of the mean-centered log-mortality rate as

$$m_{x,t} - \bar{m}_{x,t} \equiv \tilde{m}_{x,t} = b_x k_t + \epsilon_{x,t}. \quad (\text{A.2})$$

Grouping all the $\tilde{m}_{x,t}$ in a unique $(X \times T)$ matrix \tilde{m} (where the columns are mortality rates at time- t ordered by age groups and the rows are mortality rates through time for a specific age-group x), leads naturally to use SVD to obtain estimates of b_x and k_t . In particular, if \tilde{m} can be decomposed as $\tilde{m} = USV'$, $b = [b_0, b_1, \dots, b_X]$ is represented by the normalized first column of U, $u_1 = [u_{0,1}, u_{1,1}, \dots, u_{X,1}]$, so that

$$b = \frac{u_1}{\sum_{x=0}^X u_{x,1}}.$$

On the other hand the mortality index vector $k = [k_1, k_2, \dots, k_T]$ is given by

$$k = \lambda_1 \left(\sum_{x=0}^X u_{x,1} \right) \nu_1$$

where $\nu_1 = [\nu_{1,1}, \nu_{1,2}, \dots, \nu_{1,T}]'$ is the first column of the V matrix and λ_1 is the highest eigenvalue of the matrix S . The values of mortality rates obtained with this method will not, in general, be equal to the actual number of deaths. In Lee and Carter (1992), the authors hence re-estimate k_t in a second step, taking the values of a_x and b_x as given from the first-step SVD estimate and using actual mortality rates. The new values of k are obtained so that, for each year, the actual death rates are equal to the implied ones. This two-step procedure allows to take into account the population age distribution, providing a very good fit for 13 of the 19 age groups in the authors' sample, where more than 95% of the variance over time is explained. For seven of these, more than 98% of the variance is explained.

Appendix B. Financial asset returns and their predictors: the basic specification of the VAR model for traditional financial investments

Consider the continuously compounded security market returns from time t to time $t+1$, \mathbf{r}_{t+1} . Define μ_t , the conditional expected log return given information up to time t as follows:

$$\mathbf{r}_{t+1} = \mu + \mathbf{u}_{t+1},$$

where \mathbf{u}_{t+1} is the unexpected log return. Define the τ -period cumulative return from period $t+1$ through period $t+\tau$, as

$$\mathbf{r}_{t,t+\tau} = \sum_{i=1}^{\tau} \mathbf{r}_{t+i}.$$

The term structure of risk is defined as the conditional variance of cumulative returns, given the investor's information set, scaled by the investment horizon

$$\Sigma_r(\tau) \equiv \frac{1}{\tau} \text{Var}(\mathbf{r}_{t,t+\tau} \mid D_t), \quad (\text{B.1})$$

where $D_t^{Mkt} \equiv \sigma\{z_\tau^{Mkt} : \tau \leq t\}$ consists of the full histories of returns as well as predictors that investors use in forecasting returns. Following Barberis (2000) and Campbell and Viceira (2002, 2005), we describe asset return dynamics by means of a first-order vector autoregressive or VAR(1) model. We choose a VAR(1) as the inclusion of additional lags, even if easily implemented, would reduce the precision of the estimates:

$$\mathbf{z}_t^{Mkt} = \Phi_0^{Mkt} + \Phi_1^{Mkt} \mathbf{z}_{t-1}^{Mkt} + \nu_t^{Mkt}, \quad (\text{B.2})$$

where

$$\mathbf{z}_t^{Mkt} = \begin{bmatrix} r_{0t} \\ \mathbf{x}_t^{Mkt} \\ \mathbf{s}_t^{Mkt} \end{bmatrix}$$

is a $(m \times 1)$ vector, with r_{0t} being the log real return on the asset used as a benchmark to compute excess returns on all other asset classes, \mathbf{x}_t^{Mkt} being the $(n \times 1)$ vector of log

excess returns on all other asset classes with respects to the benchmark, and \mathbf{s}_t^{Mkt} is the $((m - n - 1) \times 1)$ vector of returns predictors. In the VAR(1) specification, Φ_0^{Mkt} is a $(m \times 1)$ vector of intercepts and Φ_1^{Mkt} is a $(m \times m)$ matrix of slopes. Finally, ν_t^{Mkt} is a $(m \times 1)$ vector of innovations in asset returns and returns' predictors for which standard assumptions apply, i.e.

$$\nu_t^{Mkt} \sim \mathcal{N}(0, \Sigma_\nu^{Mkt}), \quad (\text{B.3})$$

where Σ_ν^{Mkt} is the $(m \times m)$ variance-covariance matrix. Note that

$$\Sigma_\nu^{Mkt} = \begin{bmatrix} \sigma_0^2 & \sigma'_{0x} & \sigma'_{0s} \\ \sigma_{0x} & \Sigma_{xx} & \Sigma'_{xs} \\ \sigma_{0s} & \Sigma_{xs} & \Sigma_{ss} \end{bmatrix}$$

and the unconditional mean and variances-covariance matrix of \mathbf{z}_t , assuming that the VAR is stationary and therefore that these moments are well-defined, can be represented as follows:

$$\begin{aligned} \mu_z^{Mkt} &= (I_m - \Phi_1^{Mkt})^{-1} \Phi_0^{Mkt} \\ \text{vec}(\Sigma_{zz}^{Mkt}) &= (I_{m^2} - \Phi_1^{Mkt} \otimes \Phi_1^{Mkt})^{-1} \text{vec}(\Sigma_\nu^{Mkt}). \end{aligned}$$

The conditional mean of the cumulative asset returns at different horizons are instead

$$E_t(z_{t+1}^{Mkt} + \dots + z_{t+\tau}^{Mkt}) = \left(\sum_{i=0}^{\tau-1} (\tau - i) (\Phi_1^{Mkt})^i \right) \Phi_0^{Mkt} + \left(\sum_{j=0}^{\tau} (\Phi_1^{Mkt})^j \right) \mathbf{z}_t^{Mkt},$$

and their variance is:

$$\begin{aligned} \text{Var}_t(z_{t+1}^{Mkt} + \dots + z_{t+\tau}^{Mkt}) &= \Sigma_\nu^{Mkt} + (I + \Phi_1^{Mkt}) \Sigma_\nu^{Mkt} (I + \Phi_1^{Mkt})' + \\ &\quad (I + \Phi_1^{Mkt} + (\Phi_1^{Mkt})^2) \Sigma_\nu^{Mkt} (I + \Phi_1^{Mkt} + (\Phi_1^{Mkt})^2)' + \dots \\ &\quad + (I + \Phi_1^{Mkt} + \dots + (\Phi_1^{Mkt})^{\tau-1}) \Sigma_\nu^{Mkt} (I + \Phi_1^{Mkt} + \dots + (\Phi_1^{Mkt})^{\tau-1})'. \end{aligned} \quad (\text{B.4})$$

Once the conditional moments of excess returns are available the following selector matrix

extracts for each period, τ -period conditional moments of log real returns

$$M_r = \begin{bmatrix} 1 & 0_{1 \times n} & 0_{1 \times (m-n-1)} \\ \iota_{n \times 1} & I_{n \times n} & 0_{n \times (m-n-1)} \end{bmatrix}$$

which implies

$$\frac{1}{\tau} \begin{bmatrix} E_t(r_{0,t+1}^\tau) \\ E_t(\mathbf{r}_{t+1}^\tau) \end{bmatrix} = \frac{1}{\tau} M_r E_t(z_{t+1}^{Mkt} + \dots + z_{t+\tau}^{Mkt})$$

$$\frac{1}{\tau} \begin{bmatrix} Var_t(r_{0,t+1}^\tau) \\ Var_t(\mathbf{r}_{t+1}^\tau) \end{bmatrix} = \frac{1}{\tau} M_r Var_t(z_{t+1}^{Mkt} + \dots + z_{t+\tau}^{Mkt}) M_r'$$

Therefore after the estimation of the VAR it is possible to derive unconditional and conditional moments for returns and excess returns at all different investment horizons. These moments deliver the dynamics of returns and the risk of different assets across investment horizons. This information forms the input for portfolio allocation. Following Campbell and Viceira (2005), we consider a benchmark portfolio to be obtained by attributing optimal weights to bond, stock and T-bills. Therefore we include in \mathbf{x}_t^{Mkt} excess returns on stocks and bonds, real returns on T-bills, while we include in \mathbf{s}_t^{Mkt} three factors commonly recognized as good predictors of these assets' returns. In particular, the predictors are the nominal short-term interest rate, the dividend price ratio and the yield spread between long-term and short term bonds.

Appendix C. The dataset

Appendix C.1. Annuity prices

We estimate the change in the annuity price included in our VAR(1) model as

$$\Delta p_{65,t+1} = \ln \left(\frac{P_{t+1,65}^A}{P_{t,65}^A} \right),$$

where $P_{t+1,65}^A$ is the annuity price offered on the US market for 1\$ monthly life annuities written on 65-year-old males. Consequently Δp_{t+1}^A is the yearly log price change of a standardized 65 year-old annuity between time t and $t + 1$. Our annuity prices consist of the sample average premiums for immediate \$1 monthly life annuities for 65-year old US males issued during the 1952-2010 period. In order to have the longest time series of prices, we collected premia from different sources. Following Warshawsky (1988) and Friedman and Warshawsky (1988), premia over the 1952-1967 period come from successive annual issues of Spectator’s Handy Guide and A.M. Best’s Flitcraft Compend , whereas premia over the 1968-1985 years come from the successive annual issues of A.M. Best’s Flitcraft Compend. Following Koijen and Yogo (2012), Cox and Lin (2007) and Brown et al. (2002), we compile premia for the 1986-2010 years from the semiannual issues of Annuity Shopper. These data are integrated with those obtained from the annual issues of the Life/Health editions of Best’s Review for the 1995-1998 period.

Given the length of the sample period and the different sources of data we use, our sample premia refer to an unbalanced panel of companies. Although the correct approach should be to use only rates reported by the same companies, this would substantially reduce the number of premia available each year for computing the minimum, the maximum and the average annual annuity premium which, consequently, might not be reflective of the true value of annuity price. The pricing approach adopted in the present paper is essentially based on the assumption that price changes of annuities reflect changes of fundamental risks priced by buyers and sellers. A side product of this estimation analysis is a direct empirical measure of the effective price reaction to changes in aggregate mortality trends. Clearly, empirical evidence of this connection would support the hypothesis that some form of competition drives the prices in the market for annuities.

Appendix C.2. Marketed securities and the state variables

The Campbell and Viceira (2005) model is developed using quarterly data. As adapting the mortality series to this frequency is both hardly feasible (lack of mortality data for frequencies higher than yearly) and less meaningful (for example, some months of the year

experience higher mortality rates than others), we focus our analysis on annual data. We download the financial data from Robert Shiller's website¹⁴ for the postwar period 1952-2010 and, following Campbell et al. (2003) construct the financial time series as:

- Short-term ex-post real rate: return on 6-month commercial paper bought in January and rolled over July, minus the Producer Price Index (PPI).
- Excess return on stocks: log return on the S&P 500 Stocks, from which the short-term interest rate is subtracted.
- Excess return on bonds: returns are obtained using the loglinear approximation described in Campbell et al. (1998)

$$r_{n,t+1} = D_{n,t}y_{n,t} - (D_{n,t} - 1)y_{n-1,t+1},$$

where n is the Bond maturity, the Bond yield is $Y_{n,t}$, the log Bond yield is $y_{n,t} = (1 + Y_{n,t})$ and $D_{n,t}$ is the Bond duration, calculated at time t as

$$D_{n,t} \approx \frac{1 - (1 + Y_{n,t})^{-n}}{1 - (1 + Y_{n,t})^{-1}}$$

with n set to 20 years and $y_{n-1,t+1}$ approximated by $y_{n,t+1}$.

- Excess annuity prices' growth, as described in the previous subsection of this Appendix. Finally the following set of state variables are included in the VAR to parametrize the opportunity set faced by the investor
- Nominal T-bill rate: return on 6-month commercial paper bought in January and rolled over July.
- Log dividend-price ratio: natural logarithm of the S&P 500 dividend series minus the logarithm of the S&P 500 price series.

¹⁴<http://www.econ.yale.edu/~shiller/data.htm>

- Yield spread: difference between the log yield of the long Bond and the short yield on the commercial paper.
- Aggregate longevity shocks: qk_t as defined in Section 2, are average differences between predicted and fitted mortality rates for the cohorts underlying life annuities (in our case, 65-110).

Appendix D. Figures and Tables

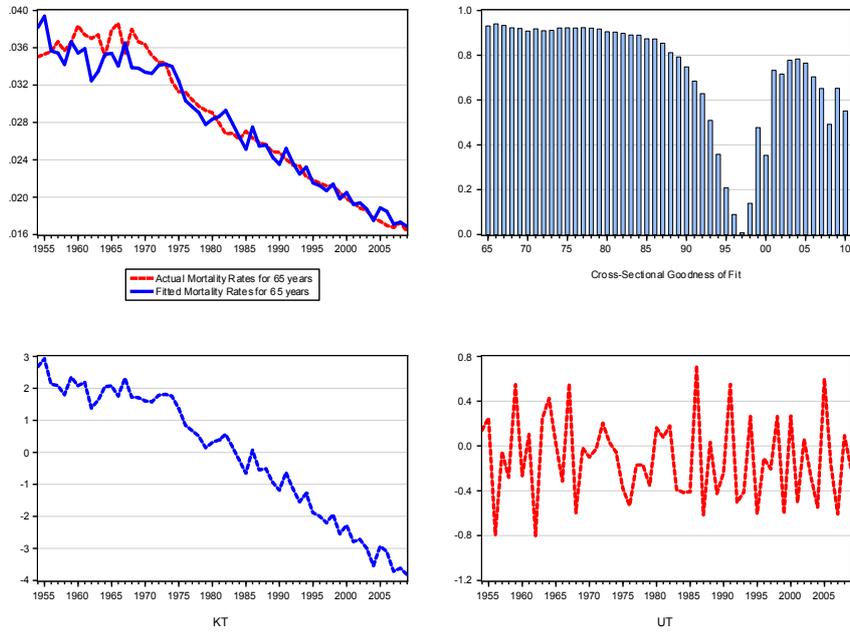


Figure D.1: Lee-Carter Fitted mortality. Figure 1.1: Fitted mortality rates at 65. Figure 1.2: Cross-Sectional R^2 of Equation (3) Figure 1.3: The unobservable mortality index k_t . Figure 1.4: Innovations in k_t .

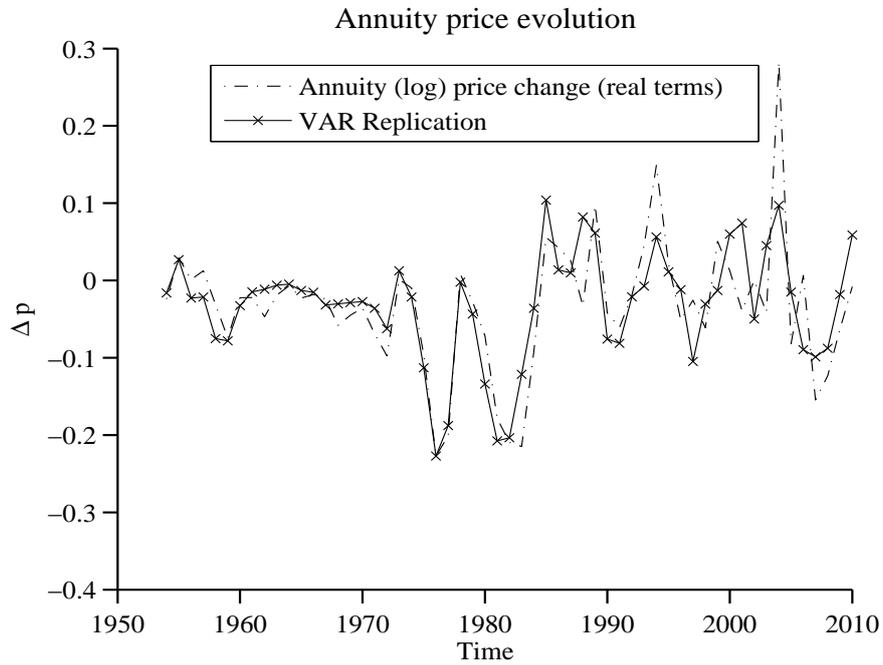


Figure D.2: Time series of historical (real) logarithmic price changes (dashed line) vs the replication as operated by the VAR dynamic model.

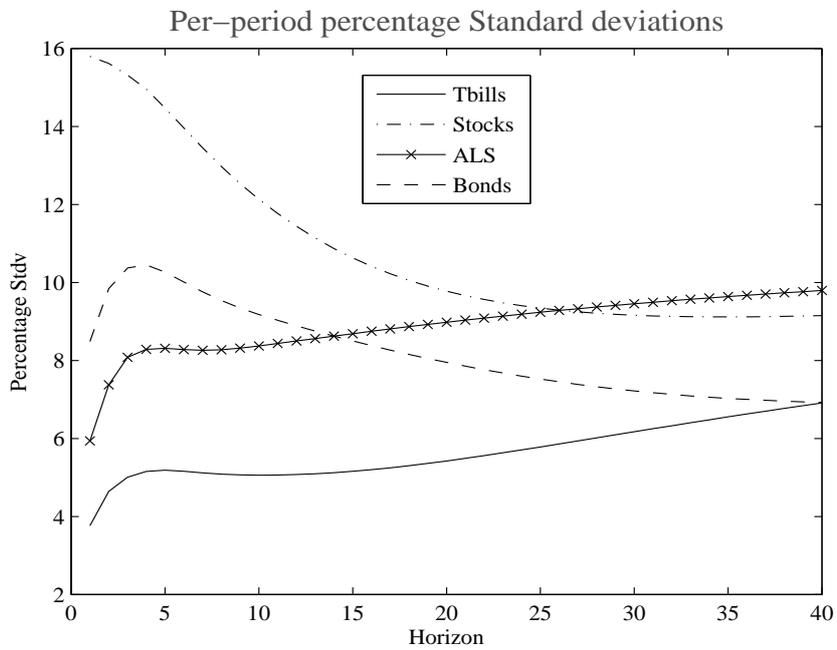


Figure D.3: Term structure of risks for the securities included in the Extended VAR model.

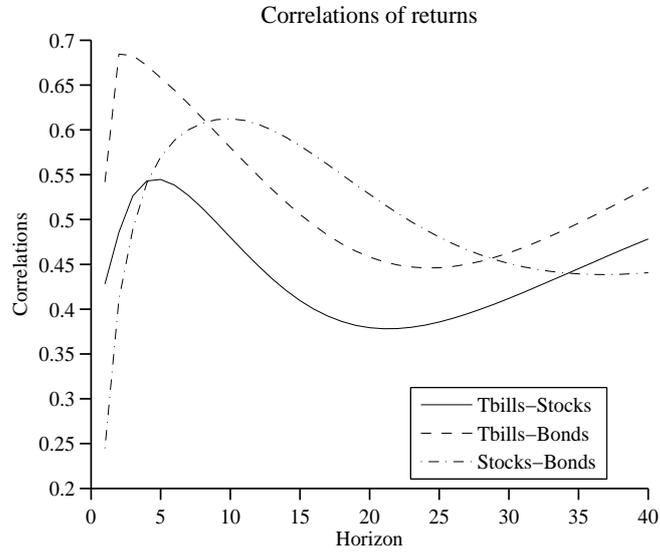


Figure D.4: Term Structure of correlations between financial securities included in the Extended VAR model

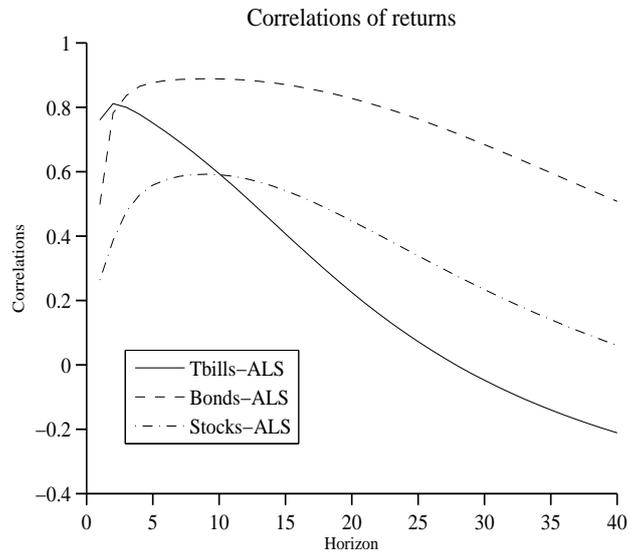


Figure D.5: Term Structure of correlations between financial securities and the Annuity-Linked Security.

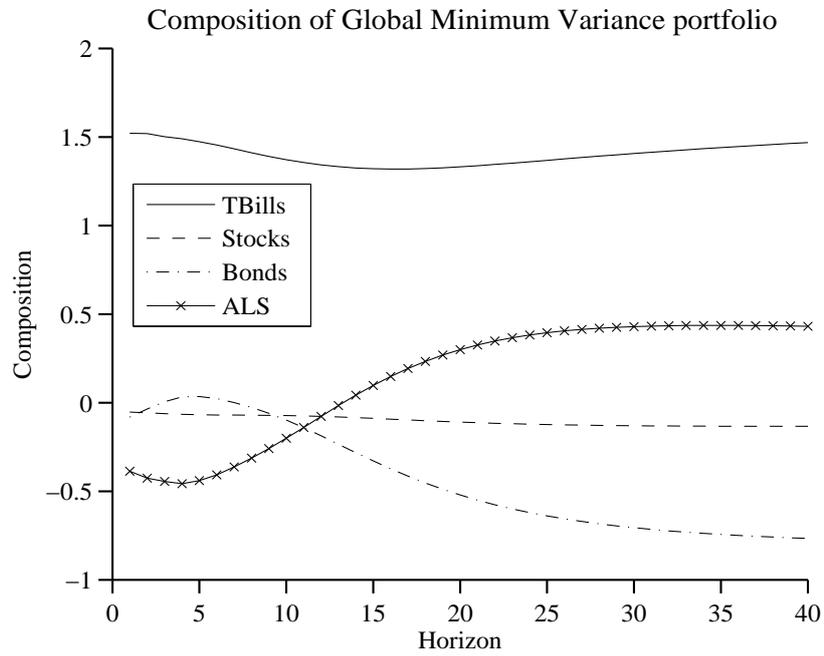


Figure D.6: Term structure of allocations forming the GMV portfolio at different horizons.

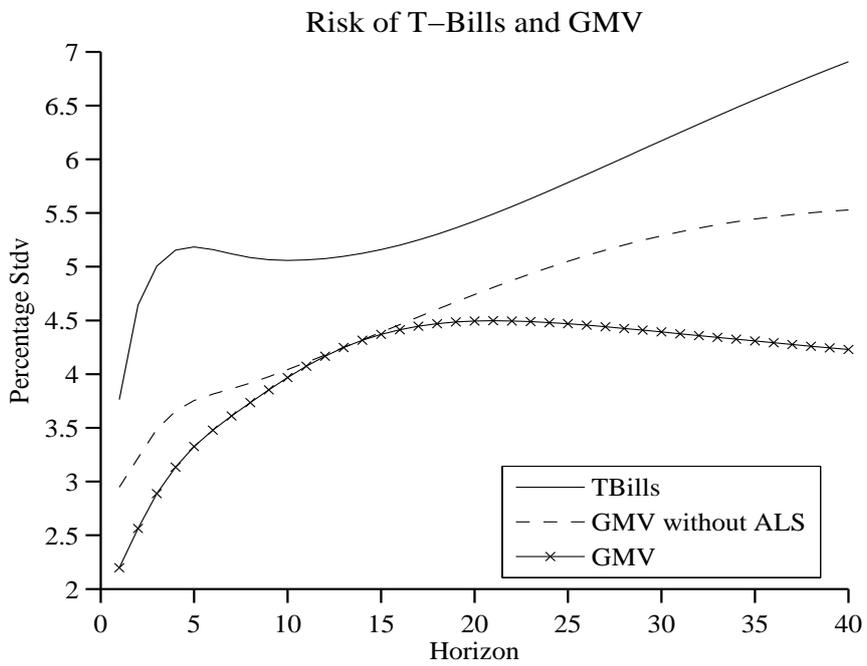


Figure D.7: Term structure of risks for an allocation in T-Bill (continuous line), in the GMV portfolio restricted to financial securities (dashed line), in the GMV portfolio including also the Annuity-Linked Security.

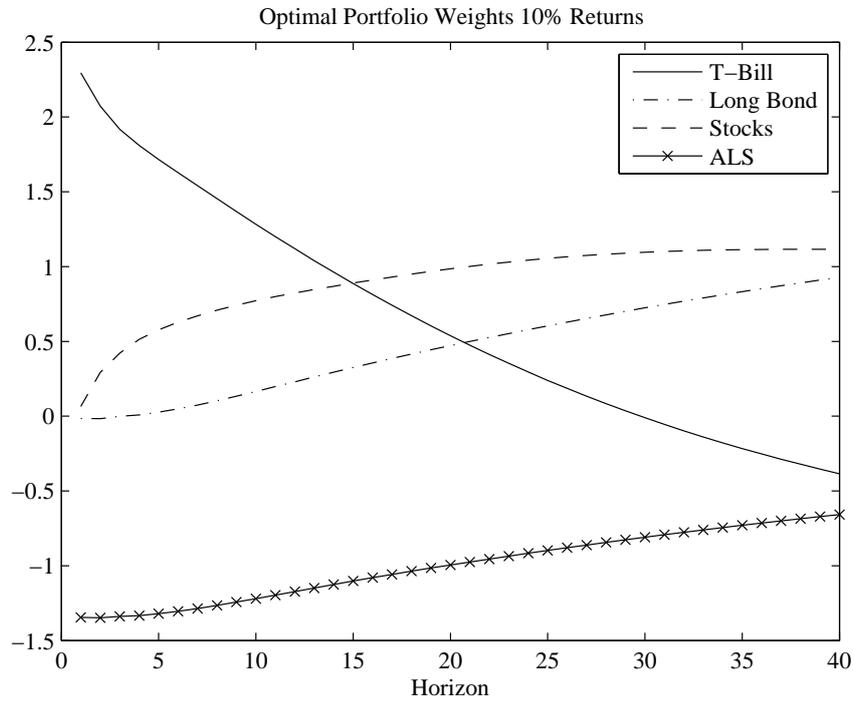


Figure D.8: Term structure of the efficient allocation with target expected return of 10%.

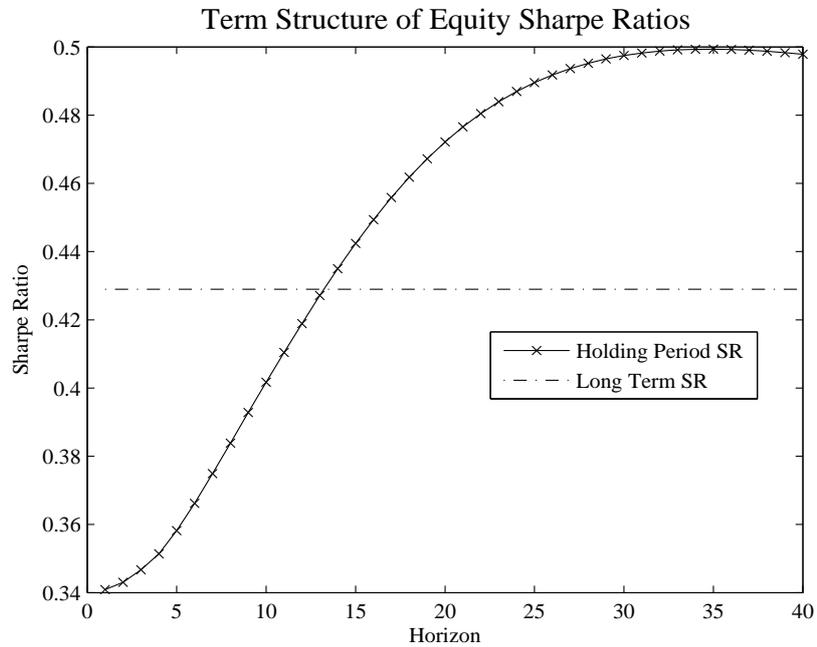


Figure D.9: Term Structure of the Equity Sharpe Ratio where the initial level of the state variables is set to their long term expected value.

	<i>rtb</i>	<i>xr</i>	<i>xb</i>	<i>xΔpr</i>	<i>y</i>	<i>d - p</i>	<i>spr</i>	<i>qk</i>
Mean	0.027	0.056	0.0046	-0.0275	0.056	-3.493	0.004	0.00(4)
Std Dev	0.052	0.160	0.108	0.057	0.031	0.411	0.014	0.041
Sharpe Ratio	-	0.353	0.042593	-0.48246	-	-	-	-

Table D.1: Summary Statistics. Mean returns is computed including the Jensen correction term, thus are computed as $\mu + 0.5\sigma^2$. Sharpe Ratio is computed as the ratio between Mean and Std Dev. **Note:** *rtb* = ex post real T-bill rate, *xr* = excess stock return, *xb* = excess bond return, $(d - p)$ = log dividend-price ratio, *y* = nominal T-bill yield, *spr* = yield spread.

VAR(1) - Matrix Φ_1 - Yearly Sample 1953-2010. Original Financial Variables

	rtb_t	xr_t	xb_t	y_t	$(d-p)_t$	spr_t	R^2	$adjR^2$
	(t)							
rtb_{t+1}	0.567 (3.825)	-0.027 (-0.662)	-0.110 (-1.843)	0.414 (2.350)	-0.008 (-0.462)	-0.640 (-1.507)	0.530	0.471
xr_{t+1}	0.099 (0.192)	0.021 (0.130)	-0.003 (-0.013)	-0.935 (-1.056)	0.107 (1.963)	0.271 (0.171)	0.087	-0.027
xb_{t+1}	0.830 (3.649)	-0.087 (-1.286)	-0.508 (-4.315)	0.246 (0.558)	-0.036 (-1.235)	4.684 (5.132)	0.546	0.489
y_{t+1}	-0.218 (-3.686)	0.036 (2.463)	0.043 (2.053)	1.006 (9.821)	0.005 (1.337)	0.211 (1.214)	0.784	0.757
$(d-p)_{t+1}$	-0.548 (-0.993)	0.071 (0.402)	0.049 (0.237)	0.398 (0.417)	0.939 (16.822)	1.414 (0.787)	0.872	0.855
spr_{t+1}	0.126 (2.778)	-0.030 (-2.680)	0.009 (0.706)	-0.039 (-0.532)	-0.001 (-0.434)	0.405 (3.204)	0.455	0.387

Cross-Correlations of Residuals						
	rtb	xr	xb	y	$(d-p)$	spr
rtb	3.552	0.464	0.070	-0.247	-0.501	0.285
xr	-	13.507	0.170	-0.406	-0.977	0.435
xb	-	-	7.273	-0.637	-0.191	0.160
y	-	-	-	1.403	0.425	-0.845
$(d-p)$	-	-	-	-	14.480	-0.436
spr	-	-	-	-	-	1.066

Table D.2: VAR(1) coefficients with relative t -statistics and Cross-Correlations of Residuals. **Note:** rtb_t = ex post real T-bill rate, xr_t = excess stock return, xb_t = excess bond return, $(d-p)_t$ = log dividend-price ratio, y_t = nominal T-bill yield, spr_t = yield spread.

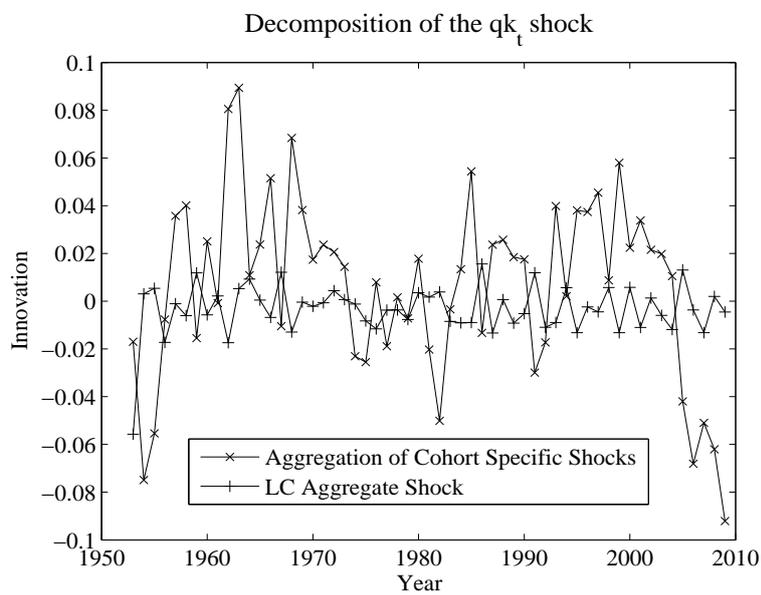


Figure D.10: Contributions to the qk_t shock: time series of aggregate vs sum of cohort-specific innovations as defined in eq.(6.)

VAR(1) - Matrix Φ_1 - Yearly Sample 1953-2010. Annuities

	rtb_t (t)	xr_t (t)	xb_t (t)	$x\Delta pr_t$ (t)	y_t (t)	$(d-p)_t$ (t)	spr_t (t)	$(qk)_t$ (t)	R^2	$adjR^2$
rtb_{t+1}	0.477 (3.265)	0.007 (0.203)	-0.106 (-1.753)	0.074 (0.640)	0.590 (2.534)	-0.010 (-0.568)	-0.680 (-1.452)	-0.016 (-0.116)	0.475	0.389
xr_{t+1}	0.190 (0.388)	0.048 (0.363)	0.131 (0.679)	0.929 (1.660)	-0.215 (-0.225)	0.154 (2.423)	-0.892 (-0.462)	-0.236 (-0.372)	0.167	0.031
xb_{t+1}	0.917 (4.121)	-0.116 (-1.972)	-0.502 (-4.311)	0.116 (0.370)	0.267 (0.550)	-0.054 (-1.921)	4.362 (4.203)	0.588 (2.721)	0.562	0.491
$x\Delta pr_{t+1}$	0.531 (4.294)	-0.099 (-2.228)	0.057 (0.970)	-0.188 (-0.968)	-1.146 (-3.950)	-0.034 (-1.871)	-0.446 (-0.675)	0.271 (2.199)	0.516	0.437
y_{t+1}	-0.204 (-3.702)	0.025 (2.477)	0.038 (1.702)	-0.046 (-0.841)	0.940 (9.253)	0.007 (1.896)	0.286 (1.315)	-0.064 (-1.728)	0.794	0.761
$(d-p)_{t+1}$	-0.812 (-1.392)	0.149 (0.808)	-0.081 (-0.363)	-0.898 (-1.504)	-0.038 (-0.035)	0.885 (13.163)	2.475 (1.141)	0.148 (0.221)	0.845	0.819
spr_{t+1}	0.104 (2.452)	-0.017 (-1.775)	0.014 (1.040)	0.048 (1.069)	0.036 (0.466)	-0.002 (-0.585)	0.339 (2.076)	0.011 (0.404)	0.456	0.367
$(qk)_{t+1}$	-0.113 (-1.341)	-0.051 (-1.671)	0.000 (0.002)	-0.059 (-0.617)	-0.008 (-0.034)	-0.009 (-1.172)	0.373 (1.292)	0.747 (6.625)	0.575	0.506

Cross-Correlations of Residuals

	rtb	xr	xb	Δpr	y	$(d-p)$	spr	(qk)
rtb	3.763	0.206	0.117	0.192	-0.298	-0.234	0.318	-0.104
xr	-	14.584	0.040	-0.063	-0.248	-0.970	0.295	-0.118
xb	-	-	7.184	0.176	-0.620	-0.089	0.132	0.014
Δpr	-	-	-	3.927	-0.547	0.022	0.583	0.044
y	-	-	-	-	1.368	0.262	-0.843	0.041
$(d-p)$	-	-	-	-	-	15.768	-0.277	0.098
spr	-	-	-	-	-	-	1.047	-0.047
(qk)	-	-	-	-	-	-	-	2.617

Table D.3: VAR(1) coefficients with relative t -statistics and Cross-Correlations of Residuals. **Note:** rtb_t = ex post real T-Bill rate, xr_t = excess stock return, xb_t = excess bond return, $x\Delta p_t$ = log difference on annuities premium $(d-p)_t$ = log dividend-price ratio, y_t = nominal T-bill yield, spr_t = yield spread, qk = aggregate longevity shock.

<i>HedgingPortfolio</i>	w_{xr}	w_{rtb}	w_{xb}	$w_{x\Delta pr}$
<i>Unconstrained</i>	-0.042	-0.014	0.009	0.047
<i>Constrained_1_</i> $w_{xr} = 0$	0	-0.062	0.004	0.058
<i>Constrained_2_</i> $w_{rtb} = w_{xr} = 0$	0	0	-0.059	0.059

Table D.4: Optimal Allocation for Aggregate Longevity Hedging Portfolio under different constraints on the set of financial securities included in the hedging portfolio.

SR_{∞}^{Equity}	$SR_{\infty}^{LongBond}$	SR_{∞}^{ALS}
0.429	0.0323	(-) 0.437

Table D.5: Long term Sharpe Ratios for securities included in the Extended VAR. Values are computed as illustrated in eq.9.

SR_{∞}^{Unc}	$SR_{\infty}^{Cons.1}$	$SR_{\infty}^{Cons.2}$
(-) 0.504	(-) 0.421	(-) 0.333

Table D.6: Long term Sharpe Ratios for Aggregate Longevity Hedging Portfolios. Values are computed as illustrated in eq.9.