Observatoire de l'Epargne Européenne

Can Risk Be Shared Across Investor Cohorts? Evidence from a Popular Savings Product

September 2019

Johan Hombert[1], Victor Lyonnet[2]

1 HEC Paris, CEPR
2 The Ohio State University

74 Avenue Paul Doumer • 75116 Paris • France • +33 (0) 9 72 62 91 58
contact@oee.fr • www.oee.fr • Siret: 424 667 947 00024

Can Risk Be Shared Across Investor Cohorts?

Evidence from a Popular Savings Product*

Johan Hombert† Victor Lyonnet‡

September 9, 2019

Abstract

This paper shows how one of the most popular savings products in Europe – life insurance financial products – shares market risk across investor cohorts. Insurers smooth returns by varying reserves that offset fluctuations in asset returns. Reserves are passed on between successive investor cohorts, causing redistribution across cohorts. Using regulatory and survey data on the 1.4 trillion euro French market, we estimate this redistribution to be quantitatively large: 1.4% of savings value per year on average, or 0.8% of GDP. These findings challenge a large theoretical literature that assumes inter-cohort risk sharing is impossible. We develop and provide evidence for a model in which the elasticity of investor demand to predictable returns determines the amount of risk sharing that is possible. The evidence is consistent with low elasticity, sustaining inter-cohort risk sharing despite predictable returns. Demand elasticity is higher for investors with a larger investment amount, suggesting that low investor sophistication enables inter-cohort risk sharing.

*A preliminary version of this paper circulated under the title "Intergenerational Risk Sharing in Life Insurance: Evidence from France." We thank Joseph Briggs, Anne-Laure Delatte, Christian Gollier, Valentin Haddad, Anastasia Kartasheva, Robert Novy-Marx, Ishita Sen, Boris Vallée, conference participants at the 2017 NBER Insurance Project Workshop, European Winter Finance Conference, FIRS, CSEF-IGIER Conference, SFS Cavalcade, Paul Woolley Annual Conference, Finance Meets Insurance Conference, EFA, seminar participants at Autorité de Contrôle Prudentiel et de Résolution, Chinese University of Hong Kong, ESCP Europe, Hong Kong University, Hong Kong University of Science and Technology, Singapore Management University, Nanyang Technological University, Deutsche Bundesbank, University of Toronto, and MIT-Sloan for helpful comments. We thank Autorité de Contrôle Prudentiel et de Résolution and in particular we thank Frédéric Ahado, Anne-Lise Bontemps-Chanel, Fabrice Borel-Mathurin, Charles-Henri Carlier, Edouard Chrétien, Pierre-Emmanuel Darpeix, Olivier DeBandt, Dominique Durant, Henri Fraisse, Samuel Slama, for access and assistance in putting together the data as well as comments. All errors remain our own.

†HEC Paris and CEPR, hombert@hec.fr

‡The Ohio State University, lyonnet.1@osu.edu

1 Introduction

Even in well-developed financial markets, aggregate risk can only be shared among investors participating in the market when it is realized. This limit to risk sharing sometimes results in significant losses: In 2008, a perfectly diversified portfolio of stocks lost 40% of its value. Superior risk sharing can be achieved by diversifying risk intertemporally across investor cohorts (Gordon and Varian, 1988), but financial markets do not allow current and future investor cohorts to trade with each other (i.e., financial markets are incomplete).[1] In principle, long-lived intermediaries can complete the market by transferring risk between successive cohorts. However, Allen and Gale (1997) show that inter-cohort risk sharing implemented by a financial intermediary unravels in the presence of perfect competition in the savings market. This paper explores, both theoretically and empirically, how inter-cohort risk sharing can be achieved when competition in the savings market is not perfect.

The first contribution of this paper is to show how one of the most popular savings products in Europe shares market risk across investor cohorts. These products are sold by life insurers to retail investors.[2] We focus on the 1.4-trillion-euro French market, where they are called *euro contracts* and are pure savings products (i.e., they are not traditional life insurance products). Euro contracts work as follows. When a retail investor buys a contract, an account is created, on which she can invest and withdraw cash at any time. In turn, each insurer pools the cash deposited by all its investors into a single fund invested in a portfolio of assets.

Funds hold reserves that vary to offset shocks to asset returns. Reserves increase when asset returns are high and decrease when asset returns are low, so that euro contract returns are an order of magnitude less volatile than funds asset returns. Crucially, reserves belong collectively to investors and are passed on between successive investor cohorts. Therefore, they are shared across investors from *all* cohorts, causing redistribution across cohorts.

Investors receive a transfer from reserves when asset returns are low, and contribute to reserves when asset returns are high. Part of these transfers net out within investors' holding period. The *net* transfer received from or contributed to reserves over investors' holding period represents inter-investor-cohort redistribution. Consider the following illustrative example. There are three periods. Investor A invests 100 in a euro contract in period 1 only, and investor B invests 100 in the same

[1] Gollier (2008) estimates that inter-cohort risk sharing increases the certainty equivalent of capital income by 25% relative to an economy without inter-cohort risk sharing.

[2] Their name varies by country: "euro contracts," "participating contracts," and so on. As of 2017, these products represent 60% of life insurers' provisions in Europe (source: statistics from EIOPA). In France, as of 2015, these products represented 80% of life insurers' provisions, and 30% of total household financial wealth.

contract in periods 2 and 3. The asset return is 5% in periods 1 and 2, and minus 1% in period 3. The euro contract return is 3% in every period. Therefore, 2 are contributed to reserves in periods 1 and 2, and 4 are distributed from reserves in period 3.[3] Investor A receives a net transfer of minus 2. Investor B receives a net transfer of 2 over two periods, that is, 1 per period. In this example, the average amount of inter-investor-cohort redistribution is equal to $(|-2|+|1|+|1|)/3 \simeq 1.3$ per period. Using regulatory and survey data from France, we estimate that inter-cohort redistribution amounts to 1.4% of total account value per year, which represents 17 billion euros redistributed across investor cohorts every year, or 0.8% of GDP.

These findings challenge a large theoretical literature that assumes inter-cohort risk sharing is impossible. Indeed, a two-sided commitment problem must be overcome to allow for inter-cohort risk sharing. First, insurers must credibly commit not to run away with reserves, which they might be tempted to do when the level of reserves is high. Regulation solves this side of the commitment problem (see Section 3.1), ensuring reserves are owed and distributed to investors. Second, investors must remain invested in contracts even when reserves are low. The existing literature has emphasized how this side of the commitment problem prevents inter-cohort risk sharing. Allen and Gale (1997) study savings contracts that share market risk across investor cohorts through a reserve mechanism similar to that of euro contracts. They show that (a) a financial intermediary can implement inter-cohort risk sharing if it is protected from competition, that is, if investors must invest with the intermediary regardless of the reserves level, but (b) inter-cohort risk sharing would unravel under perfect competition, because investors then opportunistically opt out of the contracts after negative shocks to asset returns.

How can inter-cohort risk sharing then be sustained in euro contracts? The assumptions made in the existing literature do not fit the institutional framework of these contracts. First, euro contracts are offered by multiple intermediaries competing with each other as well as with alternative investment options so that (a) does not apply. Second, the large amount of inter-cohort redistribution we observe in the data rules out the assumption of perfect competition in the savings market that would imply (b). To our knowledge, no theoretical framework of inter-cohort risk sharing exists when competition in the savings market is imperfect. The second contribution of this paper is to offer one.

We develop a model in which long-lived intermediaries compete in selling savings products

[3] In this example, we ignore the fact that reserves are invested in assets that generate returns. We also consider non-overlapping investors. The method we use in Section 3.4 to measure inter-cohort redistribution accounts for these features of real-world contracts.

to successive cohorts of investors. We characterize how the amount of inter-cohort risk sharing depends on investor demand elasticity, that is, on the degree of competition. The model nests the two polar cases of perfectly inelastic demand (no competition) and perfectly elastic demand (perfect competition) that have been studied in the literature. In line with this literature, we show asset risk can be perfectly shared across investor cohorts when demand is inelastic to predictable contract returns. Instead, when demand is elastic, investors behave opportunistically and exploit the predictability of contract returns: They flow into (out of) contracts when reserves are high (low), partially unravelling risk sharing across cohorts. In the limit where demand is perfectly elastic, inter-cohort risk sharing fully unravels so that the savings products are akin to pass-through mutual funds (as in Allen and Gale, 1997). In a nutshell, the equilibrium level of inter-cohort risk sharing crucially (and monotonically) depends on demand elasticity.

Our model shows we can estimate demand elasticity from two moments in our data. The first moment is the regression coefficient of contract return on contemporaneous asset return, conditional on the level of reserves. When demand is inelastic, euro contracts share risk across investor cohorts. In this case, the contract return depends on the level of reserves but not on contemporaneous asset return beyond its effect on reserves. The intuition is similar to that of the permanent income hypothesis, whereby optimal consumption does not depend on current income beyond its effect on permanent income. By contrast, when demand is elastic, little inter-cohort risk sharing occurs and the contract return depends strongly on the contemporaneous asset return. We estimate panel regressions and, controlling for the level of reserves, we show the contract return does not depend on the asset return in the current year. Therefore, the evidence is consistent with low demand elasticity.

The second moment that is informative about demand elasticity is the regression coefficient of investor flows on reserves. A high level of reserves predicts high expected contract returns, so that the sensitivity of flows to reserves is directly related to the elasticity of demand to expected returns. We run panel regressions and find the sensitivity of flows to reserves is a precisely estimated zero, again consistent with low demand elasticity. One issue when regressing flows on reserves is that reserves are potentially endogenous to unobserved demand shocks—a standard issue when one estimates demand functions by regressing quantity on price. Our model shows the past asset return is a valid instrument for reserves to estimate the sensitivity of flow to reserves. Instrumenting reserves, the sensitivity of flows to reserves is again a precisely estimated zero.

Why are flows inelastic to expected returns, allowing for inter-cohort risk sharing? We rule out explanations based on taxes distorting investor choices. In particular, we study investors buying a new contract. New investors do not face any tax distortion. Despite no tax distortion and the fact that high reserves predict high contract returns, we find new investors' flows are not significantly correlated with reserves.

We hypothesize that investor demand is inelastic to reserves because investors lack the knowledge to predict contract returns using reserves. In line with this hypothesis, we show that contracts held by investors with a small investment amount (below €250,000) have a flow-reserves sensitivity indistinguishable from zero, whereas contracts with a large investment amount (above €250,000) exhibit a positive and statistically significant flow-reserves sensitivity. However, the economic magnitude of this sensitivity remains small. This result is consistent with interpreting the investment amount as a proxy for wealth and financial sophistication, whereby less sophisticated investors fail to predict contract returns using reserves.[4] Differences in demand elasticity across investors can arise if, for instance, investors must incur a fixed cost to acquire the knowledge or information necessary to understand the sources of predictability (Lusardi and Mitchell, 2014).

Previous research has analyzed other arrangements implemented by financial intermediaries to share aggregate risk.

Another way to provide households with insurance against aggregate market risk is to share this risk cross-sectionally between the financial intermediary and households. The intermediary bears part of the risk while (partially) hedging households' returns from market risk. Examples of savings products implementing cross-sectional risk sharing include guaranteed variable annuities sold by US life insurers (Koijen and Yogo, 2015) and structured products sold by European banks (Célérier and Vallée, 2017). Cross-sectional risk sharing is also at play in euro contracts, but we show it is an order of magnitude smaller than inter-cohort risk sharing. Crucially, cross-sectional risk sharing hinges on intermediaries' risk-bearing capacity—typically their capital position (Koijen and Yogo, 2018)—and creates insolvency risk. By contrast, euro contracts shift most of the risk to households and share it across cohorts.

Similar to euro contracts, defined benefits (DB) pension plans contain an element of inter-cohort risk sharing, because DB sponsors have the option to increase the contributions of futures employees or may be bailed out by future taxpayers (Novy-Marx and Rauh, 2011, 2014). One important difference, however, is that DB sponsors fully commit to a rate of return for households, whereas insurers

[4]Using data from a French life insurer, Bianchi (2018) studies households' portfolio allocation between mutual funds and euro contracts. He constructs a survey-based measure of financial literacy and shows this measure is highly correlated with household wealth.

selling euro contracts do not commit to a pre-defined contract return, thus retaining substantial flexibility to spread shocks to asset returns across cohorts. The implications of the behavior of insurance companies and pension funds for asset prices is studied by Greenwood and Vissing-Jorgensen (2018), and the implications for the structure of the financial system by Scharfstein (2018).

Our paper also adds to the recent literature on consumer inertia in insurance and banking markets. Handel (2013) shows consumer health plan choice inertia can be beneficial because it reduces adverse selection, hence allowing for better (cross-sectional) risk sharing. In our paper, investor inertia is beneficial because it mitigates the free-riding problem that investors would only invest when other investors have accumulated enough reserves, allowing for better (inter-cohort) risk sharing. Drechsler, Savov, and Schnabl (2017, 2018) show household deposits are insensitive to the spread between the deposit rate and the Fed funds rate, enabling banks to pass on changes in short-term interest rates to households, thereby partially immunizing banks from interest rate risk.

We also contribute to the theoretical literature on the private implementation of inter-cohort risk sharing. The notion that financial markets cannot implement inter-cohort risk sharing because they do not allow current and future investor cohorts to trade with each other goes back at least to Stiglitz (1983) and Gordon and Varian (1988), whereas Ball and Mankiw (2007) study inter-cohort risk sharing in a hypothetical economy in which current investors could trade with future investors. Allen and Gale (1997) and Gollier (2008) study how inter-cohort risk sharing can be implemented by an intermediary having monopoly power over households' savings. Crucially, Allen and Gale (1997) show inter-cohort risk sharing unravels in the presence of perfect competition in the savings market. We extend this literature by considering the case of imperfect competition in the savings market, showing the equilibrium level of inter-cohort risk sharing decreases monotonically from perfect to nonexistent as competition increases from zero to perfect.

The rest of the paper is organized as follows. We present the model in Section 2. We document and quantify inter-cohort redistribution in euro contracts in Section 3, and in Section 4, we study why it does not unravel. Section 5 concludes. Proofs and additional material are provided in the appendix.

2 Model

2.1 Setup

Every period $t = 1, 2, \ldots, +\infty$, a mass one of investors are born, live for one period, and have one unit of wealth to invest over that period. $J \geq 1$ long-lived intermediaries, indexed by $j = 1, \ldots, J$, offer one-period saving contracts every period.[5] The contract offered by intermediary j in period t promises a return $y_{j,t}$ contingent on all information observable at the end of period t.

At the beginning of period t, each intermediary j has reserves $R_{j,t-1}$ and collects $V_{j,t-1}$ from investors. We refer to $V_{j,t-1}$ as investors' account value. The intermediary has total assets $V_{j,t-1} + R_{j,t-1}$, which generate an exogenous return $x_{j,t}$ with $E_{t-1}[x_{j,t}] = r$, where E_{t-1} denotes expectation conditional on information at the beginning of period t. Asset risk may include both a systematic component and an idiosyncratic component determined by the covariance structure of $x_t \equiv (x_{0,t}, \ldots, x_{J,t})$, where $j = 0$ defines investors' outside option described below.

Intermediary j's budget constraint in period t is given by:

$$x_{j,t}(V_{j,t-1} + R_{j,t-1}) = y_{j,t}V_{j,t-1} + \Pi_{j,t} + (R_{j,t} - R_{j,t-1}), \tag{1}$$

where $\Pi_{j,t}$ is the intermediary profit. The budget constraint describes how asset income (on the left-hand side) is split between current investors (first term on the right-hand side), the intermediary (second term), and reserves (third term). We normalize initial reserves $R_{j,0}$ to zero. To rule out Ponzi schemes, reserves must satisfy the transversality condition

$$\lim_{t \to +\infty} \frac{R_{j,t}}{(1+r)^t} \geq 0. \tag{2}$$

Reserves are used to redistribute wealth across investor cohorts. As discussed in the introduction, inter-cohort risk sharing requires overcoming a two-sided commitment problem: Intermediaries must not channel funds from the reserves to their profits, and investors must invest with the intermediary even when reserves are low. A natural solution to the commitment problem on the intermediaries' side is regulation. We assume a specific regulation that captures the main features of the actual regulatory framework in several European countries (including France), and leave the

[5] Intermediaries cannot offer multi-period contracts, because investors only live for one period. An equivalent interpretation is that investors live for several periods, have additively time-separable utility as given by Equation (5), and intermediaries are restricted to offer one-period contracts (e.g., by regulation). This interpretation fits best our empirical framework.

issue of optimal regulation for future research. Specifically, we assume regulation imposes that intermediaries' profit is equal to a fraction $\phi \in (0, 1)$ of investors' account value:[6]

$$\Pi_{j,t} = \phi \, V_{j,t-1}. \tag{3}$$

The regulatory constraint (3) ensures intermediaries cannot appropriate the reserves, implying reserves are owed to (current and future) investors. Thus, the change in the level of reserves $(R_{j,t} - R_{j,t-1})$ in (1) represents the payoff for past and future investor cohorts. Indeed, past investor cohorts have contributed $R_{j,t-1}$, and current reserves $R_{j,t}$ will be distributed to future investor cohorts. This observation has two implications regarding risk sharing.

First, the budget constraint (1) highlights that both cross-sectional risk sharing and inter-cohort risk sharing can be at play. Current investors may share asset risk with the intermediary (cross-sectional risk sharing) and with the reserves, that is, with past and future investor cohorts (inter-cohort risk sharing). The possibility of inter-cohort risk sharing stands in contrast to structured savings products that rely solely on cross-sectional risk sharing between investors and intermediaries, such as those sold by banks in Europe (see Célérier and Vallée, 2017) and those sold by life insurers is the US (see Koijen and Yogo, 2018).

Second, the regulatory constraint (3) rules out cross-sectional risk sharing between the intermediary and investors, because the intermediary's profit is constrained to be proportional to account value. Our focus is thus on inter-cohort risk sharing, which is the economically relevant dimension of risk sharing in our empirical analysis (see Section 3.3). In line with this focus, we assume intermediaries are risk neutral. Intermediaries maximize expected profit discounted at the expected rate of asset return

$$\sum_{t=1}^{+\infty} \frac{E_0[\Pi_{j,t}]}{(1+r)^t}. \tag{4}$$

We model investor demand for contracts using a multinomial logit model. Investor i from cohort t investing with intermediary j obtains utility

$$\alpha \, u(y_{j,t}) + \xi_j + \psi_{i,j,t-1}. \tag{5}$$

[6] The actual regulation in our empirical setup is that intermediaries' profit cannot exceed a fraction of asset income (see Section 3.1). We make two simplifying assumptions relative to actual regulation. First, we assume a constant coefficient ϕ, whereas in practice, the coefficient is equal to a fraction of the realized asset return. ϕ can thus be interpreted as that fraction of the average asset return if the intermediary is risk neutral, which we assume below. Second, Equation (3) is written with "=", whereas the actual regulatory constraint requires "$\leq$". In Appendix A.3, we derive a sufficient condition for the regulatory constraint to be binding and argue this condition is likely satisfied in our empirical setup.

$\alpha\, u(y_{j,t})$ is the indirect utility provided by contract return $y_{j,t}$, where $\alpha > 0$ parameterizes the sensitivity of investor utility to return, $u' > 0$, $u'' < 0$, and w.l.o.g. we normalize $u'(r - \phi) = 1$. ξ_j is preference for intermediary j shared across all investors.[7] $\psi_{i,j,t-1}$ is investor i's idiosyncratic preference for intermediary j. $\psi_{i,j,t-1}$ is indexed by $t-1$ to reflect the fact that it is realized at the beginning of period t. It is distributed i.i.d. extreme value across investors in cohort t.

Investors also have access to an outside investment opportunity indexed by $j = 0$, which yields utility given by (5) with ξ_0 normalized to zero, $\psi_{i,0,t-1}$ distributed i.i.d. extreme value, and return $y_{0,t} = x_{0,t} - \phi_0$. $\phi_0 > 0$ captures fees and other costs of investing in the outside investment opportunity. The outside investment opportunity can be thought of as investment in mutual funds or direct investment in financial markets. We assume the cost of investing through the outside option is the same as the cost of investing through intermediaries, that is, $\phi_0 = \phi$.[8]

Contract return in period t is contingent on all observable information at the end of period t, which includes the history of asset returns of all intermediaries. Thus, $y_{j,t}$ is a function of $x^t \equiv (x_1, \ldots, x_t)$. Each investor i buys the contract that provides them with the largest expected utility, $j \in \arg\max_{k=0,1,\ldots,J} \alpha\, u(y_{k,t}) + \xi_k + \psi_{i,k,t-1}$, which yields the logit demand function:

$$V_{j,t-1} = \frac{\exp\{\alpha\, E_{t-1}[u(y_{j,t})] + \xi_j\}}{\sum_{k=0}^{J} \exp\{\alpha\, E_{t-1}[u(y_{k,t})] + \xi_k\}}. \tag{6}$$

The problem of an intermediary is to maximize profit (4) by choice of a contract return policy subject to the budget constraint (1), transversality condition (2), profit function (3), and demand function (6). Each intermediary takes other intermediaries' contract return policies as given. An equilibrium is a fixed point of this problem.[9]

Finding a general analytical solution to this problem is difficult. To simplify the problem and obtain an explicit solution, we solve the model using a first-order approximation.[10] We assume asset return shocks have bounded support, that is, $|x_{j,t} - r| \le \sigma$ for all j and t, and that, for some

[7] To streamline the exposition, we present the case with time-invariant ξ_j in the main text. All proofs in the appendix are derived with time-varying ξ_j. We discuss the impact of time-varying ξ_j at the end of Section 2.3.

[8] In several countries, including France, life insurers sell mutual funds through unit-linked contracts that are subject to the same fee structure and tax treatment as euro contracts. In such cases, $\phi_0 = \phi$ by design.

[9] Because we do not clear the capital market, our model is in partial equilibrium. The model is equivalent to a general equilibrium model with constant returns to capital as in Allen and Gale (1997) and Ball and Mankiw (2007). Suppose each intermediary j can lend capital to competitive firms using a linear production function $Y_{j,t} = (1 + x_{j,t})K_{j,t-1}$, where $Y_{j,t}$ is output and $K_{j,t-1}$ is capital. In such an economy, an increase in reserves leads to an increase in the aggregate capital stock. An alternative interpretation of our model is that of a small open economy, in which case an increase in reserves leads to a capital account deficit.

[10] The advantage of using a first-order approximation is that it eliminates any possible interaction between the shocks occurring in different periods. Ball and Mankiw (2007) use a similar method to solve the complete-market equilibrium in which investors are allowed to trade with future investor cohorts.

period T, $x_{j,t} = r$ for $t > T$. The value of T can be any positive integer, so that our analysis covers any finite number of shocks, however large.

2.2 Equilibrium

To obtain an explicit solution for contract returns, we use a first-order approximation that is valid as long as asset return shocks, σ, are small. That is, we derive the equilibrium when fluctuations in asset returns are small.

Proposition 1. *Contract return of intermediary j in period t is given by*

$$y_{j,t} \simeq r - \phi + \sum_{s=1}^{t} \beta_{j,t}(s)\,(x_{j,s} - r) + f_{j,t}(\overline{x}^t - r), \tag{7}$$

where

$$\beta_{j,t}(s) = \frac{\gamma_j}{\alpha + \frac{1+r}{r}\gamma_j} \qquad \text{for } s < t, \tag{8}$$

$$\beta_{j,t}(t) = \frac{\alpha + \gamma_j}{\alpha + \frac{1+r}{r}\gamma_j}, \tag{9}$$

$\gamma_j > 0$ is a constant independent of α, and $f_{j,t}(.)$ is a function of the history of weighted-average asset return shocks $\overline{x}^t - r$. Closed-form expressions for these variables are in Appendix A.2.

Equation (7) shows the contract return is equal to the expected asset return, r, minus the compensation of the intermediary, ϕ, plus a function of the history of shocks to the intermediary's asset return, $x_{j,s} - r$, and a function of the history of average asset returns, $\overline{x}^t$. The key coefficients are the $\beta_{j,t}(s)$, which pin down the extent of risk sharing across investor cohorts. $\beta_{j,t}(s)$ measures the sensitivity of period-t contract return, $y_{j,t}$, to period-s asset return, $x_{j,s}$. When $\beta_{j,t}(s) > 0$, the period-t investor cohort bears some of period-s asset risk.

The contract return policy (7) implies period-s asset risk is shared between the current (period-s) cohort (because $\beta_{j,s}(s) > 0$) and all future cohorts (because $\beta_{j,t}(s) > 0$ for $t > s$). In turn, Equations (8) and (9) show the extent of inter-cohort risk sharing depends on the elasticity of investor demand to contract return, α. When demand is inelastic ($\alpha \simeq 0$), $\beta_{j,s}(s) = \beta_{j,t}(s)$ for all $t > s$ so that asset risk is perfectly shared between the current and future cohorts. When demand is elastic ($\alpha > 0$), $\beta_{j,s}(s) > \beta_{j,t}(s)$ for $t > s$ and more asset risk $x_{j,s}$ is shifted to the contemporaneous (period-s) cohort. As a result, asset risk is imperfectly shared across investor cohorts when demand is elastic.

The intuition behind Proposition 1 is that when demand is elastic, future investor cohorts behave opportunistically by investing more (less) when reserves are higher (lower). For instance, when the asset return is high, the intermediary would like to share in gains with future investor cohorts by hoarding part of the return as reserves. When demand is elastic, however, future investors flow in, diluting reserves and undoing the sharing of gains. Conversely, when the asset return is low, the intermediary would like to share in losses with future investor cohorts by tapping reserves and replenishing them in future periods. In this case, future investors flow out, preventing the intermediary from replenishing reserves and undoing the sharing of losses. In the limit case where demand is perfectly elastic ($\alpha \simeq \infty$), inter-cohort risk sharing unravels completely: $\beta_{j,s}(s) = 1$ and $\beta_{j,t}(s) = 0$ for $t > s$.

We denote by R_{j,t^-} the level of reserves at the end of period t just before distribution to investors. It is equal to beginning-of-period reserves plus asset income:

$$R_{j,t^-} = R_{j,t-1} + x_{j,t}(V_{j,t-1} + R_{j,t-1}). \tag{10}$$

We also denote by $\mathcal{R}_{j,t^-} \equiv R_{j,t^-}/V_{j,t-1}$ the reserve ratio to total account value. Our next result is that period-t contract return given by (7) depends on the history of past asset returns, $\overline{x}^{t-1}$, *only through* its effect on the reserve ratio.

Proposition 2. *Contract return of intermediary j in period t is given by*

$$y_{j,t} \;\; \simeq \;\; r - \phi + \frac{1}{1+r} \frac{\alpha}{\alpha + \frac{1+r}{r}\gamma_j} (x_{j,t} - r) + \frac{r}{1+r} (\mathcal{R}_{j,t^-} - r) + \mu_j \, (\overline{x}_t - r), \tag{11}$$

where $\gamma_j > 0$ is a constant independent of α, $\mu_j < 0$ goes to zero when α goes to zero or infinity, and $\overline{x}_t$ is a weighted average of $x_{k,t}$ over $k = 1, \ldots, J$.

Proposition 2 shows how the share of asset risk borne by current investors depends on the elasticity of demand, α. When demand is inelastic ($\alpha \simeq 0$), the coefficient in front of $x_{j,t}$ in (11) is equal to zero. The contract return then does not depend on the current asset return beyond its effect on the reserve ratio; that is, asset risk is perfectly shared across investor cohorts. When demand is elastic ($\alpha > 0$), the coefficient in front of $x_{j,t}$ is strictly positive. The intermediary then shifts more asset risk to the current cohort. In this case, the contract return depends on the current asset return above and beyond its effect on the end-of-period reserve ratio; that is, asset risk is imperfectly shared across investor cohorts.

One important implication of Proposition 2 is that the reserve ratio $\mathcal{R}_{j,t^-}$ is a sufficient statistic for the history of shocks. All that matters for setting the contract return is the current reserves ratio, not the path leading to that ratio. Indeed, the contract return in (11) does not depend on past shocks beyond their effect on the reserve ratio. The sensitivity of the contract return to the reserve ratio results from the following tradeoff faced by the intermediary. On the one hand, paying out a larger fraction of reserves to current investors leads to higher demand and thus higher profit in the current period. On the other hand, tapping reserves today implies paying lower returns to future investors, lowering future demand and hence future profit. The optimal choice is to pay out to current investors a fraction of reserves equal to their weight in intertemporal profit, equal to $1/\sum_{\tau=t}^{+\infty} \frac{1}{(1+r)^{\tau-t}} = \frac{r}{1+r}$.

Proposition 2 also implies an intermediary's contract return depends negatively on other intermediaries' asset returns, because $\mu_j < 0$. Intuitively, when other intermediaries have high asset returns, they increase contract returns both in the current period and in future periods, which reduces intermediary j's future demand, but not its current-period demand, because it is realized before asset returns. Intermediary j's optimal response is then to increase future contract returns to avoid losing too large future market shares,[11] by lowering the current contract return. This effect vanishes when demand is perfectly inelastic ($\alpha \simeq 0$), because intermediary j then has no incentives to react; and it vanishes when demand is infinitely elastic ($\alpha \simeq \infty$), because other intermediaries then do not change the future contract return in response to asset return shocks.

2.3 Empirical implications

We now show that two relations that can be estimated in the data are directly informative about the elasticity of demand, α, which in turn is the key determinant of equilibrium risk sharing.

The first relation is the contract return policy. The coefficients γ_j and μ_j in (11) are intermediary-specific, because the optimal contract return depends on the elasticity of demand, itself a function of the intermediary's market share due to logit demand. A closer inspection of the coefficients (reported in the appendix) reveals they only depend on market shares up to second-order terms. When market shares are not too large, equilibrium contract returns can be approximated as follows:

Implication 1 (contract return policy). *For small market shares, the period-t contract return*

[11] This best-response reflects the strategic complementarity property of logit demand, that is, the property whereby contract return best-response functions are increasing in other intermediaries' contract returns.

of intermediary j is

$$y_{j,t} \simeq cste + \frac{1}{1+r} \frac{\alpha}{\alpha + \frac{1+r}{r}\gamma} x_{j,t} + \frac{r}{1+r} \mathcal{R}_{j,t^-} + \varepsilon_{j,t}, \tag{12}$$

where $\gamma = \frac{-u''(r-\phi)}{u'(r-\phi)} > 0$ *is the coefficient of absolute risk aversion, and the cross-sectional covariance between* $\varepsilon_{j,t}$ *and* $\left(x_{j,t}, \mathcal{R}_{j,t^-}\right)$ *is approximately zero.*

Implication 1 allows the coefficients in the equilibrium contract return policy (12) to be estimated by running a linear regression with time fixed effects in a panel of intermediaries. Implication 1 also implies the model can be easily rejected, because it predicts the coefficient in front of the reserve ratio should be commensurate with the interest rate. The coefficient in front of the current asset return is informative about α.

The second relation is the flow-reserves relationship. Our next result is that if demand is elastic to contract return, that is, if $\alpha > 0$, flows depend on the beginning-of-period reserve ratio $\mathcal{R}_{j,t-1} \equiv R_{j,t-1}/V_{j,t-1}$:

Implication 2 (flow-reserves relation). *For small market shares, net flows to intermediary j in period t are given by*

$$\log(V_{j,t-1}) \simeq \nu_j + \alpha\, r\, \mathcal{R}_{j,t-1}, \tag{13}$$

where ν_j *is a constant independent of* α*.*

We know from Proposition 2 that the contract return paid at the end of period t depends on the end-of-period reserve ratio, itself determined by the beginning-of-period reserve ratio. Therefore, the period-t contract return is predicted by the reserve ratio at the beginning of period t. Correspondingly, Implication 2 states that if investors are elastic to contract returns ($\alpha > 0$), investor demand in period t depends on the reserve ratio at the beginning of period t. The sensitivity of investor demand to the beginning-of-period reserve ratio in (13) is equal to the product of the sensitivity of investor demand to expected contract return (equal to α) by the sensitivity of expected contract return to the beginning-of-period reserve ratio (equal to r).

The coefficient $\alpha\, r$ in the flow-reserves relation (13) can be estimated by running a linear regression. The OLS estimate is unbiased because we have assumed the absence of any demand shocks; that is, ξ_j is time-invariant, which ensures no error term exists in (13). In Appendix A.7, we show that introducing demand shocks (time-varying ξ_j) introduces an error term in (13) that is negatively correlated with $\mathcal{R}_{j,t-1}$. Intuitively, when the intermediary anticipates a negative demand shock, it optimally increases reserves to increase future contract returns and lean against the demand shock, generating a spurious negative correlation between reserves and demand. This correlation creates a downward bias in the OLS estimate of $\alpha\, r$. We show this bias can be corrected by instrumenting $\mathcal{R}_{j,t-1}$ using lagged asset returns. Intuitively, lagged asset returns affect reserves because a fraction of asset returns are hoarded as reserves (relevance condition), but they are not directly correlated with demand shocks beyond their effect on reserves (exclusion restriction).

2.4 Extension: Arbitrageurs

Is inter-cohort risk sharing robust to the presence of arbitrageurs? Suppose $\alpha > 0$, such that inter-cohort risk sharing can be sustained in equilibrium. Suppose further that there exist arbitrageurs who can take long positions in intermediaries' contracts, and long and short positions in the same assets as intermediaries and in a risk-free asset. We assume the risk-free rate is below the expected return on intermediaries' assets, $r_f < r$, that is, intermediaries earn a positive risk premium on their investments. Arbitrageurs implement risk-free arbitrage strategies, if such strategies exist.

Let us determine whether arbitrage strategies exist. Proposition 2 shows that contract returns can be replicated with a portfolio composed of the intermediaries' assets and the risk-free asset. Therefore, if an arbitrage strategy exists, it would consist in long positions in euro contracts, hedged with short positions in intermediaries' assets and some position in the risk-free asset. As in common in most countries, including France, we assume interest expenses are not tax deductible for households.[12] Therefore, the return on the long leg of the arbitrage strategy is taxable, whereas the return paid on the short leg is not tax deductible. We denote the capital income tax rate by τ.[13]

We show in Appendix A.8 that one euro invested long in contract j hedged by short positions in the underlying assets generates an arbitrage profit

$$\pi_{j,t}^{arb} \simeq \left[1 - (1 - \tau)\left(\frac{\alpha + \gamma_j}{\alpha + \frac{1+r}{r}\gamma_j} + \mu_j\right)\right](r - r_f) + (1 - \tau)r\,\mathcal{R}_{j,t-1} - \tau r - (1 - \tau)\phi. \quad (14)$$

[12] In some countries, including France and the US, interest paid on mortgages, student loans and business loans often are tax deductible, but interest expenses in levered financial investments usually are not. In line with the institutional framework in Europe, we assume contracts can only be purchased by households, so the relevant tax regime is that of households.

[13] Investor utility (5) accounts for the capital income tax, even though τ does not appear in (5), because the indirect utility function $u(.)$ is defined over before-tax returns.

Equation (14) highlights two distinct sources of arbitrage profits. First, contracts are partially hedged against asset risk, yet they earn the risk premium. Indeed, an arbitrageur going long the contract and short the underlying assets earns part of the risk premium without bearing the associated risk. This source of arbitrage profits is reflected in the first term of (14): The term in brackets is equal to one minus the exposure of the after-tax contract return to asset risk, and is positive because contract returns are partially hedged against asset risk; and $r - r_f > 0$ is the risk premium. Thereby, the arbitrageur extracts some of the welfare surplus created by inter-cohort risk sharing. The second source of arbitrage profits comes from the predictable distribution of reserves to contract holders. It is reflected in the second term of (14), which is proportional to the beginning-of-period reserve ratio. The costs of the arbitrage strategy are the tax on the expected asset return (third term of (14)) and the compensation of the intermediary (fourth term). In the absence of taxes or fees, arbitrage opportunities exist, and inter-cohort risk sharing unravels as in Allen and Gale (1997).

The key insight is that a capital income tax is sufficient to eliminate arbitrage opportunities; that is, $\pi_{j,t}^{arb} < 0$ if τ is large enough. This result does not rely on risk-sharing contracts benefiting from a tax advantage, nor from any form of tax distortion relative to other investments, because we assumed returns on all long positions are taxed at a uniform rate τ. The result does not rely either on contracts being expensive, because it holds even when ϕ is arbitrarily close to zero. In the empirical analysis, we will calibrate the terms in (14) to determine the capital income tax rate necessary to eliminate arbitrage opportunities, and we will compare it to the applicable tax rate.

3 Euro Contracts

European life insurers sell savings contracts designed to implement inter-cohort risk sharing. We study the market for these contracts in France, where they are called *euro contracts*. We present the institutional framework in Section 3.1, and the data and summary statistics in Section 3.2. We document the importance of reserve management in Section 3.3 and quantify inter-cohort redistribution in Section 3.4.

3.1 Institutional framework

Euro contracts account for one third of French households' financial wealth.[14] The size of the market for euro contracts was 1.4 trillion euros in 2015 (ACPR, 2016) out of the aggregate 4.5 trillion euros of household financial wealth. Euro contracts are sold by life insurers, which can be subsidiaries of insurance holding companies, subsidiaries of bank holding companies, or stand-alone life insurance companies.[15]

Despite being offered by life insurers, euro contracts are pure savings products, and do not entail insurance against longevity or mortality risk. When an investor buys a euro contract, she opens an account with the insurer on which she can deposit and withdraw cash at any time. Insurers usually charge entry fees when cash is deposited (front-end loads), and annual management fees, but are not allowed to charge exit fees. The insurer pools all the cash from all investors in a fund that is invested in a portfolio of assets.

At the end of each calendar year, each account is credited by an amount equal to the account value multiplied by a rate of return (*taux de revalorisation*), which we refer to as the contract return. The key feature of euro contracts is that the contract return can be different from the asset return. The difference between the asset return and contract return is used (or funded if negative) in two ways. One part is paid to the insurer, whereas the rest is credited to or debited from the fund's reserves, just as in Equation (1). The economic balance sheet of the fund is as follows.[16] Assets are equal to the market value of the asset portfolio. Liabilities are equal to total account value plus reserves. By definition, reserves are equal to the difference between total asset market value and total account value.[17]

The insurer has full discretion over the timing of contract returns, subject to the regulatory constraint that the insurer must distribute at least 85% of asset income to investors. The insurer can choose how much income from assets is credited immediately to investors' account (first term on the RHS of (1)), how much is credited to the insurer's equity (second term), and how much is credited to or debited from reserves (third term). Regulation imposes that the sum of the first

[14] The two other thirds are risky securities and investment funds on the one hand, and short-term instruments on the other hand (Insee, 2016).

[15] Mutual insurance companies, pension institutions, and reinsurance companies can also offer euro contracts. These institutions are subject to a different regulation and account for only 4% of aggregate provisions (ACPR, 2016). We abstract from them in the empirical analysis.

[16] The economic balance sheet differs from the accounting balance sheet because the former is marked-to-market, whereas life insurance accounting principles are mostly based on historical cost accounting.

[17] Three types of reserves exist, which we describe in detail in Appendix B. The decomposition of total reserves into these three categories reflects accounting, not economic differences. Thus, our analysis focuses on total reserves.

and third components must be at least 85% of asset income.[18] This regulation implies reserves are effectively owed to investors, because the regulatory constraint allows insurers to move funds between reserves and investors' accounts, but not from reserves to their own equity. As a result, insurers can choose the timing of contract returns by dynamically managing the level of reserves.

The key feature of reserves that gives rise to redistribution across investor cohorts is that reserves are pooled across all investors, rather than tied to individual investor accounts. In particular, new investors share in reserves accumulated by previous investors, and investors redeeming their contracts give up their share of reserves. The pooling of reserves across investor cohorts happens because all investors holding the same contract offered by a given insurer receive the same contract return regardless of when they entered into the contract.

Insurers often offer a range of contracts, for instance, a basic contract and a premium contract with a minimum investment amount and a lower fee rate. Insurers are allowed to pay different returns on different contracts. In principle, insurers could close existing contracts to new subscriptions when reserves are high, and create a new vintage of contracts to which they will pay different returns. Doing so would undo reserve pooling across investor cohorts. Using data at the contract level, we show in Section 4.1.1 that insurers do not do so; therefore, reserves are effectively pooled across investor cohorts.

Two other features of euro contracts are worth mentioning. First, regulation also imposes that insurers must distribute to investors at least 90% of their technical income if it is positive, or 100% if it is negative. Again, this amount can be paid to investors immediately by crediting investors' accounts, or later by crediting reserves. Technical income is equal to fees paid by investors minus the fund's operating costs. The implication of this regulation is that insurers cannot extract money from the reserves by raising fees on new investors, because 90% (or 100%) of these fees must eventually be returned to investors.

Second, euro contracts have a minimum guaranteed return fixed at the subscription of the contract. This rate has virtually been zero since the 1990s (Darpeix, 2016), and is typically not binding in our sample (see Section 3.2).

Overall, euro contracts closely resemble the theoretical contracts analyzed in Section 2. Both have a reserve mechanism allowing for smoothing of contract returns and inter-cohort risk sharing. In addition, euro contracts are liquid open-end investments whose return is decided ex post at the discretion of the insurer, which is equivalent to having one-period contracts as in the model. As

[18] See Appendix B for a detailed description of the regulatory framework for reserves.

discussed above, the only potential deviation from short-term contracts would happen if insurers strategically closed and opened contracts in order to pay different returns to different investor cohorts. We show in Section 4.1.1 that this does not happen in the data.

3.2 Data and summary statistics

Our main source of data comprises regulatory filings obtained from the national insurance supervisor (*Autorité de Contrôle Prudentiel et de Résolution*) for the years 1999 to 2015. The data cover all companies with life insurance operations in France and contains detailed financial statements.[19] We focus on stock insurance companies with more than 10 million euros of life insurance provisions. Because we need lagged values to calculate the change in reserves, the sample period of our analysis is 2000–2015. The final sample contains 76 insurers and 978 insurer-year observations.

Panel A of Table 1 reports summary statistics from the regulatory filings. The average (median) insurer has 13.9 (3.1) billion euros of account value. Inflows (premiums), which include cash deposited in newly opened contracts and in existing contracts, represent, on average, 10.5% of account value per year. Outflows, which include partial and full redemptions, either voluntarily or at contract termination (investor death), represent on average 8.1% of account value per year. The combination of positive net flows and compounded contract returns generates an increasing trend in aggregate account value plotted in Figure 1. Aggregate account value grows from 505 billion euros in 2000 to 1,200 billion euros in 2015 (all amounts are in constant 2015 euros). Aggregate growth reflects the internal growth of existing life insurers rather than the entry of new insurers. The number of insurers in the sample is 65 at the beginning of the period and 61 at the end. Market concentration is relatively low, with a Herfindahl-Hirschman Index around 800 and total market shares of the top five life insurers slightly below 50%.

The average reserve ratio is 10.9%. On the asset side, 80.4% of funds' portfolios are invested in sovereign and corporate bonds, 13.5% in stocks, and the rest in real estate, loans, and cash. The average asset return is 4.9% per year. The average contract return before fees is 4.0% per year.

Three factors can explain the wedge between the average asset return and the average contract return. First, as noted in Section 3.1, the insurer can keep up to 15% of the asset return as profit, which represents about 75 basis points on average. Second, part of the asset returns has been retained to offset the dilution of reserves induced by positive net flows over the sample period.

[19] See Appendix C.1 for details about all the data used in the paper and variable construction. The data are available through Banque de France's open data room (click on this link).

Given the average net flow rate of 2.4% per year and the average reserve ratio of 10.9%, insurers would have had to retain $0.024 \times 0.109 \simeq 25$ basis points of asset returns per year to maintain the reserve-ratio constant. Third, the average reserve ratio is actually about 3.5 percentage points higher at the end of the period than at the beginning (see Appendix Figure B.1), which implies insurers have retained over this 15-year period an additional $0.035/15 \simeq 25$ basis points per year on average.

We complement the regulatory data with contract-level information from two sources. First, we retrieve information on fees from the data provider *Profideo*, which collects information on contract characteristics from contract prospectuses. Our data are a snapshot of contracts with positive outstanding account value in 2017, even if the contract is closed to new subscriptions at that date. The fee structure is fixed at the subscription of the contract and written in the contract prospectus. Given that for every contract there are always some investors who hold their contract for many years, it is sufficient to have a snapshot of outstanding contracts in 2017 to retrieve the fee structure of contracts sold throughout our sample period 2000–2015. The data also includes information on the time period during which contracts were open to new subscriptions. We keep contracts for which this period overlaps with our sample period 2000–2015. 57% percent of insurers, representing 68% of account value in the regulatory filings, can be matched with this dataset.

Panel B of Table 1 shows summary statistics on fees aggregated at the level of insurer-years in which the contract is open to new subscriptions, which is the level at which we run regressions using these data. Management fees are, on average, 70 basis points of account value. Entry fees are, on average, 3.3%.[20]

Our second source of contract-level information is a survey (*Enquête Revalo*) that the insurance supervisor conducted every year from 2011 to 2015 among all the main insurers. The data cover 81% of aggregate account value in the regulatory filings. We retrieve information on net-of-fees contract returns, minimum guaranteed return, total account value, and number of investors, which allows us to calculate the average invested amount for every contract.

Panel C of Table 1 presents summary statistics from this dataset at the contract-survey year level. The average net-of-fees contract return is 2.7%.[21] The average (75th percentile) minimum guaranteed return is 35 basis points (0), which is well below the average contract return of 2.7

[20] Remember that fees do not map one for one into insurer profit, because 90% of fees have to be returned to investors.

[21] It is lower than the average before-fees contract return in regulatory filings (4% in Panel A) minus average management fees (0.7% in Panel B), because the sample period is 2011–2015 for the survey data, whereas it is 2000–2015 for the regulatory filings, and contract returns are lower towards the end of the sample period (see Figure 2).

percentage points over the same period. Thus, the minimum guaranteed rate is typically not binding: The net-of-fees contract return is strictly larger than the guaranteed return for 98% of contracts. This figure actually overstates the extent to which the minimum guaranteed return is binding, because the guaranteed return is before-fees. Assuming uniform management fees at the sample average of 70 basis points, over 99% of contracts have a non-binding minimum guaranteed return.

3.3 Reserve management

Figure 2 shows the time series of the value-weighted average asset return and contract return across insurers. The key pattern is that the contract return is an order of magnitude less volatile than the return on underlying assets. Thus, euro contracts provide insurance against market risk.

As illustrated by Equation (1), the contract return can be hedged against variation in the asset return both by offsetting transfers from the insurer, or by offsetting transfers from reserves. To assess the contribution of reserves to the provision of insurance, Figure 3 shows two series. The solid blue line is the difference between the amount credited on investors' account and asset income ($yV - x(V+R)$ using the notation of the model). It represents the total transfer to current investors, that is, transfer from the insurer plus transfer from reserves. The dashed red line plots the opposite of the change in reserves ($-\Delta R$). This number represents the transfer from reserves. Both series are aggregated across all insurers and normalized by aggregate beginning-of-year account value (V). The figure shows the two series track each other very closely; that is, variation in reserves absorbs almost all of the difference between the asset return and contract return. Therefore, insurance against market risk is implemented by transfers from reserves.

3.4 Inter-cohort redistribution

Transfers from reserves do not mechanically imply inter-cohort redistribution, because part of these transfers net out within investors' holding period. To illustrate this point with a stylized example, consider an investor holding a contract for two years during which the asset return and contract return are as follows:

	Year 1	Year 2
Asset return	0	6
Contract return	4	4

Reserves absorb the difference between the asset return and contract return. In year 1, the investor receives a positive transfer from reserves equal to 4. In year 2, the investor makes a transfer to reserves equal to 2. Therefore, part of the year-on-year transfers net out over the investor's holding period. The net transfer to the investor is then $4 - 2 = 2$ over two years, or 1 per year. Our methodology to quantify inter-cohort redistribution follows the same logic as in this example, netting out transfers within investors' holding period in order to isolate the inter-cohort component.

To quantify inter-cohort redistribution induced by reserve management, we compare the actual contract return paid out to investors with the return they would obtain in a counterfactual with constant reserves, the same asset return, and the same insurer profit as in the data. Relative to the counterfactual, investors holding a contract with insurer j in year t receive a transfer from reserves equal to $-\Delta R_{j,t}$. Consider investor i holding a contract from beginning of year t_0 to end of year t_1, and denote by $V_{i,j,\tau-1}$ her account value at the beginning of year τ. She receives in year τ a transfer proportional to her weight in the insurer's total account value, equal to $\frac{V_{i,j,\tau-1}}{V_{j,\tau-1}}(-\Delta R_{j,\tau})$. Summing over her holding period as in the simple two-period example above, we obtain investor i's lifetime net transfer, which we apportion to each year in proportion to the beginning-of-year account value:[22]

$$NetTransfer_{i,j,t} = \frac{V_{i,j,t-1}}{\sum_{\tau=t_0}^{t_1} V_{i,j,\tau-1}} \sum_{\tau=t_0}^{t_1} \frac{V_{i,j,\tau-1}}{V_{j,\tau-1}}(-\Delta R_{j,\tau}). \tag{15}$$

The net transfer received by an investor depends on her holding period, that is, the year in which she starts investing and the year she redeems (and on the time profile of her investment within the holding period). Investors with same holding period are on the same side of redistribution. By contrast, investors with different holding periods may be on opposite sides of redistribution. Transfers across investors therefore reflect transfers across cohorts. The total amount transferred across cohorts each year t is obtained by summing up across investors:

$$InterCohortTransfer_{j,t} = \sum_i |NetTransfer_{i,j,t}|. \tag{16}$$

We estimate the amount of inter-cohort transfer on the sample of insurers for which we have data throughout 1999–2015, which leads us to make two adjustments to the sample. First, when an insurer acquires another insurer, their reserves are pooled together. In this case, we consolidate

[22] Transfers taking place in different years are not discounted differently, because (85% of) asset returns are due to investors irrespective of the level of reserves; that is, investors are entitled to the same share of asset returns whether assets are credited to the reserves or to their accounts. Therefore, only the total amount of reserve distribution matters, but not its timing within an investor's holding period.

both entities into a single one before the acquisition date such that we have a single insurer with a constant scope throughout the sample period. Second, we drop a few insurers that enter or exit during the sample period or have missing data in some years. The final sample has 50 insurers that we observe continuously from 1999 to 2015 and that account for 94% of the aggregate account value in the initial sample.

Panel A of Table 2 shows the net transfer (15) received by an investor as a function of her holding period, for every possible holding period within the sample period 2000–2015. We calculate the net transfer for an investor who holds the value-weighted average contract and keeps a constant investment amount of 100 by withdrawing interest paid at the end of each year. The numbers in the table represent the additional annual returns of the representative euro contract relative to a counterfactual with constant reserves. For instance, an investor buying a euro contract at the beginning of 2006 and redeeming it at the end of 2011 earned an additional 1.5 percentage points per year relative to a counterfactual with no smoothing, because insurers tapped reserves during the 2008 stock market crash and the 2011 sovereign debt crisis. Conversely, transfers turn negative for holding periods spanning the recent period of decreasing interest rates because insurers hoarded the high bond returns as reserves.

Before calculating the total inter-cohort transfer using (16), we show how a simple back-of-the-envelop calculation can already provide a rough estimate. Suppose all investors have T-year holding periods and that the annual transfer from reserves $-\Delta R_{j,t}$ is i.i.d. across time and normally distributed with zero mean. Then, the expected annualized net transfer amount over T years (i.e., expected $\left|\sum_{t=1}^{T} -\Delta R_{j,t}/T\right|$) is equal to $1/\sqrt{T}$ times the expected yearly transfer amount from reserves (i.e., expected $|-\Delta R_{j,t}|$). Intuitively, a longer holding period reduces the impact of contract return smoothing because a larger fraction of transfers from reserves net out over investors' holding period. The average outflow rate is 8.1% per year, which implies an average holding period of 12 years. The average yearly transfer amount from reserves is 3.7% of account value, implying an average inter-cohort transfer amount of the order of $3.7\%/\sqrt{12} \simeq 1.1\%$ of account value per year. Accounting for holding period heterogeneity would lead to larger inter-cohort transfers because of the convexity of $1/\sqrt{T}$.

To have an exact measure of inter-cohort transfers (16), we would need to observe the entire investment history of all investors, which is not possible, because the investment history of investors still holding a contract at the end of the sample period is not over. Two data limitations also exist. First, regulatory data start in 1999; therefore, we do not observe the entire investment history of investors who entered their contract before 1999. We can calculate the net transfer for investors with holding periods within 2000–2015 (we need one lagged year to calculate asset returns). Second, we observe inflows and outflows at the insurer level but not at the investor level, which implies we know the average holding period but not its entire distribution. To calculate inter-cohort transfers, we assume the outflow rate is constant across cohorts at the insurer-year level and that investors only make one-off investments.[23] Under this assumption, we can reconstruct the investment history of all cohorts of investors and calculate the total inter-cohort transfer.

The value-weighted average amount of inter-cohort transfer is 1.4% of account value (Panel B of Table 2). Evaluated at the 2015 level of aggregate account value of 1,200 billion euros, it amounts to an annual 17 billion euros that shift across cohorts of investors on average, or 0.8% of GDP.[24] In the next section, we study how such a large amount of inter-cohort redistribution can be sustained in a competitive environment.

4 How Can Inter-Cohort Risk Sharing Be Sustained?

The key insight from the model is that the equilibrium level of inter-cohort market risk sharing depends on the elasticity of demand, α. In particular, $\alpha \simeq 0$ allows for perfect inter-cohort risk sharing, and $\alpha \simeq \infty$ implies investors' demand elasticity unravels inter-cohort risk sharing. We now want to measure this elasticity in the data. From the model, we know it can be identified from two moments given by Implications 1 and 2. We estimate each of these moments in turn, by running panel regressions.

4.1 Implication 1: Contract return policy

The first implication of the model is that after controlling for the current reserve ratio, equilibrium contract returns depend positively on current asset returns if $\alpha > 0$, and do not depend on current asset returns if $\alpha \simeq 0$. We estimate the contract return policy (12) given by Implication 1, by

[23] Formally, denoting by $V_{j,t}(t_0)$ the year t-total account value of contracts subscribed from insurer j in year t_0, we assume $V_{j,t}(t_0) = (1 - \theta_{j,t})(1 + y_{j,t})V_{j,t-1}(t_0)$ for all $t_0 < t$, where the outflow rate $\theta_{j,t}$ is calculated to match observed outflows for insurer j in year t, that is, $\sum_{t_0<t} \theta_{j,t}(1+y_{j,t})V_{j,t-1}(t_0) = Outflow_{j,t}$; and account value of new contracts is calculated to match observed inflows to insurer j in year t, that is, $V_{j,t}(t) = Inflow_{j,t}$. See Appendix C.2 for details.

[24] The assumption of outflow rates independent of contract age is likely to underestimate the amount of inter-cohort transfer. Actual outflow rates are decreasing in contract age (FFSA-GEMA, 2013), implying the true dispersion of holding periods is higher than the dispersion obtained under the assumption of the age-independent outflow rate. Again, because expected annualized life transfer is convex in the holding period, underestimating the dispersion of holding periods leads to underestimating inter-cohort transfer.

running a panel regression with year fixed effects. According to our model, insurer fixed effects are not necessary, because the model assumes no heterogeneity in expected asset return across insurers. Our preferred specification includes insurer fixed effects to account for this heterogeneity in the data.[25] We estimate weighted regressions using the insurer share of account value in aggregate account value as weights.[26] We calculate standard errors two-way clustered by insurer and by year.

Results are reported in Table 3. In line with the model, the coefficient on the reserve ratio is positive and statistically significant at the 1% level, in both specifications. In our preferred specification with insurer fixed effects (Column 2), the point estimate implies a one-percentage-point increase in the reserve ratio is associated with a 3.5-basis-point increase in the annual contract return. That is, out of each additional euro of reserves, 3.5 cents per year are credited to investor accounts. The model predicts a regression coefficient equal to $r/(1+r)$. Thus, the estimate of 0.035 implies $r = 3.6\%$, which is reasonable for our sample period 2000–2015.

The coefficient on the asset return is not statistically different from zero when insurer fixed effects are not included (Column 1). In other words, the contract return does not depend on the contemporaneous asset return beyond its effect on the reserve ratio, which is consistent with $\alpha \simeq 0$. The coefficient on the asset return is slightly negative and even becomes statistically significant when insurer fixed effects are included (Column 2). Now, recall that the contemporaneous asset return enters positively into the reserve ratio (Equation (10)). Therefore, the contract return depends positively on the contemporaneous asset return because the sum of the coefficients on the reserve ratio and on the contemporaneous return is positive (equal to 0.17 with p-value at 0.15). The negative coefficient implies contract return in year t is more sensitive to lagged asset returns (in year $s < t$) than to contemporaneous asset returns (in year t). Two institutional factors can explain this seemingly surprising result. First, amounts withdrawn through the calendar year are usually credited a pro rata return calculated based on the lagged contract return. Second, insurers sometimes guarantee to new clients a higher return in the first year of the contract for marketing purposes. As a result, contract returns associated with inflows and outflows do not depend on the current year asset return, which weakens the relation between the contract return and contemporaneous asset return.

In conclusion, the empirical contract return policy rejects $\alpha > 0$ and is instead consistent with

[25] In the model, including insurer fixed effects does not lead to a misspecified regression, because it only adds regressors uncorrelated with the dependent variable and with the other explanatory variables. In the data, insurer fixed effects in contract return regressions are always jointly significant at statistical levels below 1%.

[26] We obtain similar results when we estimate non-weighted regressions (untabulated).

$\alpha \simeq 0$.

4.1.1 Are reserves really pooled across investor cohorts?

Inter-cohort risk sharing arises to the extent that reserves are pooled across investor cohorts. As discussed in Section 3.1, insurers could in principle undo reserves pooling by closing existing contracts to new subscriptions when reserves are high, and creating a new vintage of contracts for new investors in order to price the high level of reserves. Pricing of reserves could be done by creating new contracts with (a) higher entry fees, (b) higher management fees, (c) lower before-fees contracts return, or any combination of (a), (b), and (c), when reserves are higher.

We use two different sources of contract-level information to test whether insurers do (a), (b), or (c). First, we use data on fees to test for (a) and (b). The data are a snapshot of contracts with positive outstanding account value in 2017 (even if the contract is no longer commercialized in 2017). The fee structure is fixed at the subscription of the contract and written in the contract prospectus. Because, for a given contract, a number of investors will always hold their contract for many years, it is sufficient to have a snapshot of outstanding contracts in 2017 to retrieve the fee structure of contracts sold throughout our sample period 2000–2015. The data also report the time period during which each contract was open to new subscriptions. For each insurer j and each year t over 2000–2015, we calculate the average entry fee and average management fee across all contracts offered by insurer j and open to new subscriptions in year t. We regress the average fee on the insurer's beginning-of-year reserve ratio. If insurers price reserves into fees, the coefficient on the reserve ratio would be positive. Results in Table 4 show that insurers do not adjust either entry fees (Column 1) or management fees (Column 2) to the level of reserves.

Second, we use data on net-of-(management-)fees returns at the contract level to test whether insurers do a combination of (b) and (c). The data are from a survey that has been run by the insurance supervisor since 2011. Each survey is a snapshot of contracts with positive outstanding contract value (even if the contract is no longer commercialized in the survey year) with information on the contract net-of-fees return.

The data report the first year in which the contract was commercialized. For each contract c of vintage s, we retrieve the insurer's reserve ratio at the beginning of year s from the regulatory filings. We obtain a panel at the contract (c) $\times$ vintage year (s) $\times$ return year (t) level, where vintage years run throughout our sample period 2000–2015 and return years are from 2011 to 2015. We regress the net-of-fees return (of contract c in year t) on the reserve ratio in the contracts' vintage year (at beginning of year s) with insurer and vintage year fixed effects.[27] If insurers price reserves by adjusting future net-of-fees contract returns, the coefficient on the reserve ratio in the contract's vintage year would be negative. Column 3 of Table 4 shows insurers do not discriminate across investor cohorts based on the level of reserves when investors enter into the contract.

In conclusion, reserves are indeed pooled across investor cohorts.

4.2 Implication 2: Flow-reserves relation

The second implication of the model is that investor flows depend positively on the reserve ratio if $\alpha > 0$, whereas flows are insensitive to reserves if $\alpha = 0$. In the model, the flow-reserves relation (13) has the log level of the invested amount as the dependent variable, rather than the change in the invested amount as flows are usually defined. The reason is that investments are assumed to be one-period in the model, so that the outflow rate is 100% at the end of each period. Instead, real-world contracts are automatically renewed from year to year unless the investor redeems shares. Accordingly, we estimate the flow-reserves relation using the usual concept of net flows, defined as inflows minus outflows divided by beginning-of-year account value. We estimate panel regressions with insurer and year fixed effects. We run separate regressions for net flows and for the three components of net flows: (plus) inflows, that is, premia, which come either from investors already holding a contract and adding money to their account, or from new investors; (minus) redemptions, which are voluntary outflows; and (minus) payments at contract termination, which are involuntary outflows (due to investor death). One should expect the former two to respond to the level of reserves if $\alpha > 0$, but not the latter. One might also expect inflows to be more sensitive to reserves than redemptions, because redemptions may be more likely to be driven by liquidity motives.

Table 5 contains the results. The sensitivity of net flows to the beginning-of-year reserve ratio is not significantly different from zero (Column 1). The net-flow decomposition yields similar results: Neither inflows (Column 2) nor outflows (Columns 3 and 4) are sensitive to reserves. All the coefficients are precisely estimated zeros. We can reject at the 5% level that the regression coefficient of net flows on the reserve ratio is larger than 0.12.

To see why a flow-reserves sensitivity of 0.12 is economically small, we can compare this value with the flow-reserves sensitivity that would be required for flows to fully dilute an extra euro of reserves at a one-year horizon and thus eliminate contract return predictability. Given that reserves represent, on average, 11% of account value, a flow-reserves sensitivity of $1/0.11 \simeq 9$ would

[27] Because we stack the five snapshots of return data, we interact the fixed effects with return-year dummies.

be necessary for flows to fully dilute reserves at a one-year horizon,[28] that is, a 75-fold larger sensitivity than the one we can reject at 5%.

As shown in Appendix A.7, the OLS estimate of the flow-reserves relation is biased downwards if insurers face anticipated flow shocks. Intuitively, when an insurer anticipates negative flow shocks in future periods, it has incentives to lean against the shock by hoarding reserves in order to pay higher returns when the shock hits. This behavior creates a negative correlation between flows and reserves, biasing the OLS downwards. We further show this bias can be corrected by instrumenting reserves using past asset returns, if flow shocks are not correlated with past asset returns (beyond the causal effect of past asset returns on current reserves, thus on future contract returns, thus on future demand if $\alpha > 0$).

We run IV regressions using the previous year's asset return to instrument the beginning-of-year reserve ratio. The first stage is strongly significant (F-stat equal to 29 with standard errors two-way clustered by insurer and year). The second stage regressions are presented in Panel B of Table 5. In Column 1, the IV estimate of the net flow-reserves sensitivity is slightly higher than the OLS estimate, but it remains very small and statistically insignificant. In Column 3, the redemption-reserves sensitivity is negative (i.e., investors redeem less when reserves are higher) and significant at 10%, but the economic magnitude remains very small. Overall, the IV regressions confirm that flows are at best barely elastic to reserves. To conclude, the empirical flow-reserves relationship rejects $\alpha > 0$ and is instead consistent with $\alpha \simeq 0$.

4.3 Flows are inelastic to predictable returns

In the model, $\alpha \simeq 0$ implies (a) contract returns are predictable but (b) investor flows are inelastic to these predictable returns. In this section, we show (a) and (b) are borne out in the data.

4.3.1 Contract returns are predictable

Because reserves are not diluted by investor flows (as shown in the previous section) and are owed to investors (by regulation), the reserve ratio should predict future contract returns. We test for contract return predictability in Table 6. In Column 1, we regress the contract return paid at the end of year t on the reserve ratio at the beginning of year t in the insurer-year panel with year

[28]Denoting the flow-reserves sensitivity by η, an extra dR of reserves is fully diluted by flows if $\frac{R+dR}{V+\eta dR} = \frac{R}{V}$, i.e., if $\eta = V/R$.

fixed effects.[29] The coefficient on the beginning-of-year reserve ratio is positive and statistically significant at the 1% level. Therefore, the reserve ratio predicts the expected contract return at a one-year horizon: Contracts with higher reserves have higher expected returns.

Higher reserves predict higher expected contract return because reserves are eventually distributed to investors– not because higher reserves are associated with higher risk. As a check, we consider a zero-cost portfolio that is invested long in contracts with high reserves and short in contracts with low reserves. At the beginning of each year, we rank insurers on the $[0, 1]$ interval based on the beginning-of-year reserve ratio, and use portfolio weights proportional to insurers' rank minus one-half.

Columns 1 and 2 of Table 7 show the performance of each leg of the portfolio, and Column 3 that of the long-short portfolio. The first row confirms that higher reserves predict higher expected returns: Average returns are 34 basis points higher per year for high-reserves contracts than for low-reserves contracts.

The second and third rows report the estimates of a market model. The difference in market beta between high-reserves and low-reserves contracts is a precisely estimated zero (a difference in beta larger than 0.01 is rejected at the 1% level), implying alpha is 34 basis points higher for high-reserves contracts than low-reserves contracts, on average. Therefore, the predictability of expected contract returns does not reflect a compensation for market risk.

The fourth row reports the cross-sectional standard deviations of high- and low-reserves contracts returns, averaged over time. We find the difference between the two groups is a precisely estimated zero. Therefore, the predictability of expected contract returns does not reflect a compensation for idiosyncratic risk either.

Reserves should predict contract returns not only at one year but also at longer horizons, because reserves are only progressively distributed to investors. This progressive distribution matters because, as we explain in the next section, a profitable strategy to exploit contract return predictability involves holding contracts for several years because of switching costs. We show in Appendix D that the predictive power of reserves for future contract returns decays at the same rate as the one at which the reserve ratio mean reverts. The reserve ratio mean reverts for two reasons. First, reserves are progressively credited to investors' accounts (at a rate of 3% per year, see Columns 1–2 of Table 3). Second, inflows dilute reserves at a rate equal to the unconditional net flow rate (2.4%

[29]We do not include insurer fixed effects, because we are running a predictive regression, and insurer fixed effects would be estimated on the entire sample period. In the (untabulated) regression with insurer fixed effects, the coefficient on the lagged reserve ratio is .03 significant at 1%.

per year, see Table 1) plus a term that depends on the sensitivity of flows to reserves (equal to zero, see Table 5). Thus, the reserve ratio mean reverts at a rate of $\delta \simeq 5.4\%$ per year. The predictive power of reserves for future contract returns should also decay at a rate of 5.4% per year.

We check in Columns 2–5 of Table 6 that our calculation for the decay of the predictive power of reserves is in line with the data. We regress contract return in years t, $t+1$, …, $t+4$, on the reserve ratio at the beginning of year t. The regression coefficient on the initial reserve ratio decays at a rate of about 7%, close to our estimate of $\delta \simeq 5.4\%$. Thus, reserves predict future contract returns over many years.

4.3.2 Flows are inelastic to predictable returns

We have shown investor flows are inelastic to reserves (Table 5). Does this finding imply investors fail to exploit return predictability? Not necessarily, because this inelasticity could stem from the costs associated with strategies aimed at exploiting contract return predictability. For instance, to move their money from one insurer to another one with a higher level of reserves, existing investors face entry fees that create a switching cost that offsets the returns from contract return predictability computed in Section 4.3.1.

We show that other strategies aimed at exploiting contract return predictability are not offset by switching costs. First, an investor increasing his investment amount with money that is not already invested in a contract incurs no switching cost. Regardless of the contract chosen, the investor must pay entry fees. Because insurers do not adjust fees to the level of reserves (see Table 4), an investor seeking to maximize returns should choose the contract with the highest reserves. This argument has one twist if the investor already has a contract with an insurer. Contract returns are taxed upon withdrawal at a rate that depends on the age of the contract at the time of withdrawal, where contract age is defined as the number of years since the investor's first investment in the contract. The tax rate is decreasing in contract age for the first eight years of the contract. Therefore, if an investor considers increasing his investment amount and already owns a contract, she has tax incentives to invest with her existing insurer if she has an investment horizon shorter than eight years. Now, the average investment horizon is 12 years, that is, longer than the eight-year period over which the tax distortion applies. As a result, the tax distortion is unlikely to explain why inflows by investors already holding a contract are inelastic to reserves.[30]

Second, new investors are not subject to any distortion. Conditional on purchasing a new

[30]See Appendix E for a description of the tax treatment of euro contracts and a quantification of the tax distortion.

contract, an investor seeking to maximize returns should unambiguously choose the contract with the highest reserves. We estimate the sensitivity of purchases of new contracts to the level of reserves. The regulatory filings contain information on the number of new contracts sold by the insurer in the current year. Insurers have been required to report this information since 2006; therefore, the sample period for this test is restricted to 2006–2015. We regress the number of new contracts sold divided by the number of outstanding contracts on the beginning-of-year reserve ratio. Table 8 shows that both in our OLS and IV estimations, new investors' inflows are not sensitive to the level of reserves.

We conclude that investors fail to exploit contract return predictability.

4.4 Why are flows inelastic to predictable returns?

We explore the hypothesis that flows do not react to predictable returns because investors lack the knowledge to predict contract returns using reserves. The reason may be that investors simply do not understand that reserves predict returns, or perhaps investors are not able to obtain information on the level of reserves.[31] To test that hypothesis, we study whether the flow-reserves sensitivity varies across investors with different levels of financial sophistication. We proxy for investor sophistication using the investment amount. The idea is that financial sophistication is correlated with wealth, for instance, if investors must incur a fixed cost to acquire the knowledge necessary to predict returns (Lusardi and Mitchell, 2014).

We construct the proxy for investor sophistication using contract-level data collected by the insurance supervisor for the years 2011 to 2015. The data contains information on the number of investors, the total account value, and the net-of-fees return for every contract. We calculate the average individual account value as the total account value divided by the number of investors. We classify contracts into three size bins according to the average account value: below 50,000 euros, 50,000–250,000 euros, and above 250,000 euros. We also construct net flows at the contract level.

We exploit cross-sectional variation in investor sophistication along two dimensions. First, we exploit it across insurers. Some insurers cater to wealthier, and hence more sophisticated, clienteles. Second, we exploit variation across contracts within a given insurer. As described in Section 3.1, insurers often offer different contracts with different minimum investment amounts that target different clientele. A crucial feature of the institutional framework is that reserves are pooled across

[31] Although insurers' annual reports contain information on the level of reserves, it often is incomplete or consolidated at the group level.

all contracts of a given insurer, so that reserves predict returns for all contracts. Therefore, we can exploit cross-contract variation in investor sophistication to test whether the flow-reserves sensitivity varies within a given insurer-year. We regress net flows at the contract level on the beginning-of-year reserve ratio interacted with dummy variables for each bin of average account value (and on the non-interacted dummy variables).

The first specification (Column 1 of Table 9) does not include insurer-year fixed effects and thus exploits cross-insurer variation in investor sophistication. The flow-reserves sensitivity is small and statistically insignificant both for contracts with small and intermediate average account value (below 250,000 euros per investor). By contrast, the flow-reserves sensitivity is positive and statistically significant at the 10% level for contracts with larger average account value (above 250,000 euros per investor).

The second specification (Column 2 of Table 9) includes insurer-year fixed effects and thus isolates cross-contract variation in investor sophistication within insurer-years. In that case, the absolute level of the flow-reserves sensitivity is no longer identified because it is defined at the insurer-year level. We use the small-average-account-value category as the reference group. The results are consistent with those obtained in the first specification: The flow-reserves sensitivity is larger for contracts with large account values than for contracts with smaller account values. The difference is significant at the 1% level. The IV estimates yield similar results (Columns 3 and 4).

These results suggest investors with sufficient skills or incentives to predict returns are able to time reserves to some extent. However, note that although the flow-reserves relation becomes statistically significant among investors with large invested amounts, the economic magnitude remains small. Recall from Section 4.2 that a flow-reserves sensitivity of 9 would be necessary for flows to fully dilute reserves and undo return smoothing. In comparison, the estimated sensitivity in Table 9 never exceeds 1 even among the most sophisticated investors.

4.5 Do arbitrage opportunities exist?

Even though, in the data, we find $\alpha \simeq 0$ for the bulk of investors, we know from Section 2.4 that inter-cohort risk sharing also requires there exist no arbitrage strategies that consist in going long in euro contracts and short in the same assets as intermediaries and in a risk-free asset. If such strategies were profitable, a single arbitrageur would upset the inter-cohort risk sharing equilibrium. However, these arbitrage strategies are unprofitable if the capital income tax rate is large enough. In this section, we calibrate the parameters in the arbitrage profit given by Equation (14) to determine the tax rate necessary to eliminate arbitrage opportunities.

There are two distinct sources of arbitrage profits. The first one is the risk premium that can be earned without bearing the associated risk. When $\alpha \simeq 0$, the contract return is almost risk-free, and the first component of the arbitrage profit is approximately equal to the risk premium, $r - r_f$. We calibrate the expected asset return using the sample average asset return (4.9%, see Table 1), and noting that it is likely realized asset returns have blue above expected returns during the sample period. As discussed in Section 3.2, the reserve ratio rose by 25 basis points per year, while positive net flows should have diluted reserves at a rate of 25 basis points per year. Therefore, insurers have retained in reserves approximately $25 + 25 = 50$ basis points of the realized asset returns in excess of expected returns. Therefore, we set $r = 4.4\%$. Using $r_f = 3\%$, the risk premium is 1.4%.

The second source of arbitrage profit is the predictable distribution of reserves, which depends on the beginning-of-year reserve ratio. To focus on a situation that makes the arbitrage most profitable, we assume the reserve ratio is 10 percentage points above target. This represents 1.5 standard deviations of the reserve ratio (see Table 1). It also amounts to the difference between the highest point of the aggregate reserve ratio (reached in 2014, see Appendix Figure B.1) and its sample average. Using the fact that reserves are distributed to investors at a rate of 3% per year (see Columns 1 and 2 of Table 3), the second term of the arbitrage profit is $0.3\% \times (1 - \tau)$, where τ is the capital income tax rate.

If we calibrate insurer compensation to the French regulatory framework, in which insurers can keep up to 15% of asset returns, we have $\phi = 0.15 \times r \simeq 0.7\%$, which is equal to the average management fee in the data (see Table 1). In this case, arbitrage opportunities are eliminated if $1.4 + 0.3 \times (1 - \tau) - 4.4 \times \tau - 0.7 \times (1 - \tau) < 0$, that is, if the capital income tax rate is greater than 25%. As described in Appendix E.1, the applicable tax rate depends on the contract holding period. At the end of the sample period, the lowest possible tax rate is 23% (15.5% of social security contributions plus 7.5% of income tax). Hence, the actual minimum tax rate is close to our estimate of the minimum tax rate necessary to eliminate arbitrage opportunities.

As noted in Section 2.4, arbitrage opportunities can be eliminated even if insurer compensation is zero. If $\phi = 0$, arbitrage opportunities are eliminated if $1.4 + 0.3 \times (1 - \tau) - 4.4 \times \tau < 0$, that is, if the capital income tax rate is greater than 36%.

Remark that the absence of arbitrage opportunities is not contradictory with our finding in Section 4.4, where we show that the flow-reserves relation is statistically significant among investors with large invested amounts (see Table 9). Indeed, sophisticated households who have positive amounts of savings always should buy contracts with high reserves rather than contracts with low reserves (see Section 4.3.2), even in the presence of a capital income tax. Yet, buying contracts and shorting underlying assets and the risk-free asset is not profitable in the presence of a large enough capital income tax.

5 Conclusion

We provide the first evidence of a large scale, and private, implementation of inter-cohort risk sharing. The evidence implies that financial intermediaries can complete markets, by allowing investors with different holding periods to share risk, which they cannot achieve even in fully developed financial markets. Such inter-cohort risk sharing is desirable from an *ex-ante* welfare perspective, that is, under the Rawlsian veil of ignorance (Gordon and Varian, 1988; Ball and Mankiw, 2007).

Private implementation of inter-cohort risk sharing requires a two-sided commitment problem to be overcome (Allen and Gale, 1997). First, regulation ensures that intermediaries eventually return reserves to households. This suggests a reason why inter-cohort risk-sharing savings products exist in several European countries, where such regulation exists, but not in the US, where it does not.

Second, investors must remain invested in contracts even when reserves are low. We show that investment flows are inelastic to predictable reserves, and do not tumble when reserves are low. This low elasticity is more prevalent among households who are expected to have lower financial sophistication. Therefore, perhaps counter-intuitively, lower investor sophistication enables a better sharing of risk—across investor cohorts—than what would be possible in a frictionless economy.

These results have implications for real investment, which we leave for future research. First, spreading aggregate risk across cohorts implies that aggregate consumption is smoothed over time, which requires the capital stock to increase in good time and to decrease in bad time. Hence, inter-cohort risk sharing has implications for the cyclicality of aggregate investment. Second, as studied theoretically by Gollier (2008), intermediaries can invest in more risky assets when risk is shared across cohorts. Therefore, inter-cohort risk sharing has implications for the composition of aggregate investment.

References

ACPR. 2016. "Les chiffres du marché français de la banque et de l'assurance."

Allen, Franklin and Douglas Gale. 1997. "Financial markets, intermediaries, and intertemporal smoothing." *Journal of Political Economy* 105 (3):523–546.

Aubier, Maud, Frédéric Cherbonnier, and Daniel Turquety. 2005. "Influence de la fiscalité sur les comportements d'épargne." *Economie & prévision* (3):321–329.

Ball, Laurence and N Gregory Mankiw. 2007. "Intergenerational risk sharing in the spirit of arrow, debreu, and rawls, with applications to social security design." *Journal of Political Economy* 115 (4):523–547.

Bianchi, Milo. 2018. "Financial literacy and portfolio dynamics." *Journal of Finance* 73 (2):831–859.

Célérier, Claire and Boris Vallée. 2017. "Catering to investors through security design: headline rate and complexity." *Quarterly Journal of Economics* (Forthcoming).

Darpeix, Pierre-Emmanuel. 2016. "Le taux technique en assurance vie (code des assurances)." *Analyses et Synthèses (Banque de France)* .

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl. 2017. "The deposits channel of monetary policy." *The Quarterly Journal of Economics* 132 (4):1819–1876.

———. 2018. "Banking on deposits: Maturity transformation without interest rate risk." *National Bureau of Economic Research Working Paper* .

FFSA-GEMA. 2013. "Les rachats des contrats d'assurance vie après 60 ans." *Document de travail* Conseil d'Orientation des Retraites (8).

Gollier, Christian. 2008. "Intergenerational risk-sharing and risk-taking of a pension fund." *Journal of Public Economics* 92 (5):1463–1485.

Gordon, Roger H and Hal R Varian. 1988. "Intergenerational risk sharing." *Journal of Public Economics* 37 (2):185–202.

Greenwood, Robin M and Annette Vissing-Jorgensen. 2018. "The impact of pensions and insurance on global yield curves." .

Handel, Benjamin R. 2013. "Adverse selection and inertia in health insurance markets: When nudging hurts." *American Economic Review* 103 (7):2643–82.

Insee. 2016. "Les revenus et le patrimoine des ménages." *Insee Références* édition 2016.

Koijen, Ralph and Motohiro Yogo. 2018. "The fragility of market risk insurance." .

Koijen, Ralph SJ and Motohiro Yogo. 2015. "The cost of financial frictions for life insurers." *American Economic Review* 105 (1):445–475.

Lusardi, Annamaria and Olivia S Mitchell. 2014. "The economic importance of financial literacy: Theory and evidence." *Journal of economic literature* 52 (1):5–44.

Novy-Marx, Robert and Joshua Rauh. 2011. "Public pension promises: how big are they and what are they worth?" *Journal of Finance* 66 (4):1211–1249.

Novy-Marx, Robert and Joshua D Rauh. 2014. "Linking benefits to investment performance in US public pension systems." *Journal of Public Economics* 116:47–61.

Scharfstein, David S. 2018. "Presidential address: Pension policy and the financial system." *The Journal of Finance* 73 (4):1463–1512.

Stiglitz, J. 1983. "On the relevance of public financial policy: indexation, price rigidities and optimal monetary policies."

Tables and Figures

Figure 1: Aggregate Account Value. The figure shows aggregate account value of euro contracts in billion 2015 (solid blue) euros and the number of insurers in the sample (dashed red).

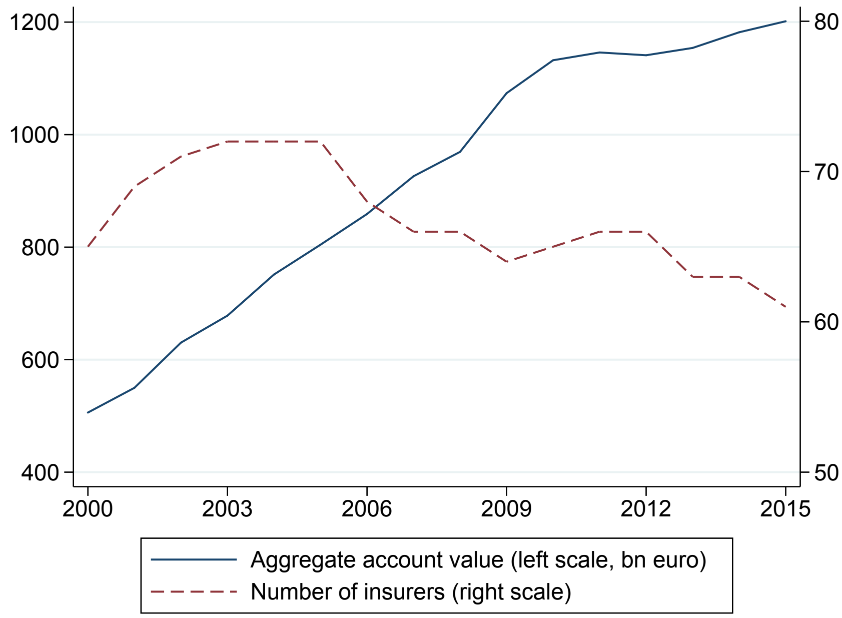


Figure 2: Asset Return vs. Contract Return. The figure shows value-weighted average contract return (solid blue) and value-weighted average asset return (dashed red).

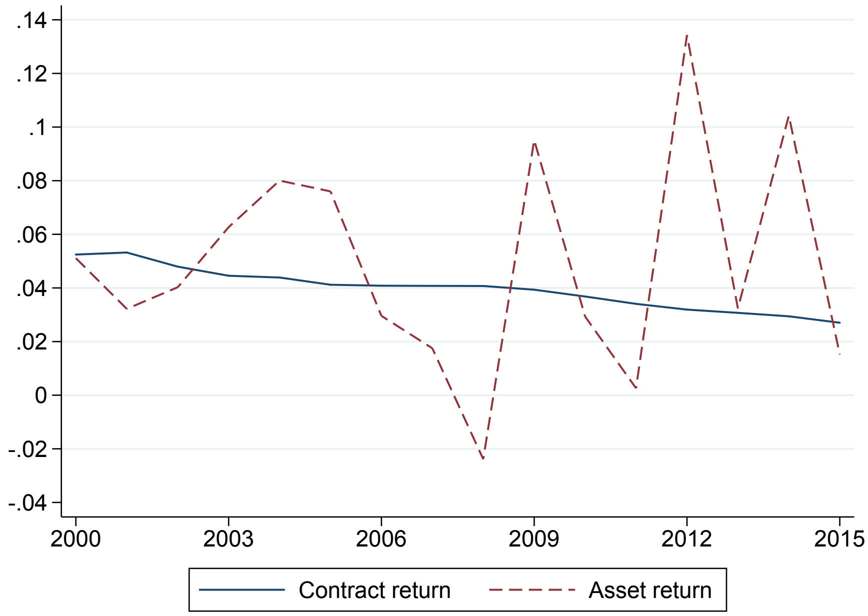


Figure 3: Reserves Absorb Asset Return Fluctuations. The figure shows the difference between aggregate contract return and asset return normalized by account value $(y_t V_{t-1} - x_t A_{t-1})/V_{t-1}$ (solid blue) and aggregate transfer from reserves normalized by account value $-\Delta R_t / V_{t-1}$ (dashed red).

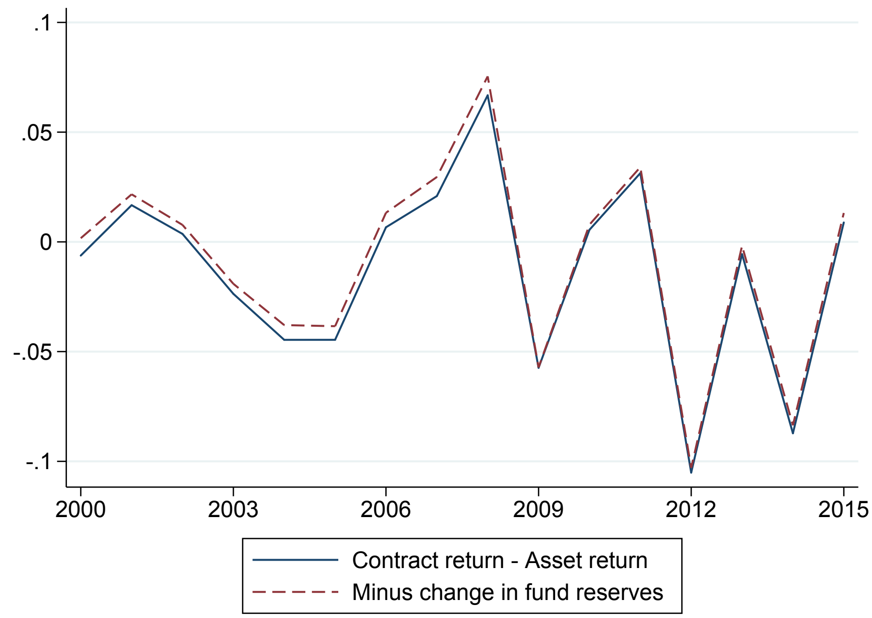


Table 1: Summary Statistics Panel A presents regulatory filings data at the insurer-year level for 76 insurers over 2000–2015. All statistics (except for account value) are weighted by the insurer share in aggregate account value in the current year. *Account value* is total account value at year-end in constant 2015 billion euros. *Inflows* are inflows (premia) divided by beginning-of-year account value plus one-half of net flows. *Outflows* are outflows (redemptions plus payment at contract termination) divided by beginning-of-year account value plus one-half of net flows. *Reserves* is total reserves divided by year-end account value. *Portfolio share: bonds* is the share of (corporate and sovereign) bonds, held either directly or through funds, in the asset portfolio. *Portfolio share: stocks* is the share of stocks, held either directly or through funds, in the asset portfolio. *Asset return* is the asset return. *Contract return* is the average before-fees contract return. Panel B presents prospectus data on fees at the insurer-year level for 48 insurers over 2000–2015. *Management fees* is the average management fees across contracts offered by the insurer and open to new subscriptions in the current year. *Entry fees* is the average entry fees across contracts offered by the insurer and open to new subscriptions in the current year. Panel C presents survey data at the contract-year level for about 2,700 outstanding contracts per year from 56 insurers over 2011–2015. *Net-of-fees return* is the contract net-of-fees return. *Minimum guaranteed return* is the before-fees minimum return guaranteed by the insurer.

	Mean	S.D.	P25	P50	P75	N
Panel A: Regulatory Filings						
Account value (bn euro)	13.9	30.1	0.9	3.1	11.9	978
Inflows (% account value)	10.5	3.8	7.8	10.5	12.3	978
Outflows (% account value)	8.1	2.0	7.1	7.9	8.8	978
Reserves (% account value)	10.9	6.8	6.7	10.5	14.3	978
Portfolio share: bonds (%)	80.4	8.0	75.4	81.5	85.6	978
Portfolio share: stocks (%)	13.5	6.3	10.0	12.5	15.7	978
Asset return (%)	4.9	4.4	2.1	4.4	7.5	978
Contract return (%)	4.0	0.9	3.3	4.0	4.5	978
Panel B: Prospectus Data						
Management fee (%)	.7	.13	.64	.73	.77	578
Entry fee (%)	3.3	.87	3	3.5	3.8	578
Panel C: Survey Data						
Net-of-fees return (%)	2.7	.45	2.4	2.8	3	13,672
Minimum guaranteed return (%)	.35	.73	0	0	0	13,672

Table 2: Inter-Cohort Redistribution. In Panel A, *Net transfer* is defined in (15) for an investor buying a contract at the beginning of year t_0 (rows) and redeeming it at the end of year t_1 (columns). Reading: An investor buying a contract at the beginning of 2006 and redeeming it at the end of 2011 received an additional 1.5 percentage points per year relative to a counterfactual with constant reserves. In Panel B, *Inter-cohort transfer* is defined in (16) and equal to the sum of lifetime net transfer across investors divided by total account value.

Panel A: Net transfer by investor cohort

entry year \ exit year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
2000	0.2	1.1	0.6	0	−0.7	−1.2	−0.9	−0.5	0.4	−0.2	−0.2	0.1	−0.7	−0.6	−1.1	−1
2001		2	0.8	−0.1	−0.9	−1.5	−1.1	−0.6	0.4	−0.3	−0.2	0.1	−0.7	−0.7	−1.2	−1.1
2002			−0.4	−1.2	−1.9	−2.4	−1.7	−1.1	0.2	−0.6	−0.4	0	−1	−0.9	−1.5	−1.3
2003				−2	−2.6	−3	−2	−1.2	0.2	−0.6	−0.4	0	−1.1	−1	−1.6	−1.4
2004					−3.2	−3.5	−2	−1	0.7	−0.4	−0.2	0.2	−1	−0.9	−1.5	−1.3
2005						−3.8	−1.4	−0.2	1.7	0.2	0.3	0.7	−0.7	−0.6	−1.4	−1.1
2006							0.9	1.6	3.5	1.2	1.1	1.5	−0.2	−0.2	−1.1	−0.8
2007								2.2	4.8	1.3	1.1	1.6	−0.4	−0.4	−1.4	−1.1
2008									7.5	0.9	0.8	1.4	−0.9	−0.8	−1.9	−1.5
2009										−5.7	−2.6	−0.6	−3	−2.5	−3.4	−2.7
2010											0.6	2	−2.1	−1.6	−3	−2.2
2011												3.4	−3.5	−2.4	−3.8	−2.8
2012													−10.4	−5.3	−6.3	−4.3
2013														−0.2	−4.2	−2.4
2014															−8.2	−3.4
2015																1.4

Panel B: Inter-cohort redistribution

Inter-cohort transfer	
in % account value	1.4
in 2015 euros	17 billion
in % GDP	0.8

Table 3: Contract Returns. Panel regressions at the insurer-year level for 76 insurers over 2000–2015. *Contract return* is the annual before-fees contract return paid at the end of year t. *Reserve ratio* is total reserves at the end of year t just before annual distribution normalized by total account value. *Asset return* is asset return in year t. All regressions are weighted by the insurer share in aggregate account value in the current year. Standard errors two-way clustered by insurer and year are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	Contract return ($y_{j,t}$)	
	(1)	(2)
Reserve ratio (R_{j,t^-})	.026***	.035***
	(.0078)	(.0081)
Asset return ($x_{j,t}$)	-.017	-.018**
	(.011)	(.0079)
Year FE	✓	✓
Insurer FE		✓
R^2	.69	.81
Observations	978	978

Table 4: Fees. Columns 1 and 2 present panel regressions at the insurer-year level for 48 insurers over 2000–2015. The dependent variable in Column 1 is *Entry fee* constructed as the average entry fee (frond-end load) of contracts sold by the insurer j in year t. The dependent variable in Column 2 is *Management fee* constructed as the average management fee of contracts sold by insurer j in year t. The independent variable in Columns 1 and 2 is *Lagged reserves* constructed as insurer j's reserves at beginning-of-year t normalized by total account value. The regressions in Columns 1 and 2 include insurer and year fixed effects and are weighted by the insurer share in aggregate account value in the current year. Column 3 presents a panel regression at the contract-vintage year-return year level for about 2,700 outstanding contracts per year from 56 insurers over 2011–2015. The dependent variable in Column 3 is contract return in year t of contract c of vintage year s offered by insurer j. The independent variable in Column 3 is *Lagged reserves* constructed as insurer j's reserves at beginning-of-year s normalized by total account value. The regression in Column 3 includes insurer-return year and vintage year-return year fixed effects and are weighted by the contract share in aggregate account value in the current return year. Standard errors two-way clustered by insurer and year (return year for Column 3) are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	Entry fee	Management fee	Net-of-fee contract return
	(1)	(2)	(3)
Lagged reserves	-.016	.000054	-.005
	(.011)	(.0011)	(.0056)
Year FE	✓	✓	✓
Insurer FE	✓	✓	✓
R^2	.92	.95	.72
Observations	578	578	13,659

Table 5: Investor Flows. Panel regressions at the insurer-year level for 76 insurers over 2000–2015. *Inflows* is total premia normalized by total account value. *Redemptions* is voluntary redemptions normalized by total account value. *Termination* is involuntary redemptions at contract termination (investor death) normalized by total account value. *Net flows* is Inflows minus Redemptions minus Termination. *Lagged reserves* is the beginning-of-year level of reserves normalized by total account value. Panel A shows OLS regressions. Panel B shows IV regressions in which the insurer's beginning-of-year reserve ratio is instrumented using the insurer's asset return in the previous year (the first year of data for each insurer is therefore dropped from the second stage). All regressions are weighted by the insurer share in aggregate account value in the current year. Standard errors two-way clustered by insurer and year are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	Panel A: OLS Regressions			
	Net flows	Inflows	Redemptions	Termination
	(1)	(2)	(3)	(4)
Lagged reserves	.035	.031	-.013	.012
	(.038)	(.037)	(.019)	(.0097)
Year FE	✓	✓	✓	✓
Insurer FE	✓	✓	✓	✓
R^2	.63	.73	.74	.79
Observations	978	978	978	978

	Panel B: IV Regressions			
	Net flows	Inflows	Redemptions	Termination
	(1)	(2)	(3)	(4)
Lagged reserves	.086	-.02	-.078*	-.025
	(.098)	(.091)	(.041)	(.02)
Year FE	✓	✓	✓	✓
Insurer FE	✓	✓	✓	✓
R^2	.66	.77	.75	.8
Observations	859	859	859	859

Table 6: Contract Return Predictability. Panel regressions at the insurer-year level for 76 insurers over 2000–2015. *Contract return* is the annual before-fees contract return the end of years t (Column 1), $t+1$ (Column 2), ..., $t+4$ (Column 5). *Reserves at beginning of year t* is total reserves at the beginning-of-year t normalized by total account value. All regressions include year fixed effects and are weighted by the insurer share in aggregate account value in the current year. Standard errors two-way clustered by insurer and year are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	Contract return in year				
	t	$t+1$	$t+2$	$t+3$	$t+4$
	(1)	(2)	(3)	(4)	(5)
Reserves at beginning of year t	.025***	.024***	.023**	.019**	.019*
	(.0074)	(.0073)	(.0078)	(.0087)	(.0088)
Year FE	✓	✓	✓	✓	✓
R^2	.69	.71	.62	.61	.57
Observations	978	859	783	717	645

Table 7: High-Reserves Contracts Are Not Riskier. Performance of a portfolio long contracts with beginning-of-year reserves above median and short contracts with beginning-of-year reserves below median with portfolio weights proportional to the contract rank rescaled between minus one and one times the contract's total account value. Column 1 shows the performance of the short leg, Column 2 of the long leg, and Column 3 the performance of the long-short portfolio. *Mean return* is the average return of the leg/portfolio. *Alpha* and *Beta* are the intercept and loading on the market in the market model. *S.D. return* is the time-series average of the cross-sectional standard deviation of contract return within the leg in Columns 1 and 2, and it is the difference between that of the long leg and that of the short leg in Column 3. Newey-West standard errors with two lags are reported in parenthesis. In Column 3, ***, **, and * mean that the difference between the long leg and the short leg is statistically significant at the 1%, 5%, and 10% levels, respectively.

	Low-reserves contracts (1)	High-reserves contracts (2)	Difference High minus Low (3)
Mean return	.039	.042	.0034***
	(.0029)	(.0030)	(.00035)
Alpha	.039	.042	.0034***
	(.0027)	(.0029)	(.00032)
Beta	-.012	-.0092	.0027*
	(.0070)	(.0075)	(.00092)
S.D. return	.0051	.0042	-.00091
	(.00034)	(.00049)	(.00069)

Table 8: Inflows From New Investors. Panel regressions at the insurer-year level for 67 insurers over 2006–2015 *Purchases of new contracts* is the number of new contracts purchased in the current year divided by the beginning-of-year outstanding number of contracts. *Lagged reserves* is the beginning-of-year level of reserves normalized by total account value. Column 1 shows the OLS regression. Column 2 shows the IV regression in which the insurer's beginning-of-year reserve ratio is instrumented using the insurer's asset return in the previous year. All regressions include insurer and year fixed effects and are weighted by the insurer share in aggregate account value in the current year. Standard errors two-way clustered by insurer and year are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	Purchases of new contracts	
	OLS	IV
	(1)	(1)
Lagged reserves	.029	-.053
	(.047)	(.14)
Year FE	✓	✓
Insurer FE	✓	✓
R^2	.49	.49
Observations	581	548

Table 9: Financial Sophistication. Panel regressions at the contract-year level *Contract-level net flows* is contract net flows normalized by contract total account value. *Lagged reserves* is insurer beginning-of-year level of reserves normalized by insurer total account value. *Avg account value RANGE* is a dummy variable equal to one if the contract average account value (calculated as contract total account value divided by number of investors) lies in *RANGE*. All regressions include these non-interacted dummy variables in addition to their interaction with lagged reserves. Colums 1 and 3 include insurer and year fixed effects. Colums 2 and 4 include insurer-year fixed effects. Columns 1 and 2 show OLS regressions. Columns 3 and 4 show IV regressions in which the insurer's beginning-of-year reserve ratio is instrumented using the insurer's asset return in the previous year. All regressions are weighted by the contract share in aggregate account value in the current year. Standard errors two-way clustered by insurer and year are reported in parenthesis. ***, **, and * mean statistically significant at the 1%, 5%, and 10% levels, respectively.

	Contract-level net flows			
	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)
Lagged reserves x (Avg account value 0–50 k€)	-.059		-.29	
	(.17)		(.4)	
Lagged reserves x (Avg account value 50–250 k€)	.014	.13	-.13	.25*
	(.17)	(.076)	(.29)	(.15)
Lagged reserves x (Avg account value 250+ k€)	.36*	.41***	.43	.76***
	(.13)	(.0031)	(.26)	(.22)
Avg account value bin FE	✓	✓	✓	✓
Year FE	✓		✓	
Insurer FE	✓		✓	
Insurer-Year FE		✓		✓
R^2	.13	.16	.13	.16
Observations	7,272	7,272	7,268	7,268

Appendix (For Online Publication)

A Proofs

A.1 Demand shocks

All proofs are derived in the presence of demand shocks. Formally, we replace ξ_j in the demand function (6) with a random $\xi_{j,t-1}$ representing the preference for intermediary j in period t. $\xi_{j,t-1}$ is realized at the end of period $t-1$, hence the time index is $t-1$. We assume $\xi_{j,t}$ follows a random walk: $E_{t-1}[\xi_{j,t}] = \xi_{j,t-1}$. Similar to our assumption on asset returns, we assume demand shocks have bounded support, i.e., $|\xi_{j,t} - \xi_{j,t-1}| \leq \sigma$ for all j and t. Moreover, we assume no demand shocks after some period T, i.e., $\xi_{j,t} = \xi_{j,t-1}$ for $t > T$. We assume demand shocks are uncorrelated with contemporaneous asset return shocks: $Cov_{t-1}(x_{j,t}, \xi_{j,t}) = 0$. We denote the vector of demand shocks as $\xi_t = (\xi_{0,t}, \ldots, \xi_{J,t})$, and the history of demand shocks as $\xi^t = (\xi_1, \ldots, \xi_t)$. We assume that (x^T, ξ^T) has a finite support.

Contract return in period t is contingent on all information observable at the end of period t, that is, $y_{j,t}$ is a function of (x^t, ξ^t). Demand for contract j in period t can be written as function of the collection of expected utility of all contracts' returns, $V_{j,t-1}\big(\{E_{t-1}[u(y_{k,t})]\}_{k=0,\ldots,J}\big)$, where:

$$V_{j,t-1}\big(\{U_k\}_{k=0,\ldots,J}\big) = \frac{\exp\{\alpha\, U_j + \xi_{j,t-1}\}}{\sum_{k=0}^{J} \exp\{\alpha\, U_k + \xi_{k,t-1}\}}. \tag{A.1}$$

A.2 Proof of Proposition 1

The intermediary chooses a contract return policy $\big\{y_{j,t}(x^t, \xi^t)\big\}_{t\geq 1}$ to maximize expected discounted profit

$$\sum_{t=1}^{+\infty} \frac{1}{(1+r)^t}\, \phi\, E_0\Big[V_{j,t-1}\big(\{E_{t-1}[u(y_{k,t})]\}_{k=0,\ldots,J}\big)\Big], \tag{A.2}$$

where we omit arguments (x^t, ξ^t) in $y_{j,t}$. This maximization problem is subject to the intertemporal budget constraint

$$\sum_{t=1}^{+\infty} \frac{1}{(1+r)^t} \left[(x_{j,t} - y_{j,t} - \phi)\, V_{j,t-1}\big(\{E_{t-1}[u(y_{k,t})]\}_{k=0,\ldots,J}\big) \prod_{s=t+1}^{+\infty} \frac{1 + x_{j,s}}{1+r}\right] \geq 0, \quad \forall (x^T, \xi^T). \tag{A.3}$$

(A.2) is obtained by plugging per-period profit (3) into intertemporal profit (4). (A.3) is obtained by plugging profit (3) into the sequential budget constraint (1), consolidating the budget constraint intertemporally, and using the transversality condition (2). (A.3) must hold for all histories (x^T, ξ^T).

We denote by $\lambda_j(x^T, \xi^T)$ the Lagrange multiplier of (A.3) divided by the probability of history (x^T, ξ^T). Therefore, the Lagrangian associated with intermediary j's problem is equal to (A.2) plus the time-0 expectation of $\lambda_j(x^T, \xi^T)$ times the LHS of (A.3). The first-order condition with respect to $y_{j,t}(x^t, \xi^t)$ is

$$\left(\phi + E_{t-1}\left[\lambda_j(x_{j,t} - y_{j,t} - \phi)\prod_{s=t+1}^{+\infty}\frac{1+x_{j,s}}{1+r}\right]\right) V_{j,t-1}^{(j)}\big(\{E_{t-1}[u(y_{k,t})]\}_{k=0,\ldots,J}\big)u'(y_{j,t})$$
$$- E_t\left[\lambda_j \prod_{s=t+1}^{+\infty}\frac{1+x_{j,s}}{1+r}\right] V_{j,t-1}\big(\{E_{t-1}[u(y_{k,t})]\}_{k=0,\ldots,J}\big) = 0, \quad \text{(A.4)}$$

where we continue to omit arguments (x^t, ξ^t) in $y_{j,t}$ and arguments (x^T, ξ^T) in λ_j, $V_{j,t-1}^{(k)} = \partial V_{j,t-1}/\partial U_k$ denotes the partial derivative of demand, and $V_{j,t-1}^{(k\ell)} = \partial^2 V_{j,t-1}/(\partial U_k \partial U_\ell)$ the cross-derivative.

We write asset returns and demand shocks as follows:

$$x_{j,t} = r + \sigma\,\epsilon_{j,t}, \quad \text{(A.5)}$$
$$\xi_{j,t} = \xi_j + \sigma\,\zeta_{j,t} = \xi_j + \sigma\sum_{s=1}^{t}\eta_{j,s}, \quad \text{(A.6)}$$

for $j = 0, \ldots, J$ and $t \geq 1$, where $\sigma > 0$, $\epsilon_{j,t}$, $\zeta_{j,t}$, and $\eta_{j,t}$ are realized at the end of period t, $E_{t-1}[\epsilon_{j,t}] = E_{t-1}[\eta_{j,t}] = 0$, $\epsilon_{j,t}$ and $\eta_{j,t}$ have bounded support, and $\epsilon_{j,t} = \eta_{j,t} = 0$ for $t > T$. We solve the model using a first-order approximation for small shocks to asset returns and demand. We guess that, when σ goes to zero,

$$y_{j,t} = y_{j,t}^0 + \sigma\,y_{j,t}' + O(\sigma^2), \quad \text{(A.7)}$$
$$\lambda_j = \lambda_j^0 + \sigma\,\lambda_j' + O(\sigma^2), \quad \text{(A.8)}$$

where $y_{j,t}^0$ and $y_{j,t}'$ are functions of (ϵ^t, η^t), λ_j^0 and λ_j' are functions of (ϵ^T, η^T), and $O(\sigma^2)$ denote functions of $(\epsilon^T, \eta^T, \sigma)$ at the order of σ^2, that is, there exist $K > 0$ and $\overline{\sigma} > 0$ such that $O(\sigma^2) \leq K\sigma^2$ for all (ϵ^T, η^T) and $\sigma < \overline{\sigma}$.

We determine $y_{j,t}^0$ and λ_j^0 by letting σ go to zero in the intertemporal budget constraint (A.3) and first-order condition (A.4). The latter yields

$$\Big(\phi + \lambda_j^0(r - y_{j,t}^0 - \phi)\Big)V_j^{(j)}\big(\{u(y_{k,t}^0)\}_{k=0,\ldots,J}\big)u'(y_{j,t}^0) - \lambda_j^0 V_j\big(\{u(y_{k,t}^0)\}_{k=0,\ldots,J}\big) = 0, \tag{A.9}$$

where $V_j(.) \equiv V_{j,0}(.)$ denotes the demand function when demand shocks $\zeta_{k,t}$ are set to zero for all k. Since (A.9) does not depend on t, $y_{k,t}^0$ does not depend on t for all k. Therefore, we denote $y_{k,t}^0$ by y_k^0. Letting $\sigma \to 0$ in (A.3), we obtain

$$\sum_{t=1}^{+\infty} \frac{1}{(1+r)^t}(r - y_j^0 - \phi)V_j\big(\{u(y_k^0)\}_{k=0,\ldots,J}\big) = 0. \tag{A.10}$$

The solution to (A.10) is symmetric across intermediaries, and is given by

$$y_j^0 = r - \phi. \tag{A.11}$$

Substituting y_j^0 into (A.9), we obtain

$$\lambda_j^0 = \phi \frac{V_j^{(j)}}{V_j} u', \tag{A.12}$$

where we omit argument y_j^0 in $u(.)$ and its derivatives, and we omit argument $\{u(y_k^0)\}_{k=0,\ldots,J}$ in $V_j(.)$ and its derivatives.

We determine $y'_{j,t}$ and λ'_j by calculating first-order approximations of the intertemporal budget constraint (A.3) and first-order condition (A.4). Let us first write down a first-order approximation of the demand function (A.1):

$$\begin{aligned} V_{j,t-1}\big(\{E_{t-1}[u(y_{k,t})]\}_{k=0,\ldots,J}\big) &= V_j\big(\{u(y_{k,t}^0)\}_{k=0,\ldots,J}\big) \\ &+ \sigma \sum_{k=0}^{J} V_j^{(k)}\big(\{u(y_{k,t}^0)\}_{k=0,\ldots,J}\big)u'(y_{j,t}^0)\Big(E_{t-1}[y'_{j,t}] + \frac{\zeta_{j,t-1}}{\alpha u'(y_{j,t}^0)}\Big) + O(\sigma^2). \end{aligned} \tag{A.13}$$

The analogous approximation holds for $V_{j,t-1}^{(j)}$. A first-order approximation of the budget constraint (A.3) gives

$$\sum_{t=1}^{+\infty} \frac{1}{(1+r)^t}\big[(\epsilon_{j,t} - y'_{j,t})V_j\big] = 0. \tag{A.14}$$

We denote $y_{j,t}^s = E_s[y'_{j,t}] - E_{s-1}[y'_{j,t}]$ as the time-s innovation of $y'_{j,t}$ for all $1 \le s \le t$. Calculating

$E_s[.]$ of (A.14) minus $E_{s-1}[.]$ of (A.14), we obtain

$$\sum_{t=s}^{+\infty} \frac{y^s_{j,t}}{(1+r)^{t-s}} = \epsilon_{j,s}, \qquad s \leq t. \tag{A.15}$$

A first-order approximation of the first-order condition (A.4) gives

$$\phi V_j^{(j)} u'' y'_{j,t} - \lambda_j^0 V_j^{(j)} u' E_{t-1}[y'_{j,t}] + \phi \sum_{k=0}^{J} V_j^{(jk)} (u')^2 \Big(E_{t-1}[y'_{k,t}] + \frac{\zeta_{k,t-1}}{\alpha u'} \Big) \\ - \lambda_j^0 \sum_{k=0}^{J} V_j^{(k)} u' \Big(E_{t-1}[y'_{k,t}] + \frac{\zeta_{k,t-1}}{\alpha u'} \Big) - E_t[\lambda'_j] V_j = 0. \tag{A.16}$$

Using (A.12), we substitute λ_j^0 into (A.16). We then divide by $\phi V_j^{(j)} u''$, to obtain

$$y'_{j,t} + \frac{V_j^{(j)}}{V_j} \frac{(u')^2}{-u''} E_{t-1}[y'_{j,t}] + \sum_{k=0}^{J} \Big(\frac{V_j^{(k)}}{V_j} - \frac{V_j^{(jk)}}{V_j^{(j)}} \Big) \Big(\frac{(u')^2}{-u''} E_{t-1}[y'_{k,t}] + \frac{u'}{-\alpha u''} \zeta_{k,t-1} \Big) = \frac{E_t[\lambda'_j] V_j}{\phi V_j^{(j)} u''}. \tag{A.17}$$

We denote $\lambda_j^s = E_s[\lambda'_j] - E_{s-1}[\lambda'_j]$ the time-s innovation of λ'_j for all $1 \leq s \leq T$. Calculating (A.17) minus $E_{t-1}[.]$ of (A.17), we obtain

$$y^t_{j,t} = \frac{V_j}{\phi V_j^{(j)} u''} \lambda_j^t. \tag{A.18}$$

Calculating $E_s[.]$ of (A.17) minus $E_{s-1}[.]$ of (A.17) for $s < t$, we obtain

$$\Big(1 + \frac{V_j^{(j)}}{V_j} \frac{(u')^2}{-u''} \Big) y^s_{j,t} + \sum_{k=0}^{J} \Big(\frac{V_j^{(k)}}{V_j} - \frac{V_j^{(jk)}}{V_j^{(j)}} \Big) \Big(\frac{(u')^2}{-u''} y^s_{k,t} + \frac{u'}{-\alpha u''} \eta_{k,s} \Big) = \frac{V_j}{\phi V_j^{(j)} u''} \lambda_j^s, \qquad s < t. \tag{A.19}$$

The derivatives of the logit demand function are

$$\frac{V_j^{(j)}}{V_j} = \alpha(1-s_j), \qquad \frac{V_j^{(jj)}}{V_j^{(j)}} = \alpha(1-2s_j), \qquad \frac{V_j^{(k)}}{V_j} = -\alpha s_k, \qquad \frac{V_j^{(jk)}}{V_j^{(j)}} = -\alpha s_k \Big(1 - \frac{s_j}{1-s_j} \Big), \qquad k \neq j,$$

where

$$s_j \equiv V_j = \frac{\exp\{\xi_j\}}{\sum_{k=0}^{J} \exp\{\xi_k\}} \tag{A.20}$$

is intermediary j's market share when all intermediaries offer the same contract return and all demand shocks are set to zero. We use these expressions, and (A.18) to substitute λ_j^s on the right-hand side of (A.19). We obtain

$$y_{j,t}^s = \frac{\gamma_j}{\alpha+\gamma_j} y_{j,s}^s + \frac{\alpha\delta_j}{\alpha+\gamma_j}\sum_{k=0}^J s_k y_{k,t}^s - \frac{\frac{1}{u'}\delta_j}{\alpha+\gamma_j}\Big(\eta_{j,s} - \sum_{k=1}^J s_k\eta_{k,s}\Big), \qquad s<t, \tag{A.21}$$

where

$$\gamma_j = \frac{1}{1+\frac{s_j^2}{1-s_j}}\frac{-u''(r-\phi)}{(u'(r-\phi))^2}, \tag{A.22}$$

$$\delta_j = \frac{1}{1+\frac{s_j^2}{1-s_j}}\frac{s_j}{1-s_j}. \tag{A.23}$$

To determine $\sum_{k=0}^J s_k y_{k,t}^s$, we first note that $y_{0,t}^s = 0$ for $s<t$. Then, multiplying (A.21) by s_j, summing over $j=1,\ldots,J$, and collecting the terms $y_{k,t}^s$, we obtain

$$(1-A)\sum_{k=0}^J s_k y_{k,t}^s = \sum_{k=1}^J \frac{\gamma_k s_k}{\alpha+\gamma_k} y_{k,s}^s - \sum_{k=1}^J \frac{\frac{1}{u'}\delta_k s_k}{\alpha+\gamma_k}\Big(\eta_{k,s} - \sum_{\ell=1}^J s_\ell \eta_{\ell,s}\Big), \qquad s<t. \tag{A.24}$$

where $A=\sum_{k=1}^J \frac{\alpha\delta_k s_k}{\alpha+\gamma_k}$. Substituting the expression of $\sum_{k=0}^J s_k y_{k,t}^s$ given by (A.24) into (A.21), and collecting the terms $\eta_{k,s}$, we obtain

$$y_{j,t}^s = \frac{\gamma_j}{\alpha+\gamma_j} y_{j,s}^s + \frac{1}{1-A}\frac{\alpha\delta_j}{\alpha+\gamma_j}\sum_{k=1}^J \frac{\gamma_k s_k}{\alpha+\gamma_k} y_{k,s}^s - \frac{\frac{1}{u'}\delta_j}{\alpha+\gamma_j}\big(\eta_{j,s} - \overline{\eta}_s\big), \qquad s<t, \tag{A.25}$$

where

$$\overline{\eta}_s = \frac{1}{1-A}\sum_{k=1}^J \Big(1-\frac{\alpha\delta_k}{\alpha+\gamma_k}\Big) s_k \eta_{k,s}. \tag{A.26}$$

Substituting the expression of $y_{j,t}^s$ given by (A.25) into the budget constraint (A.15), we obtain

$$\frac{\alpha+\frac{1+r}{r}\gamma_j}{\alpha+\gamma_j} y_{j,s}^s + \frac{1}{1-A}\frac{\frac{1}{r}\alpha\delta_j}{\alpha+\gamma_j}\sum_{k=1}^J \frac{\gamma_k s_k}{\alpha+\gamma_k} y_{k,s}^s = \epsilon_{j,s} + \frac{\frac{1}{ru'}\delta_j}{\alpha+\gamma_j}\big(\eta_{j,s} - \overline{\eta}_s\big). \tag{A.27}$$

To determine $\sum_{k=1}^J \frac{\gamma_k s_k}{\alpha+\gamma_k} y_{k,s}^s$, we multiply both sides of (A.27) by $\frac{\gamma_j s_j}{\alpha+\frac{1+r}{r}\gamma_j}$, sum over $j=1,\ldots,J$, and rearrange terms, to obtain

$$\frac{1-B}{1-A}\sum_{k=1}^J \frac{\gamma_k s_k}{\alpha+\gamma_k} y_{k,s}^s = \widehat{\epsilon}_s + \frac{1}{\alpha u'}\widehat{\eta}_s, \tag{A.28}$$

where $B = \sum_{k=1}^{J} \frac{\alpha\delta_k s_k}{\alpha + \frac{1+r}{r}\gamma_k}$, and

$$\widehat{\epsilon}_s = \sum_{k=1}^{J} \frac{\gamma_k s_k}{\alpha + \frac{1+r}{r}\gamma_k} \epsilon_{k,s}, \tag{A.29}$$

$$\widehat{\eta}_s = \sum_{k=1}^{J} \frac{\gamma_k s_k}{\alpha + \frac{1+r}{r}\gamma_k} \frac{\frac{\alpha}{r}\delta_k}{\alpha + \gamma_k} \big(\eta_{k,s} - \overline{\eta}_s\big). \tag{A.30}$$

Substituting (A.28) back into (A.27), we obtain

$$y_{j,s}^s = \frac{\alpha + \gamma_j}{\alpha + \frac{1+r}{r}\gamma_j} \epsilon_{j,s} - \frac{\frac{1}{r}\alpha\delta_j}{\alpha + \frac{1+r}{r}\gamma_j} \frac{1}{1-B} \widehat{\epsilon}_s + \frac{\frac{1}{ru'}\delta_j}{\alpha + \frac{1+r}{r}\gamma_j} \Big(\eta_{j,s} - \overline{\eta}_s - \frac{1}{1-B}\widehat{\eta}_s\Big). \tag{A.31}$$

Substituting the expression of $y_{j,s}^s$ given by (A.31) into (A.25) we obtain

$$y_{j,t}^s = \frac{\gamma_j}{\alpha + \frac{1+r}{r}\gamma_j} \epsilon_{j,s} + \frac{\alpha\delta_j}{\alpha + \frac{1+r}{r}\gamma_j} \frac{1}{1-B} \widehat{\epsilon}_s - \frac{\frac{1}{u'}\delta_j}{\alpha + \frac{1+r}{r}\gamma_j} \Big(\eta_{j,s} - \overline{\eta}_s - \frac{1}{1-B}\widehat{\eta}_s\Big), \qquad s < t. \tag{A.32}$$

Finally, we use (A.11), (A.31), and (A.32) to calculate $y_{j,t} = y_{j,t}^0 + \sigma \sum_{s=1}^{t} y_{j,t}^s + O(\sigma^2)$. We obtain:

$$y_{j,t} = r - \phi + \sum_{s=1}^{t-1} \frac{\gamma_j}{\alpha + \frac{1+r}{r}\gamma_j}(x_{j,s} - r) + \frac{\alpha + \gamma_j}{\alpha + \frac{1+r}{r}\gamma_j}(x_{j,t} - r) + f_{j,t}(\overline{x}^t - r, \eta^t) + O(\sigma^2), \tag{A.33}$$

where $f_{j,t}(.)$ is a function of the history of average asset return shocks $\overline{x}^t - r$ and demand shocks η^t:

$$f_{j,t}(\overline{x}^t - r, \eta^t) = \frac{\alpha\delta_j}{\alpha + \frac{1+r}{r}\gamma_j} \frac{\sum_{k=1}^{J} \frac{\gamma_k s_k}{\alpha + \frac{1+r}{r}\gamma_k}}{1 - \sum_{k=1}^{J} \frac{\alpha\delta_k s_k}{\alpha + \frac{1+r}{r}\gamma_k}} \left(\sum_{s=1}^{t-1}(\overline{x}_s - r) - \frac{1}{r}(\overline{x}_t - r)\right) + g_{j,t}(\eta^t) \tag{A.34}$$

$$\overline{x}_t = \sum_{k=1}^{J} \frac{\frac{\gamma_k s_k}{\alpha + \frac{1+r}{r}\gamma_k}}{\sum_{\ell=1}^{J} \frac{\gamma_\ell s_\ell}{\alpha + \frac{1+r}{r}\gamma_\ell}} x_{k,t}, \tag{A.35}$$

and s_j, γ_j, and δ_j are given by (A.20), (A.22), and (A.23), respectively.

A.3 Sufficient condition for binding regulatory constraint

Suppose that instead of holding with an equality ("="), we model the regulatory constraint (3) as an inequality ("≤"). This modeling would be closer to the regulation of euro contracts (see

Section 3.1). In this case, we show that a sufficient condition for this constraint to be binding is:

$$\phi\,\alpha < 1. \tag{A.36}$$

Suppose (A.36) holds. Let $\kappa \in (1, \frac{1}{\phi\,\alpha})$. To show that the regulatory constraint is binding, we need to show that intermediaries can increase their intertemporal profit by violating the constraint. Consider a marginal increase $d\phi_{j,t} > 0$ in the fraction of account value that goes to intermediary j in period t and a reduction in contract return $dy_{j,t} = -\kappa d\phi_{j,t}$. Investor demand in period t changes by $-\kappa d\phi_{j,t}[V_j^{(j)} + O(\sigma)]$. The budget constraint (A.3) is strictly relaxed because $\kappa > 1$ and $x_{j,t} - y_{j,t} - \phi = O(\sigma)$. The intermediary's intertemporal profit (A.2) changes by

$$\frac{1}{(1+r)^t}\left[V_j d\phi_{j,t} - \phi\,\kappa V_j^{(j)} d\phi_{j,t}\right] = \frac{1}{(1+r)^t}\left[1 - \phi\,\kappa\,\alpha\,(1-s_j)\right]V_j d\phi_{j,t},$$

which is positive, because $\kappa < \frac{1}{\phi\,\alpha}$ and $1 - s_j < 1$. Therefore, the regulatory constraint is binding.

(A.36) is arguably satisfied in our empirical setup since we estimate $\alpha \simeq 0$.

A.4 Proof of Proposition 2

Reserves evolve according to $R_{j,t} = (1 + x_{j,t})R_{j,t-1} + (x_{j,t} - \phi - y_{j,t})V_{j,t}$ with $R_{j,0} = 0$. Therefore, a first-order approximation of $R_{j,t}$ is

$$R_{j,t} = \sigma\,h_{j,t}V_j + O(\sigma^2), \tag{A.37}$$

where

$$\begin{aligned}
h_{j,t} &= \sum_{s=1}^{t}(1+r)^{t-s}(\epsilon_{j,s}-y'_{j,s}) \\
&= \sum_{s=1}^{t}(1+r)^{t-s}(\epsilon_{j,s}-y^{s}_{j,s}) - \sum_{s=1}^{t}(1+r)^{t-s}\sum_{\tau=1}^{s-1}y^{\tau}_{j,s} \\
&= \sum_{s=1}^{t}(1+r)^{t-s}(\epsilon_{j,s}-y^{s}_{j,s}) - \sum_{s=1}^{t}\sum_{\tau=s+1}^{t}(1+r)^{t-\tau}y^{s}_{j,\tau} \\
&= \sum_{s=1}^{t}(1+r)^{t-s}\Big(\epsilon_{j,s}-\sum_{\tau=s}^{t}(1+r)^{s-\tau}y^{s}_{j,\tau}\Big) \\
&= \sum_{s=1}^{t}(1+r)^{t-s}\sum_{\tau=t+1}^{+\infty}(1+r)^{s-\tau}y^{s}_{j,\tau} \\
h_{j,t} &= \frac{1}{r}\sum_{s=1}^{t}y^{s}_{j,t+1}, \qquad \text{(A.38)}
\end{aligned}$$

where we move from the first line to the second line using $y'_{j,s} = \sum_{\tau=1}^{s} y^{\tau}_{j,s}$, to the third line by switching indices s and τ, to the fourth line by putting the two sums over s together, to the fifth line using the budget constraint (A.15), and to the sixth line using that $y^{s}_{j,\tau}$ does not depend on τ for all $\tau > s$, so $y^{s}_{j,\tau}$ can be replaced by $y^{s}_{j,t+1}$.

We have

$$\begin{aligned}
y_{j,t}-(r-\phi) &= \sigma y'_{j,t}+O(\sigma^2) \;=\; \sigma\sum_{s=1}^{t-1}y^{s}_{j,t}+\sigma y^{t}_{j,t}+O(\sigma^2) \\
&= \frac{r}{1+r}\Big((1+r)\frac{R_{j,t-1}}{V_j}+\sigma\epsilon_{j,t}\Big)+\frac{1}{1+r}\,\frac{\alpha}{\alpha+\frac{1+r}{r}\gamma_j}\sigma\epsilon_{j,t} \\
&\quad -\frac{\frac{1}{r}\alpha\delta_j}{\alpha+\frac{1+r}{r}\gamma_j}\frac{1}{1-B}\sigma\widehat{\epsilon}_t+\frac{\frac{1}{r}\delta_j}{\alpha+\frac{1+r}{r}\gamma_j}\sigma\Big(\eta_{j,t}-\overline{\eta}_t-\frac{1}{1-B}\widehat{\eta}_t\Big)+O(\sigma^2)
\end{aligned}$$

where we move from the first line to the second line using (A.37) and (A.38) to substitute $\sum_{s=1}^{t-1} y^{s}_{j,t}$, and using (A.31) to substitute $y^{t}_{j,t}$. Thus:

$$\begin{aligned}
&y_{j,t} = \\
&\quad r-\phi+\frac{1}{1+r}\,\frac{\alpha}{\alpha+\frac{1+r}{r}\gamma_j}(x_{j,t}-r)+\frac{r}{1+r}\left(\frac{R_{j,t-}}{V_{j,t-1}}-r\right)+\mu_j(\overline{x}_t-r)+\omega_j\widetilde{\eta}_{j,t}+O(\sigma^2), \quad \text{(A.39)}
\end{aligned}$$

where

$$\mu_j \quad = \quad -\frac{\frac{1}{r}\alpha\delta_j}{\alpha+\frac{1+r}{r}\gamma_j}\frac{\sum_{k=1}^{J}\frac{\gamma_k s_k}{\alpha+\frac{1+r}{r}\gamma_k}}{1-\sum_{k=1}^{J}\frac{\delta_k s_k}{\alpha+\frac{1+r}{r}\gamma_k}}, \tag{A.40}$$

s_j, γ_j, and δ_j are given by (A.20), (A.22), and (A.23), respectively, $\overline{x}_t$ is a weighted average of asset returns $x_{k,t}$ over $k = 1, \ldots, J$ defined in (A.35), $\omega_j = \frac{1}{r}\delta_j / \left(\alpha + \frac{1+r}{r}\gamma_j\right)$, $\widetilde{\eta}_{j,t} = \eta_{j,t} - \overline{\eta}_t - \frac{1}{1-B}\widehat{\eta}_t$, and $\overline{\eta}_t$ and $\widehat{\eta}_t$ are weighted averages of demand shocks $\eta_{k,t}$ over $k = 1, \ldots, J$ defined in (A.26) and (A.30), respectively.

A.5 Proof of Implication 1

We consider the case where the number of intermediaries, J, is large, so that market shares, s_j, are small. Formally, we assume there exists $\overline{s} > 0$ such that $s_j < \overline{s}J^{-1}$ for all $J > 1$. Let $O(J^{-1})$ denote functions of the order of J^{-1}, that is, there exist $K > 0$ and $\overline{J}$ such that $O(J^{-1}) \leq KJ^{-1}$ for all $J > \overline{J}$. It follows from (A.22) that $\gamma_j = -u''/(u')^2 + O(J^{-2})$, from (A.23) that $\delta_j = O(J^{-1})$, and from (A.40) that $\mu_j = \mu + O(J^{-1})$. Therefore, (A.39) can be rewritten as:

$$y_{j,t} = cste + \frac{1}{1+r}\,\frac{\alpha}{\alpha+\frac{1+r}{r}\gamma}x_{j,t} + \frac{r}{1+r}\,\mathcal{R}_{j,t^-} + \varepsilon_{j,t} + O(\sigma^2), \tag{A.41}$$

where $\gamma \equiv -u''/(u')^2 = -u''/u'$ using the normalization $u'(r-\phi) = 1$, and

$$\varepsilon_{j,t} = \mu\,\overline{\epsilon}_t + \omega_j\widetilde{\eta}_{j,t} + O(J^{-1})x_{j,t} + O(J^{-1})(\overline{x}_t - r). \tag{A.42}$$

Since demand shocks entering into the expression of $\widetilde{\eta}_{j,t}$ are uncorrelated with asset return $x_{j,t}$, the covariance between $x_{j,t}$ and $\varepsilon_{j,t}$ conditional on $\overline{\epsilon}_t$ is $O(J^{-1})$. Since $\mathcal{R}_{j,t^-} = \mathcal{R}_{j,t-1} + x_{j,t}(1+\mathcal{R}_{j,t-1})$ and $\mathcal{R}_{j,t-1}$ is uncorrelated with $\varepsilon_{j,t}$, the covariance between $\mathcal{R}_{j,t^-}$ and $\varepsilon_{j,t}$ conditional on $\overline{\epsilon}_t$ is also $O(J^{-1})$. Therefore, the cross-sectional covariance between $\left(x_{j,t}, \mathcal{R}_{j,t^-}\right)$ and $\varepsilon_{j,t}$ is $O(J^{-1})$.

A.6 Proof of Implication 2

A first-order approximation of log demand of intermediary j in period t is

$$\log(V_{j,t-1}) = \log(V_j) + \frac{V_j^{(j)}}{V_j}\Big(E_{t-1}[y_{j,t}] - (r-\phi) + \sigma\eta_{j,t-1}\Big) + O(\sigma^2). \tag{A.43}$$

The expectation of contract return (A.39) is equal to

$$E_{t-1}[y_{j,t}] = r - \phi + r\,\mathcal{R}_{j,t-1} + O(\sigma^2), \tag{A.44}$$

where we have used $\mathcal{R}_{j,t^-} = (1 + x_{j,t})\mathcal{R}_{j,t-1} + x_{j,t}$ and $E_{t-1}[x_{j,t}] = r$. Plugging (A.44) into (A.43) and calculating the derivative of logit demand, we obtain

$$\log(V_{j,t-1}) = \log(V_j) + \alpha\,(1 - s_j)\,r\,\mathcal{R}_{j,t-1} + \alpha\,(1 - s_j)\,\sigma\,\eta_{j,t-1} + O(\sigma^2).$$

When market shares are small, we obtain that

$$\log(V_{j,t-1}) = \log(V_j) + \alpha\ r\,\mathcal{R}_{j,t-1} + \alpha\,\sigma\,\eta_{j,t-1} + O(\sigma^2) + O(\sigma J^{-1}). \tag{A.45}$$

A.7 Estimation of the flow-reserves relation with demand shocks

OLS estimation of the flow-reserves relation (A.45) is unbiased when there are no demand shocks ($\eta \equiv 0$). However, in the presence of demand shocks, OLS estimation is biased because $\mathcal{R}_{j,t-1}$ is correlated with $\eta_{j,t-1}$. Indeed, the contract return policy (A.39) implies $Cov_{t-2}(y_{j,t-1}, \eta_{j,t-1}) > 0$. Using the fact that η is uncorrelated with asset returns and past demand shocks, it follows from the budget constraint (1) that $Cov_{t-2}(\mathcal{R}_{j,t-1}, \eta_{j,t-1}) < 0$. Therefore, the OLS estimate of the flow-reserves relation is biased downard in the presence of demand shocks.

Let us now show that $\mathcal{R}_{j,t-1}$ is a valid instrument for $x_{j,t-1}$. The relevance condition is satisfied, because it follows from the budget constraint (1) that $Cov_{t-2}(x_{j,t-1}, \mathcal{R}_{j,t-1}) > 0$. The exclusion restriction is satisfied, because $Cov_{t-2}(x_{j,t-1}, \eta_{j,t-1}) = 0$. Thus, the IV estimate of the flow-reserves relation (A.45) using lagged asset return to instrument for reserves is unbiased.

A.8 Arbitrageurs

Contract return (11) can be rewritten as

$$y_{j,t} \simeq r - \phi + \frac{\alpha + \gamma_j}{\alpha + \frac{1+r}{r}\gamma_j}\,(x_{j,t} - r) + r\,\mathcal{R}_{j,t-1} + \mu_j\,(\overline{x}_t - r),$$

where $\mathcal{R}_{j,t-1} = R_{j,t-1}/V_{j,t-1}$ is the beginning-of-period reserve ratio, and $\mu_j < 0$. Consider the hedged, zero-cost portfolio that goes long one euro in contract j, short $(1 - \tau)\frac{\alpha+\gamma_j}{\alpha+\frac{1+r}{r}\gamma_j}$ euros in intermediary j's asset portfolio, long $(1 - \tau)|\mu_j|$ euros in the weighted-average intermediary portfolio, and borrows $1-(1-\tau)\frac{\alpha+\gamma_j}{\alpha+\frac{1+r}{r}\gamma_j}+(1-\tau)|\mu_j|$ at the risk-free rate. The return on the long position in the euro contract is taxed at rate τ. The return on the long position in the weighted-average intermediary portfolio is taxed if the position cannot be netted against the short position in intermediary j's asset, but the part of the long position that can be netted is not taxed. We make the conservative assumption (in the sense that it maximizes the profitability of the arbitrage strategy) that the long position in the weighted-average intermediary portfolio can be fully netted against the short position in intermediary j's asset, and thus is not taxed. The arbitrage profit is equal to

$$\begin{aligned}\pi_{j,t}^{arb} &\simeq (1-\tau)y_{j,t}-(1-\tau)\frac{\alpha+\gamma_j}{\alpha+\frac{1+r}{r}\gamma_j}x_{j,t}+(1-\tau)|\mu_j|\overline{x}_t-\left(1-(1-\tau)\frac{\alpha+\gamma_j}{\alpha+\frac{1+r}{r}\gamma_j}+(1-\tau)|\mu_j|\right)r_f \\ &\simeq \left[1-(1-\tau)\frac{\alpha+\gamma_j}{\alpha+\frac{1+r}{r}\gamma_j}+(1-\tau)|\mu_j|\right](r-r_f)+(1-\tau)r\,\mathcal{R}_{j,t-1}-\tau r-(1-\tau)\phi. \quad \text{(A.46)}\end{aligned}$$

B Regulatory Framework for Reserves

Fund reserves are defined as the difference between the market value of funds' assets and total account value. Reserves are made up of three components:

1. Profit-sharing reserves At least 85% of financial income plus 90% of their technical income (or 100% if it is negative) must be distributed to investors. Financial income is equal to asset yield (dividends on non-fixed income securities plus yield on fixed income securities) plus realized gains and losses on non-fixed income securities. Technical income is equal to fees paid by investors minus operating costs. The amount distributed to investors is split into two parts: one part credited immediately to investors' accounts and another part credited to a reserve account called the profit-sharing reserve (*provision pour participation aux bénéfices*). The profit-sharing reserve account can only be used for future distribution to investor accounts. Therefore, profit-sharing reserves effectively belong to (current and future) investors. One key property of profit-sharing reserves is that they are are pooled across all contracts so that when an investor redeems her contract, she gives up her right to future distribution of the profit-sharing reserves. Conversely, when a new investor buys a contract, she shares in the existing profit-sharing reserves.[32]

2. Capitalization reserves Realized gains and losses on fixed income securities are credited to, or debited from, a reserve account called the capitalization reserve account (*réserve de capitalisation*). The capitalization reserve account can only be used to offset future losses on fixed income securities and cannot be credited to investors' accounts or to insurer income. Thus, capitalization reserves represent deferred financial income for the fund. Since at least 85% of the fund financial income must be distributed to investors, it implies that at least 85% of capitalization reserves effectively belong to (current and future) investors.[33]

3. Unrealized gains Unrealized capital gains are not booked as fund income. Therefore, accumulated capital gains and losses create a wedge between the market value and the book value of the fund assets. This constitutes the third reserves component.[34] Since unrealized gains represent

[32] By law, the insurer must distribute the profit-sharing reserve account to investors within eight years. In practice, this constraint is never binding. Profit-sharing reserves represent less than one year of contract returns on average, and two years and a half at the 99th percentile.

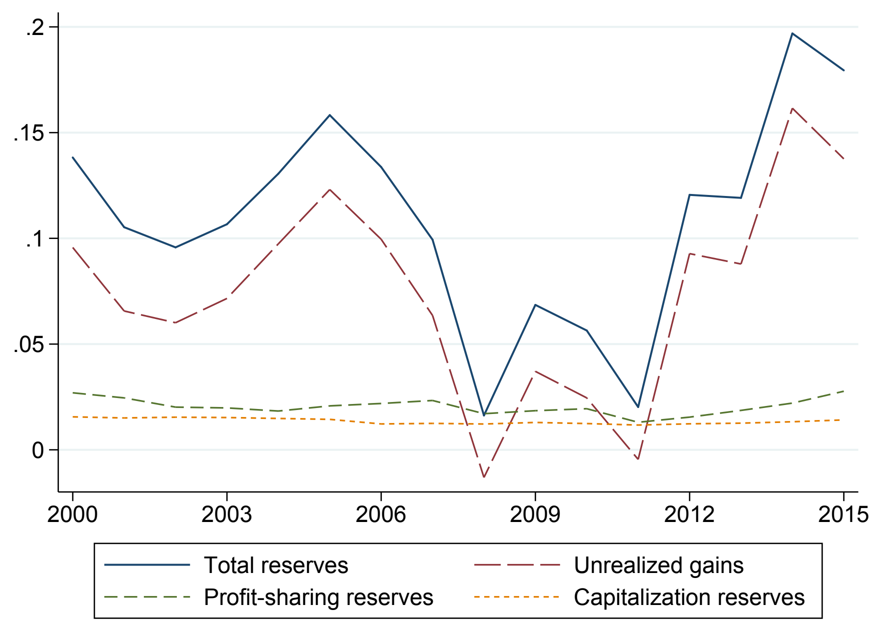
[33] Although, for accounting and regulatory purposes, capitalization reserves are booked as insurer equity.

[34] While unrealized capital gains are never booked as fund profit, there are two deviations from historical cost accounting principles that force insurers to recognize large unrealized losses. First, when an asset has "lasting and significant" unrealized capital losses, its book value is partially adjusted downwards through the creation of a provision

deferred fund financial income, at least 85% of their value effectively belong to (current and future) investors.

Summary statistics Reserves represent on average 10.9% of account value, of which about two-third are unrealized gains (7.5%) and one-third are accounting reserves (2.1% of profit-sharing reserves and 1.4% of capitalization reserves). Figure B.1 plots the time-series of aggregate reserves and its three sub-components as a fraction of account value.

Figure B.1: Reserves. The figure shows total reserves as a fraction of account value (solid blue) and the breakdown into the three components of reserves: unrealized gains (long dashed red); profit-sharing reserves (dashed green); and capitalization reserves (short dashed orange).



on the asset side of the balance sheet (*provision pour dépréciation durable*) to reflect the paper loss. This adjustment is booked as a realized loss. It thus increases unrealized gains (makes them less negative). If the return credited to contracts and to the insurer are held constant, this realized loss reduces the profit-sharing reserve account, and total reserves are not affected. The goal of this provision is to induce the insurer to reduce the return credited to investor accounts and thus reduce profit-sharing reserves by less than the realized loss, increasing total reserves.

The second deviation from historical cost accounting is that, when the market value of the fund portfolio of non-fixed income securities is less than the book value, the overall paper loss is recognized through a provision on the liability side of the balance sheet (*provision pour risque d'exigibilité*). This is booked as a loss. Therefore, if the return credited to contracts and to the insurer are held constant, this reduces the profit-sharing reserve account and thus total reserves. The goal of this provision is to induce the insurer to reduce the return credited to investor accounts and thus offset the reduction in the amount of reserves.

C Variables Construction

C.1 Regulatory filings

This section describes how we construct variables at the insurer-year level using the annual regulatory filings (*Dossiers Annuels*) from 1999 to 2015.

Account value `Provisions d'assurance vie à l'ouverture` (beginning-of-year account value) and `Provisions d'assurance vie à la clôture` (end-of-year account value) in C1V1–C1V3 statements summed over contract categories 1, 2, 4, 5, and 7, which is the set of contracts backed by the same pool of underlying assets and associated to the same pool of reserves. The main excluded contract categories are 8 and 9, which are unit-linked contracts.

Profit-sharing reserves `Provisions pour participations aux bénéfices et ristournes` in BILPV statement.

Capitalization reserves `Réserve de capitalisation` in C5P1 statement.

Unrealized gains Book value (`Valeur nette`) minus market value (`Valeur de réalisation`) of assets underlying life insurance contracts measured as `Placements représentatifs des provisions techniques` minus `Actifs représentatifs des unités de compte` in N3BJ statement.

Inflows `Sous-total primes nettes` in C1V1–C1V3 statements summed over contract categories 1, 2, 4, 5, and 7. It includes initial cash deposits at subscription and subsequent cash deposits in existing contracts. The inflow rate is calculated as inflow amount divided by beginning-of-year account value plus one half of net flows.

Outflows `Sinistres et capitaux payés` plus `Rachats payés` in C1V1–C1V3 statements summed over contract categories 1, 2, 4, 5, and 7. It includes partial and full redemptions, either voluntary or at death of investor. The outflow rate is calculated as outflow amount divided by beginning-of-year account value plus one half of net flows.

Contract return We calculate the value-weighted average contract return as the amount credited to investor accounts divided by beginning-of-year account value plus one half of net flows (i.e., we assume flows are uniformly distributed throughout the year and thus receive on average one half of the annual contract return). The amount credited to investor accounts is measured as `Intérêts techniques incorporés aux provisions d'assurance vie` plus `Participations aux bénéfices` plus `Intérêts techniques inclus dans exercice prestations` plus `Participations aux bénéfices incorporées dans exercice prestations` in C1V1–C1V3 statements summed over contract categories 1, 2, 4, 5, and 7.

Asset return We sum the three components of asset returns, which are reported separately in insurers' financial statement. First, `Produits des placements nets de charges` in C1V1–C1V3 statements summed over contract categories 1, 2, 4, 5, and 7, measures asset yield (dividends on non-fixed income securities plus yield on fixed income securities) and realized gains and losses on non-fixed income securities, net of operating costs. Second, the change in capitalization reserves account value reflects realized gains and losses on fixed income securities. Third, the change in unrealized gains captures measures unrealized gains. We calculate asset return as the sum of these three components divided by account value plus reserves.

C.2 Account value by cohort

We describe in this appendix how we estimate account value by cohort from insurer-level account value, inflows, and outflows, under parametric assumptions on the inflow rate and the outflow rate.

Regarding inflows, we assume investors only make one-off investments. They make an initial deposit when they buy a contract and never deposit additional funds at subsequent dates. Regarding outflows, we assume investors only proceed to full redemptions and that the redemption rate does not depend on contract age for a given insurer in a given year.

We call cohort (t_0, t_1) the set of investors who buy their contract in year t_0 and redeem it in year t_1, for $t_0 < t_1$. We denote $V_t(t_0, t_1)$ the account value of cohort (t_0, t_1) at the end of year t and by $V_t^+(t_0, t_1)$ and $V_t^-(t_0, t_1)$ their inflows and outflows, respectively, during year t. Under the maintained assumption that inflows and outflows are uniformly distributed throughout the year and are entitled to one half of the annual contract return, account value of cohort (t_0, t_1) evolves according to

$$V_{t_0-1}(t_0,t_1) = 0, \tag{C.1}$$

$$V_t(t_0,t_1) = (1+y_t)V_{t-1}(t_0,t_1) + (1+\frac{y_t}{2})(V_t^+(t_0,t_1) - V_t^-(t_0,t_1)), \quad t = t_0, ..., t_1-1, \tag{C.2}$$

$$V_{t_1}(t_0,t_1) = 0, \tag{C.3}$$

where y_t is the net-of-fees contract return. The assumption of no inflow after initial subscription writes

$$V_t^+(t_0,t_1) = 0, \qquad t > t_0. \tag{C.4}$$

The assumption of no partial redemption before exit writes

$$V_t^-(t_0,t_1) = 0, \qquad t < t_1. \tag{C.5}$$

The assumption of outflow rate independent of contract age at the insurer-year level writes

$$\frac{V_t^-(t_0,t)}{V_{t-1}(t_0)} = \frac{V_t^-}{V_{t-1}}, \qquad t > t_0. \tag{C.6}$$

We now describe the procedure to calculate account value by cohort.

Net-of-fees returns The data only reports gross-of-fees contract return. Since we observe beginning-of-year account value V_{t-1}, inflows V_t^+, outflows V_t^-, and end-of-year account value V_t, we back out the net-of-fees contract return y_t from the law of motion of total account value

$$V_t = (1+y_t)V_{t-1}(1+\frac{y_t}{2})(V_t^+ - V_t^-). \tag{C.7}$$

Birth-cohort-level account value Define a birth-cohort t_0 as the set of cohorts $\{(t_0,t_1) : t_1 > t_0\}$. Denoting by $T_0 = 1999$ and $T_1 = 2015$ the first year and last year when account value data are available, we redefine birth-cohort $T_0 - 1$ as the set of birth-cohorts $\{t_0 : t_0 \leq T_0 - 1\}$. We denote by $V_t(t_0)$, $V_t^+(t_0)$, and $V_t^-(t_0)$ the end-of-year, inflows, and outflows, respectively, of birth-cohort t_0.

$V_{T_0-1}(T_0-1)$ is observed in the data as beginning-of-year account value in year T_0. (C.4) implies that, for all $t_0 \geq T_0$, inflows of birth-cohort t_0 in year t_0 is $V_{t_0}^+(t_0) = V_{t_0}^+$, which is observed in the data as total outflow in year t_0.

Then, we compute birth-cohort-level end-of-year account value and outflows in all years $t \in [T_0, T_1]$ by forward iteration. Once we have computed birth-cohort-level end-of-year account value in year $t-1$, (C.5) and (C.6) imply that outflows of birth-cohort $t_0 < t$ in year t is $V_t^-(t_0) = \frac{V_{t-1}(t_0)}{V_{t-1}} V_t^-$, where the last term is total outflows in year t, which is observed in the data. End-of-year account value of birth-cohort $t_0 < t$ in year t is $V_t(t_0) = (1 + y_t)V_{t-1}(t_0) - (1 + \frac{y_t}{2})V_t^-(t_0)$. End-of-year account value of birth-cohort t in year t is $V_t(t) = (1 + \frac{y_t}{2})V_t^+(t)$.

Cohort-level account value For $t_1 \in [T_0, T_1]$, we redefine cohort $(T_0 - 1, t_1)$ as the set of cohorts $\{(t_0, t_1) : t_0 \leq T_0 - 1\}$. For $t_0 \in [T_0, T_1]$, we redefine cohort $(t_0, T_1 + 1)$ as the set of cohorts $\{(t_0, t_1) : t_1 \geq T_1 + 1\}$.

(C.5) implies that cohort-level outflows is $V_{t_1}^-(t_0, t_1) = V_{t_1}^-(t_0)$ for all $T_0 - 1 \leq t_0 < t_1 \leq T_1$. Then, we compute end-of-year account value for each cohort (t_0, t_1) in all year $t \in [t_0, t_1 - 1]$ by backward iteration. If $t_1 \leq T_1$, it follows from (C.2) and (C.3) that $V_{t_1-1}(t_0, t_1) = (1 + \frac{y_{t_1}}{2})V_{t_1}^-(t_0)/(1 + y_{t_1})$. If $t_1 = T_1 + 1$, $V_{T_1}(t_0, T_1 + 1) = V_{T_1}(t_0)$. Once we have computed the end-of-year account value of cohort (t_0, t_1) in year t, we use (C.3) to calculate it in year $t - 1$: $V_{t-1}(t_0, t_1) = V_t(t_0, t_1)/(1 + y_t)$ for all $t \in [t_0 + 1, t_1 - 1]$. Finally, for $t_0 \geq T_0$, it follows from (C.1) and (C.2) that inflows of cohort (t_0, t_1) in year t_0 is $V_{t_0}^+(t_0, t_1) = V_{t_0}(t_0, t_1)/(1 + \frac{y_{t_0}}{2})$.

D Reserves Dilution

This appendix presents the calculation of the mean reversion rate of the reserve ratio. The evolution of reserves is given by (1) as in the model. Replacing the regulatory constraint (3) by the actual regulatory constraint $\Pi_{j,t} = \pi x_t(V_{j,t-1} + R_{j,t-1})$, where $\pi = 0.15$, (1) can be rewritten:

$$(\mathcal{F}_{j,t} + 1 + y_{j,t})\mathcal{R}_{j,t} = x_{j,t} + (1 + x_{j,t})\mathcal{R}_{j,t-1} - y_{j,t} - \pi x_t(1 + \mathcal{R}_{j,t-1}), \tag{D.1}$$

where $\mathcal{R}_{j,t} = R_{j,t}/V_{j,t}$ is the reserve ratio and $\mathcal{F}_{j,t} = (V_{j,t} - (1 + y_{j,t})V_{j,t-1})/V_{j,t-1}$ is net flow. The empirical estimate of the contract return policy in Table 3 implies:

$$E_{t-1}[y_{j,t}] = y_0 + \frac{\partial y}{\partial \mathcal{R}} \times [\mathcal{R}_{j,t-1} + x_{j,t}(1 + \mathcal{R}_{j,t-1})], \qquad \frac{\partial y}{\partial \mathcal{R}} \simeq 0.03. \tag{D.2}$$

The empirical estimate of the flow-reserves relation in Table 5 implies:

$$E_{t-1}[\mathcal{F}_{j,t}] = \mathcal{F}_0 + \frac{\partial \mathcal{F}}{\partial \mathcal{R}} \times \mathcal{R}_{j,t-1}, \qquad \frac{\partial \mathcal{F}}{\partial \mathcal{R}} \simeq 0. \tag{D.3}$$

Taking the conditional expectation $E_{t-1}[.]$ of (D.1) and using (D.2) and (D.3) to substitute $E_{t-1}[y_{j,t}]$ and $E_{t-1}[\mathcal{F}_{j,t}]$, respectively, we obtain:

$$\mathcal{R}_{j,t} = \frac{x_{j,t} + (1 + x_{j,t})\mathcal{R}_{j,t-1} - y_{j,t} - \pi x_t(1 + \mathcal{R}_{j,t-1})}{\mathcal{F}_0 + \frac{\partial \mathcal{F}}{\partial \mathcal{R}} \times \mathcal{R}_{j,t-1} + 1 + y_0 + \frac{\partial y}{\partial \mathcal{R}} \times [\mathcal{R}_{j,t-1} + x_{j,t}(1 + \mathcal{R}_{j,t-1})]}. \tag{D.4}$$

We linearize (D.4) for small deviations of $(x_{j,t}, \mathcal{R}_{j,t})$ around the steady state $(\overline{x}, \overline{\mathcal{R}})$, where $\overline{x} = E_{t-1}[x_{j,t}]$ and $\overline{\mathcal{R}}$ is such that $(x_{j,t}, \mathcal{R}_{j,t-1}) = (\overline{x}, \overline{\mathcal{R}})$ implies $E_{t-1}[\mathcal{R}_{j,t}] = \overline{\mathcal{R}}$. First, we solve for the steady state. The steady-state contract return $\overline{y} = y_0 + \frac{\partial y}{\partial \mathcal{R}}[\overline{\mathcal{R}} + \overline{x}(1 + \overline{\mathcal{R}})]$ must satisfy $\overline{y} = (1 - \pi)\overline{x} - \frac{\overline{\mathcal{F}}\overline{\mathcal{R}}}{1 + \overline{\mathcal{R}}}$, where $\overline{\mathcal{F}} = \mathcal{F}_0 + \frac{\partial \mathcal{F}}{\partial \mathcal{R}}\overline{\mathcal{R}}$ is steady-state net flow. Then, we linearize (D.4) around the steady state:

$$E_{t-1}\left[\mathcal{R}_{j,t} - \overline{\mathcal{R}}\right] = (1 - \delta)(\mathcal{R}_{j,t-1} - \overline{\mathcal{R}}) + \epsilon_{j,t}, \tag{D.5}$$

where

$$\delta = \frac{(1 + \overline{\mathcal{R}})(1 + \overline{x})\frac{\partial y}{\partial \mathcal{R}} + \overline{\mathcal{F}}/(1 + \overline{\mathcal{R}}) + \frac{\partial \mathcal{F}}{\partial \mathcal{R}}\overline{\mathcal{R}}}{1 + (1 - \pi)\overline{x} + \overline{\mathcal{F}}/(1 + \overline{\mathcal{R}})} \tag{D.6}$$

and

$$\epsilon_{j,t} = \frac{(1 + \overline{\mathcal{R}})(1 - \pi - (1 + \overline{\mathcal{R}})\frac{\partial y}{\partial \mathcal{R}})}{1 + (1 - \pi)\overline{x} + \overline{\mathcal{F}}/(1 + \overline{\mathcal{R}})}(x_{j,t} - \overline{x}). \tag{D.7}$$

A first order approximation of (D.6) for small values of $\frac{\partial y}{\partial \mathcal{R}}$, $\overline{\mathcal{F}}$, $\overline{x}$ and $\overline{\mathcal{R}}$ implies:

$$\delta \simeq \frac{\partial y}{\partial \mathcal{R}} + \overline{\mathcal{F}} + \frac{\partial \mathcal{F}}{\partial \mathcal{R}} \overline{\mathcal{R}}. \tag{D.8}$$

The first term of (D.8) arises because fraction $\frac{\partial y}{\partial \mathcal{R}}$ of reserves are distributed to investors. The second and third terms reflect reserves dilution by flows. The second term arises because unconditional flows dilute reserves at a rate equal to the unconditional net flow rate $\overline{\mathcal{F}}$. The third term arises because conditional flows dilute reserves at a rate equal to the sensitivity of flows to reserves $\frac{\partial \mathcal{F}}{\partial \mathcal{R}}$ times the unconditional reserve ratio $\overline{\mathcal{R}}$. Using $\frac{\partial y}{\partial \mathcal{R}} = 0.03$ (Table 3), $\overline{\mathcal{F}} = 0.024$ (Table 1), and $\frac{\partial \mathcal{F}}{\partial \mathcal{R}} = 0$ (Table 5), the reserve ratio mean reverts at a rate of $\delta \simeq 5.4\%$ per year.

E Taxes

E.1 Tax treatment of euro contracts

Contract returns are automatically reinvested in the contract and are not taxable until cash is withdrawn. When an individual withdraws cash, contract returns associated with the withdrawal are taxable as capital income.

The French tax system for capital income has a two-tier structure. The first tier is social security contributions, which is a flat tax on capital income whose rate has progressively increased from 10% in 1999 to 15.5% in 2015. The second tier is the income tax. Households can either include capital income in their taxable income, in which case it is taxed at the marginal income tax rate (between 0% and 45% depending on total taxable income and household size). Or they can choose to pay a flat withholding tax, whose rate depends on the savings vehicle. The withholding tax rate has been in the range 16%–19% for directly held stocks and mutual funds over 2004–2015.[35] For euro contracts and unit-linked contracts, the withholding tax rate depends on the holding period of the contract: 35% if less than four years; 15% between four and eight years; 7.5% with a tax allowance of 4,600 euros if more than eight years.[36] The withholding tax is the most favorable option for the majority of households (at least in value-weighted terms).

E.2 Tax cost of switching insurer

The tax system creates a tax cost of switching insurer for two reasons. First, contract returns are taxed upon withdrawals. Therefore, switching contracts moves the tax bill forward in time, which increases the present value of the tax bill. Second, the tax rate is a (non-continuously) decreasing function of contract age upon withdrawal. Therefore, switching contract increases the applicable tax rate by resetting the tax clock. The total tax loss of switching contract depends on how long the investor has held the initial contract and how long she will hold the new contract.

To quantify the tax cost of switching insurer, consider an investor who has been holding a contract for m years and has a contract value of one euro in year t, i.e., she invested $(1+y)^{-m}$ euro in year $t-m$. We calculate the year t-present value of the tax bill in the following two scenarios: (1) she holds the contract for another n years; (2) she switches to another insurer and holds the new contract for n years. Returns are taxed upon withdrawals and the tax rate depends on the age

[35] See Aubier, Cherbonnier, and Turquety (2005).

[36] Source: www.service-public.fr/particuliers/vosdroits/F22414

of the contract at the time of withdrawals. During the sample period, the tax rate for a k year old contract is $\tau(k) = 35\%$ if k is less than four years; $\tau(k) = 15\%$ if k is between four and eight years; and $\tau(k) = 7.5\%$ if k is more than eight years. In scenario (1), the tax bill is

$$\tau(m+n)\big[(1+y)^n - (1+y)^{-m}\big] \quad \text{in year } t+n. \tag{E.1}$$

In scenario (2), the tax bill is

$$\begin{cases} \tau(m)\big[1-(1+y)^{-m}\big] & \text{in year } t, \\ \tau(n)\big[(1+y)^n - 1\big] & \text{in year } t+n. \end{cases} \tag{E.2}$$

The tax cost of switching insurer is the year t-present value of (E.2) minus that of (E.1). This tax cost is plotted in Figure E.1 as a function of n, for $m \in \{0, 4, 8\}$.

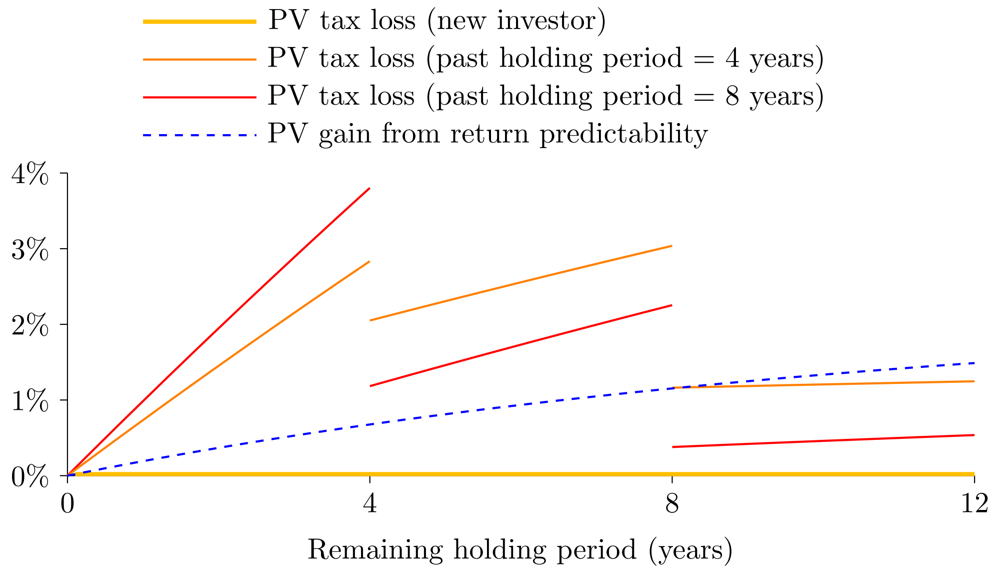
We compare the tax cost of switching insurer to the gain of switching from an insurer with a low reserve ratio $\mathcal{R}_L$ to an insurer with a higher reserve ratio $\mathcal{R}_H > \mathcal{R}_L$. The gain is calculated as present value of additional contract returns obtained by switching to the high reserve contract. We discount the expected return difference between the two contracts at the risk-free rate, because the market beta of the long high-reserves/short low-reserves portfolio is zero (see Table 7). Using that reserves are distributed to investors at a rate of $\frac{\partial y}{\partial \mathcal{R}} \simeq 3\%$ per year (see Columns 1–2 of Table 3) and that the reserve ratio decays at rate $\delta \simeq 5.4\%$ per year (see Appendix D), the present value for an investment of n years is

$$PV(n) = \frac{\partial y}{\partial \mathcal{R}} \times (\mathcal{R}_H - \mathcal{R}_L) \times \frac{1 - \big(1 - r_f - \delta\big)^n}{r_f + \delta}. \tag{E.3}$$

The present value is evaluated at the sample standard deviation of the reserve ratio $\mathcal{R}_H - \mathcal{R}_L = 0.068$ using $r_f = 3\%$. It is plotted in Figure E.1 as a function of the investment horizon n.

Two main results can be taken away from Figure E.1. First, new investors (yellow line) face no tax distortions and thus should always select contracts with higher reserves. Second, for investors already holding a contract for four yours (orange line) or eight years (red line), the tax loss outweighs the gains from predictability (dashed blue line) if investors plan to liquidate their investment within eight years whereas the gain outweighs the loss if investors plan to invest for another eight years or more. Given that the average holding period is twelve years, the majority of investors already holding a contract should find switching contract to be profitable. Thus, tax distortions do not seem

Figure E.1: Tax Cost of Switching Contract. The figure plots the tax losses of switching from an insurer with low reserves to an insurer with high reserves as a function of the remaining holding periods. The solid blue line is the present value of expected additional returns. The dashed red (orange) line is the present value of the tax loss for an investor who has held her previous contract for eight (four) years. The dashed green line is the present value of the tax loss for an investor who does not already hold a contract.



qualitatively large enough to explain inelastic flows even for investors already holding a contract.