ICIR Working Paper Series No. 22/2016

The Modern Tontine: An Innovative Instrument for Longevity Risk Management in an Aging Society

Jan-Hendrik Weinert, Helmut Gründl*

This version: July 22, 2016

Abstract

The changing social, financial and regulatory frameworks, such as an increasingly aging society, the current low interest rate environment, as well as the implementation of Solvency II, lead to the search for new product forms for private pension provision. In order to address the various issues, these product forms should reduce or avoid investment guarantees and risks stemming from longevity, still provide reliable insurance benefits and simultaneously take account of the increasing financial resources required for very high ages. In this context, we examine whether a historical concept of insurance, the tontine, entails enough innovative potential to extend and improve the prevailing privately funded pension solutions in a modern way. The tontine basically generates an age-increasing cash flow, which can help to match the increasing financing needs at old ages. However, the tontine generates volatile cash flows, so that - especially in the context of an aging society - the insurance character of the tontine cannot be guaranteed in every situation. We show that partial tontinization of retirement wealth can serve as a reliable supplement to existing pension products.

Keywords: Life Insurance, Tontines, Annuities, Asset Allocation, Retirement Welfare, Aging Society
JEL Classification: D14, D91, G11, G22, H75, J11, J14

1 Introduction

Through changes in social, financial and regulatory conditions, both life insurance policyholders and life insurers are facing big challenges. One of the large social challenges in most of the western countries is the demographic change caused by declining birth rates and an increasing longevity of the population[1]. Therefore, the retirement quotient rises; as a result pay-as-you-go (PAYG) retirement systems come under pressure while funded retirement products gain relevance. In addition, the liquidity-need increases for elderly people, which is mainly driven by increasing medical expenses at old ages. According to a study by Standard Life (2013), the liquidity-need of persons older than 85 years is six times higher than for persons below 65 years of age. As a consequence, the demand for funded retirement products that help to diminish the pension provision gap in an aging society increases.

Without doubt, life and health insurers already offer appropriate pension, health insurance and life-care products, sometimes combined to enhanced annuities. Yet insurers are exposed to changing financial and regulatory conditions. Traditional pension products often entail minimum return guarantees, which providers try to ensure by investing extensively in fixed-income securities. However, the current low interest environment clearly shows large solvency risks caused by the issuance of lifetime guarantees. The possible way out of this problem, i.e., to invest extensively in more profitable asset classes like stocks, is however restricted due to its higher risk and its limited ability to cover granted guarantees. Furthermore, providers of pension products are exposed to the longevity risk of their customers, which can only be partially passed on to them.

In the European Union and many other parts of the world, the regulatory conditions for providers of private pension products change substantially with the introduction of risk-based solvency regulation. The market-consistent valuation of investments as well as of technical provisions immediately reveals the addressed high risks of traditional life and health insurance products involved and therefore can cause severe financial imbalance for life insurers. In a traditional insurance context, managing these risks requires considerable equity capital backing or other comprehensive risk management activities like re-insurance or securitization. Such risk management ultimately has to be funded by higher insurance premiums, which might make pension planning unattractive.

The changes in social, financial and regulatory conditions therefore lead to the quest for innova-

---

[1] See for example Statistisches Bundesamt (2015) for a prognosis for the demographic change until the year 2060 in Germany.

tive instruments for private pension planning. A product innovation should optimally reduce or dispense with the need for investment guarantees and risks related to longevity, and nevertheless be able to provide reliable insurance performance. At the same time it should meet the concerns of increasing liquidity-needs at old ages.

Against this backdrop, we transfer the idea of the historic tontine to a modern context and analyze whether it can help to solve the aforementioned problems. A tontine provides a mortality driven, age-increasing payout structure. Although an insurer can easily replicate such a payout structure, the tontine has the big advantage of its simplicity and (presumably) low costs. While traditional insurance products entail large safety and administrative cost loadings[2], a tontine can be offered at low additional costs. This is because a tontine is a simple redistribution mechanism of the invested funds without the need for an active management. The investment strategy of the tontinized wealth can be decided on an individual basis according to the individual risk aversion, without the issuance of guarantees. Due to the linkage between tontine returns and individual mortality, tontine payments are very low in younger years and increase strongly for very high ages. As we observe an increasing liquidity-need at old ages, we can show that the tontine can be an appropriate instrument to serve as financial protection in the late retirement years for relatively small investment volumes at low cost.

We show that a modern tontine can be an appropriate complement to existing privately funded pension solutions with the ability to improve policyholders' welfare. To this end, we take the perspective of a tontine holder who holds a tontine for pension planning purposes. We compare the tontine and its benefit structure with that of a conventional pension annuity and derive implications for the individual tontine demand. We show that a tontine can be a cost-efficient instrument to serve the increasing monetary demand at old ages and can thereby become an interesting supplement to traditional pension planning solutions. Although tontines do not provide for any investment guarantees, their features make it possible to generate cash flows that cover the age-specific needs of retirees. In an aging society, tontines can thus become an interesting instrument of old-age provision, complementary to conventional insurance products. For assessing the advantages and disadvantages of tontines, their age-increasing payments with the ability to finance the also increasing care costs have so far not been investigated. This is the gap in the literature which we try to bridge with our contribution.

In our article, we aim at assessing the effects of tontinizing some fraction of the individual

---

[2] According to Bundesanstalt für Finanzdienstleistungsaufsicht (2014) the average acquisition and administrative costs for German life insurers are 10.7% of the gross premiums and amount to more than 20% for Non-Direct Insurers.

retirement wealth on the individual lifetime utility, considering an increasing liquidity-need at old ages. We illustrate under which conditions there are appropriate incentives for individuals to hold some fraction of retirement wealth in tontines compared to complete annuitization.

The remainder of the article is organized as follows: Section 2 introduces the general concept of tontines. Section 3 reviews the relevant literature on tontines and increasing liquidity-needs at old ages. Section 4 introduces our model framework specifying the underlying mortality dynamics, the tontine model as well as the old-age liquidity-need curve and the valuation of annuities. Finally, we propose a Cumulative Prospect Theory based valuation of lifetime utility of tontines and annuities. In Section 5, we first describe the data and the calibration we adopt and provide findings for the optimal individual wealth allocation and discuss our results. Section 6 provides implications and our conclusion.

2   Tontines

The Italian Lorenzo de Tonti invented a product to consolidate the public-sector deficit in the 1650s.[3] His ideas were based on the pooling of persons by considering their mortality risk. The innovation was that, in exchange for a lump sum payment to the Italian government, one received the right to a yearly, lifelong pension. This pension increased over time because the yields were distributed among a smaller amount of surviving beneficiaries. The last survivor thus received the pensions of all others who died before. As Manes (1932) notes, the valuation of the original tontine was inaccurate, retirees were grouped in broad age classes, and so the contract terms were not fair in an actuarial sense. In this article, we build upon a fair tontine based on Sabin (2010) that allows participants to be of any age, of any gender, and to invest a desired amount of money[4] in the tontine. Furthermore, the tontine is revolving, which means that new members can join the tontine at any age and take on the position of deceased members. Apart from that it is not allowed to leave the tontine before passing away. The tontine is a fair lottery for every member. Expected individual tontine payments equal the individual investment in the tontine, yielding an unconditional expected profit of zero. Expected tontine payments depend on the individual stake in the tontine and on the individual survival probability. On the one hand, if a member dies, he or she loses the entire stake, while, on the other, if he or she survives he or she receives some fraction of the stake of deceased members. To be fair, both expected gains and losses are equal in each period. Because the survival probability declines

---

[3] See McKeever (2009) for an overview of the history of tontines.

[4] According to Sabin (2010) a fair tontine is a tontine in which the distribution to surviving participants is made in unequal portions according to a plan that provides each participant with a fair bet.

with age and tontine payments are only paid if one survives, the probability of receiving tontine payments decreases. To counterbalance the otherwise induced reduction in expected tontine payments, the size of the payments one receives has to increase. Through this mortality-driven feature, the expected conditional tontine payments increase with age. Mortality, therefore, is the crucial factor for determining the tontine benefit structure. For example, a man born in 1981 can expect to live 84 years, while a man born in 2013 has an increased life expectancy of 89 years. This difference of 5 years translates directly into different benefit patterns, especially at old ages. Furthermore, the whole composition of the demographic structure of the tontine members changes on the basis of the population mortality, which also impacts the actual tontine benefit realization structure. Therefore, it is important to model and forecast the development of mortality and demographic structure of the tontine members. We use the one-factor model by Lee and Carter (1992) to forecast mortality, which is the standard approach to model mortality rates. Recent studies which use the Lee-Carter model are, for example, Renshaw and Haberman (2003) and Renshaw and Haberman (2008). The Lee-Carter model only considers cohort effects, i.e. changes of mortality within a cohort. Willets (2004) empirically substantiates cohort effects as characteristic of mortality dynamics.

While in a traditional annuity[5], longevity risk is transferred from the insured to the insurer (and covered by its risk management instruments), in a tontine the risk that a single participant might live longer than expected is completely borne and shared by the other tontine holders who in this case receive less cash flows than expected. Therefore, no equity capital backing is needed to cover longevity risk, and the tontine can be offered without a risk-cost loading. However, the tontine has the disadvantage that because individual shares in the tontine as well as times of death of tontine members are random, both amount and timing of tontine payments are uncertain.

Because the tontine members carry, pool and share the total risk among each other while the offering provider does not bear it, a tontine can be offered at a cheaper price than a comparable traditional life insurance product. Additionally, it generates age-increasing benefits and is therefore able to meet the increasing monetary requirements at old ages. Moreover, due to the absence of guarantees, the tontine enables participation in stock market developments. However, the tontine generates volatile payments, which means that the insurance character of a tontine might not be ensured in every situation.

So far, tontines have been considered an alternative or historic predecessor of traditional pension

---

[5] In the following we use annuity synonymously for the traditional life insurance product.

insurance. However, tontines were considered to be inferior compared to traditional pension insurance, because the latter provides less volatile payments for an equal expected return[6]. In addition, one could think of potential moral hazard problems among tontine members. However, Milevsky (2015) found no evidence that such phenomena have ever occurred in practice.

3   Literature Review

In the existing literature there are only few contributions on the evaluation of longevity-linked securities from a policyholder perspective. Sabin (2010) designs a fairly priced tontine with regard to age, gender and entry date with an equivalent to a common annuity scheme. His results exhibit a better payout pattern than a typical insurer-provided annuity not just on average, but for virtually every member who lived more than just a few years. Cairns et al. (2008) summarize various instruments that deal with longevity risks of insurance companies. Their review presents various concepts of longevity-linked securities, such as mortality swaps and longevity bonds that may serve as a hedging instrument against longevity. However, the products introduced had only moderate success due to low market acceptance and the lack of a perfect hedge. Milevsky and Salisbury (2015) derive an optimal tontine design by accounting for sensitivity of both the tontine size and the longevity risk aversion for each tontine member. By doing so, they raise the question whether an optimally designed tontine with low implications regarding capital requirements for the sponsor will gain more attention in times of risk-based capital standards and conclude that, due to higher volatility of the payments, the tontine provides a lower utility than a traditional life annuity. Forman and Sabin (2014) construct a fair transfer plan (FTP) to guarantee a fair bet for all participating investors of a tontine by accounting for each age, death expectancy and investment level. They show that a fairly designed tontine is superior to defined benefit plans in terms of funding and sponsoring of the pension system. They illustrate that a fairly developed tontine model would improve the situation of pension providers while serving the retirement income demand of the tontine participants. In an empirical work, Blanchet (2013) derives the trend of retirement consumption. He observes a shift in the expenditures towards increasing health care, entertainment and food. This work is the basis for the old-age liquidity-need function that can be served by the tontine.

---

[6] See Sabin (2010).

4 Model Framework

We first model mortality dynamics in Germany for the upcoming decades and derive in a second step possible demographic pyramids. These, in turn, are the basis for the composition of the fair revolving tontine. We then estimate the risk-free benefit profile of a standard annuity and compare it to the risky benefit profile of a tontine. We assume that each individual i has wealth W_i consisting of the total available funds and that there is no further source of income in the future. W_i can be seen as the sum of both discounted future earnings until retirement and savings up to the investment date. We assume a situation in which W_i will be completely converted into pension installments[7]. For the expected tontine benefit, we provide a closed-form solution while we determine the realized benefit by performing a Monte Carlo simulation. We then analyze to what extent tontine and traditional annuity are able to satisfy an empirically estimated, increasing old-age liquidity-need function for different settings. Furthermore, we estimate an optimal portfolio consisting of annuity and tontine which maximizes expected utility. We provide results for different demographic scenarios and mortality dynamics for Cumulative Prospect Theory and Expected Utility Theory and show the capability of tontines as instruments for retirement planning from a policyholder perspective. Finally we incorporate subjective beliefs about individual mortality to account for different perceptions about individual life expectations, which leads to a changing optimal asset allocation for retirement planning.

4.1 Mortality Model

In a first step, we project mortality rates for a forecast horizon of $t = 1, \ldots, T$ years. Our starting point is the one-factor-model for estimating mortality rates by Lee and Carter (1992). According to the Lee-Carter Model (LC model) the one-year death probability $q_{x,t}$ of a person aged x in year t is specified as

$$\ln(q_{x,t}) = a_x + b_x \cdot k_t + \varepsilon_{x,t} \Leftrightarrow q_{x,t} = e^{a_x + b_x \cdot k_t + \varepsilon_{x,t}} \tag{1}$$

where a_x and b_x are time constant parameters for a male aged x that determine the shape and the sensitivity of the mortality rate to changes in k_t, which is a time-varying parameter that captures the changes in the mortality rates over time. $\varepsilon_{x,t}$ is an error term with mean 0 and constant variance[8]. As originally proposed by Lee and Carter (1992), the estimation of the time

---

[7] This assumption can be substantiated by Yaari (1965), who shows that under the expected utility paradigm for fairly priced annuities without bequest motive, complete annuitization is optimal

[8] We use y for a female aged y years analogously.

varying parameter k_t can be performed by fitting a standard $ARIMA$ model using standard time series analysis techniques. The $ARIMA(p, d, q)$ process is given by

$$k_t = (\alpha_0 + \alpha_1 k_{t-1} + \alpha_2 k_{t-2} + ... + \alpha_p k_{t-p} + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + ... + \beta_q \varepsilon_{t-q}) + \varepsilon_t = \widehat{k_t} + \varepsilon_t \quad (2)$$

in which the error term is normally distributed $\varepsilon \sim N(0, \sigma_k)$.[9]

4.2   Demographic Structure

Based on the predicted mortality rates $q_{y,t}$ for the one-year death probability of a woman aged y in period t, and $q_{x,t}$ for the one-year death probability of a man aged x in period t, we determine the demographic structure of an economy in every period t. $Q_{y,t}$ ($Q_{x,t}$) is the total quantity of female (male) persons of a cohort aged y (x) at time t. Equation (3) shows the updating process. Newborns, or persons in their first year of life ($y, x = 1$) are determined by the sum of the age-specific fertility rate AGZ_y times the quantity of females of the respective age y in each period t, which is weighted by the fraction of newborn females f_0 and males $m_0 = 1 - f_0$. From the second year of being alive, the number of persons is the probability to survive one year of someone who was one year younger in the year before times the number of persons who were one year younger in the year before:

$$Q_{y,t} = \begin{cases} f_0 \cdot \sum_{y=2}^{\Omega} Q_{y,t} \cdot AGZ_y \\ Q_{y-1,t-1} \cdot (1 - q_{y-1,t-1}) \end{cases} \qquad Q_{x,t} = \begin{cases} m_0 \cdot \sum_{y=2}^{\Omega} Q_{y,t} \cdot AGZ_y & \text{for } y, x = 1 \\ Q_{x-1,t-1} \cdot (1 - q_{x-1,t-1}) & \text{for } y, x = 2 \ldots \Omega \end{cases} \quad (3)$$

We determine these quantities for all cohorts $y, x = 1, \ldots, \Omega$ in all periods $t = 1, \ldots, T$ for males and females to estimate the corresponding population pyramids. Equation (4) shows the Cumulative Distribution Function (CDF) of a person aged x in each t, where $\kappa_{y,t} = \frac{Q_{y,t}}{\sum_{y=1}^{\Omega} Q_{y,t}}$ and $\kappa_{x,t} = \frac{Q_{x,t}}{\sum_{x=1}^{\Omega} Q_{x,t}}$ for $y, x = 1, \ldots, \Omega$ are the fractions of each cohort of females and males of the total female and male population in each period t.

$$F_{y,t} = \begin{cases} 0 & : y < 0 \\ \sum_1^y \kappa_{y,t} & : 0 \le y < \Omega \\ 1 & : y > \Omega \end{cases} \qquad F_{x,t} = \begin{cases} 0 & : x < 0 \\ \sum_1^x \kappa_{x,t} & : 0 \le x < \Omega \\ 1 & : x > \Omega \end{cases} \quad (4)$$

---

[9] The term d indicates the grade of co-integration of the series. For further mathematical details, please refer to Brockwell and Davis (2013) pp. 273-320.

The fractions of females and males aged y and x of the total population in each t are $f_{y,t} = \frac{\kappa_{y,t}}{\kappa_{y,t} + \kappa_{x,t}}$ and $m_{x,t} = 1 - f_{y,t}$. The predicted mortality rates and demographic structures are the basis for calculating the tontine composition and tontine benefits and are the starting point to analyze the impact of the demographic development on the tontine and its possible application as an alternative retirement planning product.

4.3 Tontine Model

We model a fair revolving tontine based on Sabin (2010) that allows tontine participants to be of any gender and age, and to invest a desired one-time initial amount of money B_i at tontine entrance. B_i then is tied in the tontine and cannot be withdrawn before the tontine member's death. Furthermore we assume that there is no possibility to inject additional capital for any individual in future periods. While the original model considers infinitesimal points in time, resulting in only one member being able to die at one point in time, we adjust the model to a yearly time frame allowing for multiple deaths. We further assume the number of the tontine members N as fixed: every time a participant dies, the tontine is refilled to N. A new entrant i is randomly drawn from the period-corresponding demographic structure. We assume entrants at least to be of a certain age $\underline{y}$, $\underline{x}$ and also assume an upper limit of entering the tontine of age $\overline{y}$, $\overline{x}$ so $0 \leq \underline{y}, \underline{x} < \overline{y}, \overline{x} \leq \Omega$. The random age of an individual i entering the tontine in t is expressed by

$$y_{E,i,t}, \quad x_{E,i,t} = F_t^{-1}(z) \tag{5}$$

where F^{-1} is the inverse function of F with $Pr(y) = f_{y,t}$ and $Pr(x) = m_{x,t}$ and with $z \in \left(F_t^{-1}(\underline{y}), F_t^{-1}(\overline{y})\right)$ or $z \in \left(F_t^{-1}(\underline{x}), F_t^{-1}(\overline{x})\right)$ where z is uniformly distributed on $\mathcal{U}\left(F_t^{-1}(\underline{y}), F_t^{-1}(\overline{y})\right)$ or $\mathcal{U}\left(F_t^{-1}(\underline{x}), F_t^{-1}(\overline{x})\right)$. We further assume the establishment of the tontine in $t = 0$ and refrain from investing the tied capital to streamline the model and to be able to quantify solely the interrelation of mortality benefits and demographic change. For better readability, we denote the one-year death probability of individual i with the beforehand assigned characteristics as $q_{i,t}$ in t. The index i allows to identify each individual with its specific characteristics in each period. Furthermore we denote the age of a person as x in the following, irrespective of the gender.

Let $\{A_{i,t}\}$ be the event that i dies in t with $P(A_{i,t}) = q_{i,t}$ and $\left\{A_{i,t}^{\complement}\right\}$ be the event that i survives in t with $P\left(A_{i,t}^{\complement}\right) = 1 - q_{i,t}$. Let $\left\{A_{0,t}^{\complement}\right\}$ be the event that at least someone dies in t and $\{A_{0,t}\}$ be the event that no one dies in t. Using the inclusion-exclusion principle[10], the probability that

---

[10] see for example Graham et al. (1997).

at least someone dies in t is

$$P\left(A_{0,t}^{c}\right) = P\left(\bigcup_{i=1}^{N} A_{i,t}\right) = \sum_{j=1}^{N}\left((-1)^{j+1} \sum_{\substack{I \subseteq \{1,\dots N\} \\ |I|=j}} P\left(\bigcap_{i \in I} A_{i,t}\right)\right)$$

and the probability that no one dies in t is

$$P\left(A_{0,t}\right) = 1 - P\left(\bigcup_{i=1}^{N} A_{i,t}\right).$$

$\left\{A_{k,t} \mid A_{0,t}^{c}\right\}$ denotes the event that k dies in t conditioned that at least someone dies in t. Using the law of total probability yields for the probability that k dies in t conditioned that at least someone dies in t

$$P\left(A_{k,t} \mid A_{0,t}^{c}\right) = \frac{P\left(A_{k,t}\right)}{P\left(A_{0,t}^{c}\right)} = \frac{q_{k,t}}{\sum_{j=1}^{N}\left((-1)^{j+1} \sum_{\substack{I \subseteq \{1,\dots N\} \\ |I|=j}} P\left(\bigcap_{i \in I} A_{i,t}\right)\right)} = \rho_{k,t}. \qquad (6)$$

If member i dies, his or her balance account B_i is distributed to the survivors. To be fair, this reallocation takes place according to the specific characteristics of the surviving members: Older members and those with a larger stake in the tontine receive more. If member $k \neq i$ dies, member i receives a fraction of k's balance $a_{i,k,t}B_k$, where

$$0 \leq a_{i,k,t} \leq 1 \text{ for } i, k = 1, \dots, N \text{ and } i \neq k. \qquad (7)$$

k's balance is forfeited entirely, so

$$a_{k,k,t} = -1 \text{ for } k = 1, \dots, N. \qquad (8)$$

Equation (8) states that the dying members' stake in the tontine is distributed among the surviving members. In sum, the amount lost by k equals the sum of the distributed benefits to the surviving members, so

$$\sum_{i=1}^{N} a_{i,k,t} = 0 \text{ for } k = 1, \dots, N. \qquad (9)$$

The unconditional expected benefit received by member i in t is the return in case no one dies and the return if at least someone dies, weighted with their corresponding probabilities, thus

$$E\left[r_{i,t}\right] = E\left[r_{i,t} \mid A_{0,t}\right] P\left(A_{0,t}\right) + E\left[r_{i,t} \mid A_{0,t}^{c}\right] P\left(A_{0,t}^{c}\right) \qquad (10)$$

Since return is generated solely by mortality, if no one dies, there cannot be any return, thus $E[r_{i,t} \mid A_{0,t}] = 0$ and the expected return comprises the second term of the right-hand side of Equation (10). The expected return conditioned that at least someone dies is the sum of the conditional dying probability weighted fractions of the balance accounts over all k members in t, thus

$$E\left[r_{i,t} \mid A_{0,t}^{\mathsf{c}}\right] = \sum_{k=1}^{N} \rho_{k,t} a_{i,k,t} B_k. \tag{11}$$

To achieve a fair tontine, each member's expected benefit is zero in each year. This is because the expected loss of the own balance account in the case of the own death has to be offset by the expected gains one receives from other members' deaths, so $E\left[r_{i,t} \mid A_{0,t}^{\mathsf{c}}\right] \overset{!}{=} 0$. The older i is, the higher the death probability $q_{i,t}$, causing that $\rho_{i,t}$ increases as well and leads to a higher expected loss in case of the i-th death. This has to be compensated by an increase in the fractions $a_{i,k,t}$ one receives in the case of other members' death to counterbalance the aforementioned effect and to create a fair bet.

To satisfy the conditions of a fair bet for every tontine member $i = 1, \ldots, N$, requires to search for a set of $a_{i,k,t}$s that yield an expected benefit of zero for every tontine member and which fulfills conditions (6), (7), (8) and (9), yielding $E[r_{i,t}] = E[r_{i,t} | A_{0,t}^{\mathsf{c}}] = 0$.

As Sabin (2010) shows, such a set of $a_{i,k,t}$s exists only if no member is exposed to more than half of the total risk of the tontine and can be achieved by the introduction of a ceiling of the amounts to invest B_i. Choosing N large enough additionally reduces the threat of a single individual holding too large a fraction of risky exposure of the tontine. Here, we implement an algorithm[11] for the determination of the set of $a_{i,k,t}$s that is proposed by Sabin (2010), which approximately assigns constant $a_{i,k,t}$s, irrespective of k for $k \neq i$ and which provides best results for large N. In the following, we assume that the resulting $a_{i,k,t}$s satisfy conditions (6) - (9) and that no member holds more than half of risky exposure on the tontine, formally meaning that

$$\rho_{i,t} B_i \leq \frac{1}{2} \sum_{k=1}^{N} \rho_{k,t} B_k \text{ for } i = 1 \ldots N. \tag{12}$$

The expected return, conditioned that i survives in t, is

$$E\left[r_{i,t} | A_{i,t}^{\mathsf{c}}\right] = \sum_{\substack{k=1 \\ k \neq i}}^{N} \rho_{k,t} a_{i,k,t} B_k. \tag{13}$$

---

[11] For further algorithms to construct a tontine, see Sabin (2011).

Because Equation (11) is solved to be zero, to yield a fair bet, $\sum_{\substack{k=1 \\ k \neq i}}^{N} \rho_{k,t} a_{i,k,t} B_k = -\rho_{i,t} a_{i,i} B_i$ and

equation (13) is

$$E\left[r_{i,t} | A_{i,t}^{\mathbf{c}}\right] = \mu_{i,t} = q_{i,t} B_i. \tag{14}$$

This is an interesting property since the individual expected return in case of the own survival

is solely driven by the own mortality $q_{i,t}$ and the own investment in the tontine B_i, and does

not depend on the tontine composition.

The unconditional realized benefit for i in t is

$$r_{i,t} = \sum_{\substack{k=1 \\ k \neq i}}^{N} a_{i,k,t} B_k \mathbf{1}_{\left\{ A_{k,t} \cap A_{i,t}^{\mathbf{c}} \right\}} - B_i \mathbf{1}_{\left\{ A_{i,t} \right\}}$$

where the indicator function $\mathbf{1}_{\{\dots\}}$ takes on the value of 1 if the respective event occurs and 0

otherwise and therefore $\mathbf{1}_{\{\dots\}} \sim Ber_{P(\dots)}$. The realized return conditioned that i survives is

$$r_{i,t} | A_{i,t}^{\mathbf{c}} = \sum_{\substack{k=1 \\ k \neq i}}^{N} a_{i,k,t} B_k \mathbf{1}_{\left\{ A_{k,t} \right\}}. \tag{15}$$

Since the $a_{i,k,t}$s are approximately constant for i for every k for large N, $\overline{a}_{i,k,t} \approx \frac{q_{i,t} B_i}{\sum_{\substack{k=1 \\ k \neq i}}^{N} q_{k,t} B_k}$, and

equation (15) is

$$r_{i,t} | A_{i,t}^{\mathbf{c}} \approx q_{i,t} B_i \frac{\sum_{\substack{k=1 \\ k \neq i}}^{N} B_k \mathbf{1}_{\left\{ A_{k,t} \right\}}}{\sum_{\substack{k=1 \\ k \neq i}}^{N} B_k q_{k,t}}. \tag{16}$$

Although the volatility of the tontine converges toward zero for $N \to \infty$, for a finite and

realistic tontine size, the payouts are volatile. As Appendix A.1 shows, the tontine payouts are

approximately normally distributed, with

$$\mu_{i,t} = q_{i,t} B_i$$

and

$$\sigma_{i,t} = \frac{q_{i,t} B i}{\sqrt{M-1}} \sqrt{\sum_{m=1}^{M} \left(\frac{\sum_{\substack{k=1 \\ k \neq i}}^{N} B_k \mathbf{1}_{\left\{ A_{k,t} \right\}}^{m}}{\sum_{\substack{k=1 \\ k \neq i}}^{N} B_k q_{k,t}} - 1 \right)^2}. \tag{17}$$

with M Monte Carlo simulation paths.

4.4 Annuity Model

We assume that the individual has pension wealth W_i, no other wealth, no other source of income and the share of wealth which is not tontinized is annuitized. The individual can choose to hold any positive proportion of his wealth in an annuity or in a tontine. The individual has no heirs and no desire to leave a bequest. We refer to an annuity applied by Milevsky (2006) which requires a lump sum investment of $W_i - B_i$. The annuity then pays a stable income stream on a yearly basis until the participants' death, starting as an immediate annuity.

The conditional survival probability of a person aged x in t of surviving τ more years is defined as

$$\tau p_x^t = \prod_{j=0}^{\tau-1} \left(1 - q_{x+j}^{t+j}\right).$$

Because we refrain from interest rates in this model, the lump sum price $\bar{a}_x^t$ in t of an immediate annuity which provides an income of 1 EUR per year until death is equal to the expected remaining lifetime of a person aged x in t, $E\left[T_x^t\right]$ which is the sum of the conditional survival probabilities:

$$\bar{a}_x^t = E\left[T_x^t\right] = \sum_{\tau=1}^{\Omega} \tau p_x^{t+\tau}.$$

To simplify, we denote the price of the immediate annuity as $\bar{a}_i^t$, where the individual characteristics can be identified via i. A lump sum investment of $W_i - B_i$ in the annuity provides a stable, lifelong and yearly income stream of

$$R_{i,t} = \frac{W_i - B_i}{\bar{a}_i^t}. \tag{18}$$

The individual now can decide on the allocation of tontinization (B_i) and annuitization $(W_i - B_i)$ of the wealth W_i.

4.5 Old Age Liquidity Need Function

According to the Worldbank (2015), life expectancy at birth has increased between 1970 and 2013 from 70.6 to above 80 years. The increasing lifetime will cause the number of people over 80 years old to almost double to 9 million in Germany by the year 2060 according to forecasts by the Statistisches Bundesamt (2015). In the future, it is therefore very probable that very high ages of 100 years and even more will be achieved by a large number of people. According

to the medicalisation thesis motivated by Gruenberg (1977), the additional years that people live due to demographic change are increasingly spent in bad health condition and disability. In those additional years of life, the demand for care products and medical service increases over-proportionately. Coming from 2.6 million nursing cases in Germany in 2013, Kochskämper (2015) estimates between 1.5 and 1.9 million additional nursing cases in Germany in the year 2060 due to demographic change. By the year 2030, the demand for stationary permanent care will increase by 220,000 places in Germany.

While previous research finds a systematic decrease in the consumption level at retirement[12], incorporating the nursing care costs and medical expenses to consumption yields a so called retirement smile. When people retire they are mostly still healthy and have therefore time to spend on lifestyle. As they age, first physical constraints appear, they become more and more home-bound and thus consumption declines while supplementary and medical costs are still at a low level. As people become very old they rely more on assisted living requirements, and costly long-term nursing care is needed. Therefore the typical monetary demand for a retiree is U-shaped. Based on own empirical research[13] we model an old-age liquidity-need function, which accounts for demand for nursing care and medical service. The determination of the liquidity-need function is based on data available on consumer spending in the SOEP from 1984 to 2013. We determine the age-specific expenditure pattern for the spending categories food, living, health, care, leisure, refurbishment and miscellaneous and finally aggregate them. The considered age ranges from 60 to 95. From the age of 95 on, we extrapolate until the age of 105 due to the limited data basis for those ages. The modeling of the nursing care costs is based on the costs of an inpatient, permanent care, which occurs in nursing homes, even though a large share of care is performed by family members and nursing care services. The inpatient care costs reflect the actual potential resources needed in a more appropriate manner because in home care, the time spent by family members is not taken into account. Moreover, many refuse inpatient care only because of lacking resources. Figure 1a shows our estimates of the average nursing care costs from the age of 60 to 90 per year. In addition, the liquidity-need without nursing care costs is shown. If we aggregate both components, we obtain the characterized retirement smile, which is presented in figure 1b.

In our model, we map the old-age liquidity-need via a polynomial liquidity-need function D_t with order 2 which is calibrated based on our results, and assume an extrapolation up to the

---

[12] See for example Hamermesh (1984), Mariger (1987), Banks et al. (1998), Bernheim et al. (2001) or Haider and Stephens Jr (2007).

[13] The data used in this publication was made available to us by the German Socio-Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW), Berlin.
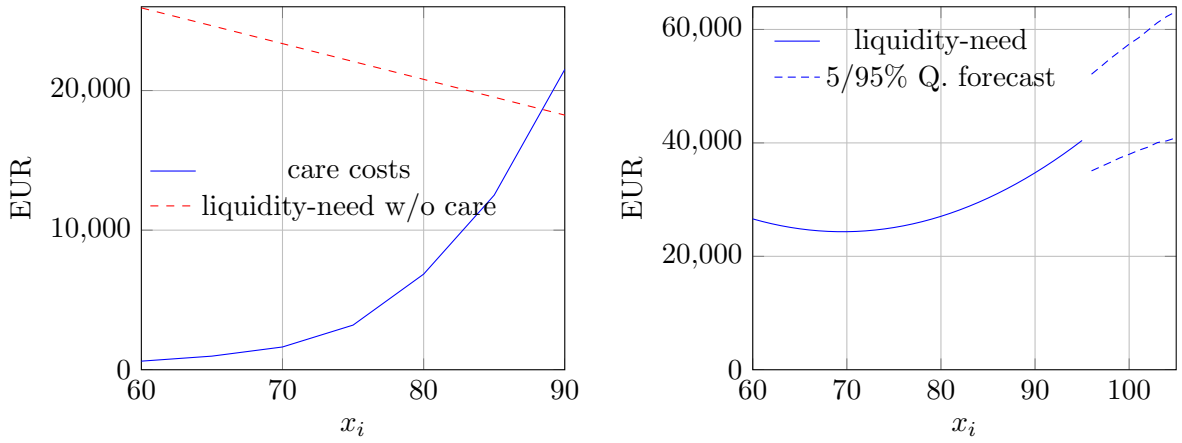
(a) Average nursing care costs and liquidity-need without nursing care per year

(b) Aggregated liquidity-need per year

Figure 1: The Retirement Smile

maximum attainable age of $x = 105$. The desired consumption level is driven by age x in t so

$$D_t = \beta_0 + \beta_1 x_t + + \beta_2 x_t^2 \epsilon_t \tag{19}$$

where the parameters β_0, β_1 and β_2 are fitted using our empirical data and ϵ_t is the error term in t.

4.6 Multi Cumulative Prospect Theory Valuation

So far, we have on the one hand an income stream which is composed of both certain and volatile annuity payments, and on the other an age-increasing liquidity-need which the income stream should cover. We aim to design the payout pattern of tontine and annuity such that the liquidity-need can be served in an optimal way. Therefore, we evaluate the income stream relative to the liquidity-need. An income stream larger than the liquidity-need is considered to be utility-generating, while an income stream lower than the liquidity-need generates disutility. In this sense, we look at the utility of the relative income stream in reference to the liquidity-need, rather than the absolute level of the income. Since the liquidity-need increases with age, a payout which is able to meet the demand in early years might not be sufficient in the later years of retirement. Therefore, to evaluate an income stream relative to a reference point, or in other words, gains and losses, the Cumulative Prospect Theory (CPT), originated by Kahneman and Tversky (1979) and enhanced by Tversky and Kahneman (1992) is highly suitable for our purpose: If the income stream is not sufficient to meet the liquidity-need, the missing difference to the reference point is valued rather than the absolute level of the income stream. Although the CPT is a descriptive rather than a normative theory, it allows us to

14

capture the aforementioned properties and to determine an optimal, CPT-utility maximizing fraction to be invested in the tontine. Schmidt (2015) shows that using CPT to value insurance demand can yield an optimum of full insurance coverage. To capture the life-cycle dynamics of the repeating payments until death, we use the Multi Cumulative Prospect Theory (MCPT) applied by Ruß and Schelling (2015) where the CPT utility is determined in every period t under consideration of a changing reference point which is represented by the respective liquidity-need D_t and finally aggregated with respect to survival prospects. The total utility of person i over his or her stochastic remaining lifespan is the weighted sum of the CPT utilities in each point in time t deflated by a subjective discount factor $\delta \leq 1$. The conditional survival probability $_\tau p_x$ of an x year old of surviving τ more years is incorporated in the CPT utility. It is combined with the occurrence probability of the respective payouts and this joint probability is valued according to the CPT, thus

$$MCPT\,(i) = \sum_{\tau}^{T-x} \delta^\tau CPT\,(Z_{i,t+\tau}) \tag{20}$$

where

$$CPT\,(Z_{i,t+\tau}) = \int_{-\infty}^{0^-} v\,(z)\,d\left(w^-\left(F_{i,t+\tau}\,(z)\right)\right) + \int_{0^+}^{\infty} v\,(z)\,d\left(-w^+\left(1 - F_{i,t+\tau}\,(z)\right)\right) \tag{21}$$

is the CPT in an adjusted version to allow for continuous probability distributions[14]. The probability weighting function $w^+\,(F)$ for gains and $w^-\,(F)$ for losses is

$$w^+\,(F) = \frac{F^\gamma}{\left(F^\gamma + (1-F)^\gamma\right)^{1/\gamma}}, \quad w^-\,(F) = \frac{F^\kappa}{\left(F^\kappa + (1-F)^\kappa\right)^{1/\kappa}} \tag{22}$$

The value function $v\,(z)$ is given by

$$v\,(z) = \begin{cases} z^a & z \geq 0 \\ -\lambda\,|z|^b & z < 0 \end{cases} \tag{23}$$

where $a, b \in (0,1)$ and $\lambda > 1$. The mixture cumulative distribution function, to account for the joint probability of conditional survival and the payout size, is

$$F_{i,t+\tau}\,(z) = (1 -_\tau p_x)\,\mathbf{1}_{[0,\infty)} +_\tau p_x \int_{-\infty}^{z} dF_{\mathcal{N}_{i,t+\tau}}\,(u) \tag{24}$$

---

[14] See for example Hens and Rieger (2010), Ågren (2006) or Ruß and Schelling (2015) who use the CPT in a continuous context.
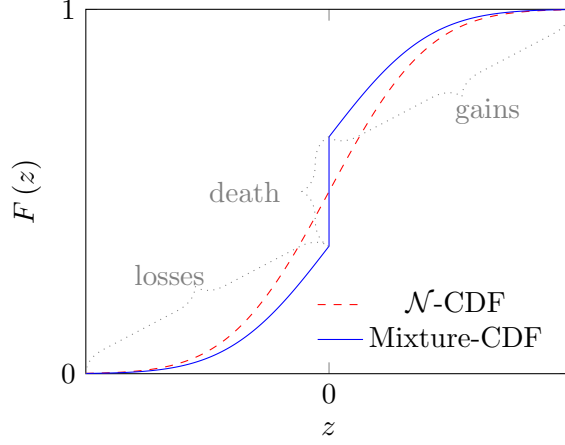
Figure 2: Mixture CDF of survival and payout size vs. Normal CDF

with the first moment

$$\mu_{i,t+\tau} = q_{i,t+\tau} B_i + \frac{W_i - B_i}{a_{i,0}} - D_{i,t+\tau} \tag{25}$$

and the second moment $\sigma_{i,t+\tau}$ resulting from equation (17) for the normally distributed gains and losses.

Figure 2 shows the intuition of equation (24) for a mean zero payout: if one is alive, one receives payments which occur with the conditional survival probability $_\tau p_x$. Up to the point where payments equal the liquidity-need, they are perceived as losses. If one dies, one receives no payments out of the tontine and annuity. Therefore, at $z = 0$, there is an immediate jump in the CDF of $1 - _\tau p_x$, because the probability of receiving less than 0 results not only from the probability of being alive and experiencing a payment of zero, but additionally from dying and therefore receiving nothing. If one is alive and payments are sufficient to satisfy the liquidity-need, they are perceived as gains. Therefore, less probability mass is assigned to losses and more probability mass is assigned to gains using the mixture CDF compared to a normal CDF. Incorporating the mixture CDF in the analysis, equation (21) becomes

$$CPT\left(Z_{i,t+\tau}\right) = _\tau p_x \left[\int_{-\infty}^{0^-} v\left(z\right) w^{-\prime}\left(F_{i,t+\tau}\left(z\right)\right) f_{\mathcal{N}_{i,t+\tau}}\left(z\right) dz + \int_{0^+}^{\infty} v\left(z\right) w^{+\prime}\left(1 - F_{i,t+\tau}\left(z\right)\right) f_{\mathcal{N}_{i,t+\tau}}\left(z\right) dz \right]$$
$$+ \left(1 - _\tau p_x\right) \left[v\left(0^-\right) w^{-\prime}\left(F_{i,t+\tau}\left(0^-\right)\right) + v\left(0^+\right) w^{+\prime}\left(1 - F_{i,t+\tau}\left(0^+\right)\right) \right] \tag{26}$$

with

$$w'\left(F\right) = \frac{F^{\theta-1}\left(F^\theta + (1-F)^\theta\right)^{-\frac{\theta+1}{\theta}}\left((F-1)(\theta-1)F^\theta + (F(\theta-1)-\theta)(1-\theta)^\theta\right)}{\theta - 1} \tag{27}$$

where $\theta \in (\gamma, \kappa)$. The first line of equation (26) is the utility in case of survival, whereas the

16

second line is the utility in case of death, which is 0, since $v(0) = 0$.

Finally, we numerically maximize equation (20) subject to the optimal level of tontine investment B_i.

$$\max_{B_i} MCPT(i)$$
$$\text{s. t. } B_i \leq W_i, B_i \geq 0 \tag{28}$$

4.7 Variation: Stochastic Liquidity Need in MCPT Valuation

From Equations (25) and (17), it follows that the combined payout of tontine and annuity is normally distributed with

$$Z_{i,t} \sim \mathcal{N}\left(\mu_{i,t}, \sigma_{i,t}^2\right).$$

To cover effects stemming from uncertainty about the future liquidity-need, we assume that the liquidity-need itself is normally distributed with mean $E(D_{i,t})$ and standard deviation $\sigma_{D_{i,t}}$, therefore

$$\mu_{i,t} \sim \mathcal{N}\left(\mu_{i,t_0}, \sigma_{\mu_{i,t}}^2\right)$$

where $\mu_{i,t_0} = q_{i,t} B_i + \frac{W_i - B_i}{a_{i,0}} - E(D_{i,t})$ and $\sigma_{\mu_{i,t}}^2 = \sigma_{D_{i,t}}^2$ is calibrated based on own empirical research. If we write

$$Z'_{i,t} = \left(Z'_{i,t} - \mu_{i,t}\right) + \mu_{i,t}$$

then

$$\left(Z'_{i,t} - \mu_{i,t}\right) \sim \mathcal{N}\left(0, \sigma_{i,t}^2\right)$$

and

$$Z'_{i,t} \sim \mathcal{N}\left(0, \sigma_{i,t}^2\right) + \mathcal{N}\left(\mu_{i,t_0}, \sigma_{\mu_{i,t}}^2\right) = \mathcal{N}\left(\mu_{i,t_0}, \sigma_{i,t}^2 + \sigma_{\mu_{i,t}}^2\right) \tag{29}$$

because the sum of two independent normally distributed random variables is also normally distributed. While the mean expected payout remains the same, the volatility increases to the sum of the volatility of the tontine payment and the volatility of the liquidity-need.

4.8 Variation: Subjective Mortality

To account for subjective beliefs about the own mortality risk, we adjust the objective forecasted mortality. It is important to understand the difference compared to the probability adjustment which CPT undertakes: while CPT accounts for a deviating perception of objective probabilities, the subjective mortality adjustment modifies the average probabilities subject to own percep-

tions about the individual health status. People who believe to live longer than the aggregate average because they feel very healthy or have an active lifestyle perceive to have lower death probabilities and thus believe to have a longer expected remaining lifetime. Therefore, the optimistic subjective death probability $q'_{x,t}$ is lower than the average, objective death probability $q_{x,t}$. Likewise, the pessimistic subjective death probability $q'_{x,t}$ for persons who believe to live shorter than the overall average (because of severe illness or the awareness of a poor lifestyle) is higher than the actual death probability $q_{x,t}$. Bissonnette et al. (2014) show that, within different groups (e.g. gender, ethnic background or education), people with similar characteristics are only slightly optimistic regarding their survival prospects compared to the average mortality within the subgroup, whereas the actual subgroup mortality itself differs tremendously from the overall population mortality. The authors conclude that the individual perceptions are very precise. Therefore, it is important to incorporate subjective survival probabilities in our analysis, because people who believe to live longer tend to live longer, and thus different retirement planning solutions are needed for different individuals. To account for subjective mortality in our model, we adjust the actual mortality rates $q_{x,t}$ by an individual mortality multiplier d, therefore the subjective mortality rate $q'_{x,t}$ is

$$
q'_{x,t} = \begin{cases} d \cdot q_{x,t} & \text{if } d \cdot q_{x,t} \leq 1 \\ 1 & \text{otherwise} \end{cases} \tag{30}
$$

where d is the realization of a random variable D and determines the subjective survival probability. For $0 < d < 1$ the individual expects to live longer than the average, if $d = 1$ the individual self assesses his or her lifetime of being average and if $d > 1$, the individual expects to live shorter than the actual mortality table predicts. Furthermore, $q_\Omega = 1$ which means that there is a limiting age Ω when the individual dies with certainty. A simple modeling approach for $d \sim D$ is shown in Appendix A.3, where D is modeled using a Gamma Distribution. Since the insurance company offering tontines and annuities uses average objective mortality rates, pricing is undertaken on the basis of average mortalities. The subjective beliefs only influence the subjective determination of individual utility.

5 Calibration and Results

We calibrate the Lee-Carter model based on data from the *Human Mortality Database*[15], and forecast mortality rates for $T = 100$ years beginning from 2011 which is denoted by $t = 1$ in the analysis. The fractions of female and male newborns we use to update population pyramids are calculated based on German birth statistics[16]. The age-specific birth rates are based on German birth statistics[17] and describe the number of newborns per year of a woman in each cohort. The maximum attainable age is set to be $\Omega = 105$ which means that at the age of $x = 105$ one dies with certainty. We consider initial wealth W_i as an independent variable in our analysis and measure its influence on the optimal investment behavior under various scenarios. The parameters of the polynomial liquidity-need function D_t are $\beta_0 = 163,984.686$, $\beta_1 = -4,000.634$ and $\beta_2 = 28.589$ and fit our empirically estimated old-age demand function based on SOEP data.

5.1 Base Case

In Table 1, we report the parameters used in the MCPT[18] analysis, which constitute the base case. We consider i as a male individual aged 62 in the year the tontine is set up ($t = 1$), and vary his initial endowment W_i. Furthermore, we assume that the remaining tontine members $k = 1 \ldots N, k \neq i$ behave optimally, i.e. the individual amounts B_k invested in the tontine are assumed to be the $MCPT_k$-utility maximizing amounts and are assumed to be uniformly distributed on $[0, 50,000]$.[19] Based on $M = 10,000$ simulations, we calculate the realized tontine returns for individual i in every period and thereby determine the moments of the normal approximation of tontine returns for member i. We set the subjective discount factor $\delta = 1$, because we assume that the future states are as important as present states for an individual who aims to secure the future standard of living[20]. We calibrate the CPT parameters a, b, λ, γ and κ according to the values proposed by Tversky and Kahneman (1992). The expected return and

---

[15] Data from 2011, Source: `http://www.mortality.org`.

[16] Data from 2000 - 2010, see Statistisches Bundesamt (2012).

[17] Data from 2011, Statistisches Bundesamt, `https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Bevoelkerung/Geburten/Tabellen/GeburtenzifferAlter.html`.

[18] We use the notation MCPT according to Equation (20) for the sum of the periodic CPT utilities and CPT according to Equation (26) for the periodic utilities.

[19] There exists a corresponding amount of initial wealth W_k for which B_k is optimal. We make this assumption due to computational reasons. The values correspond roughly to the optimal amount invested in the tontine for the considered male individual for the considered range of initial wealth W_i. Thereby we provide the optimal investment decision for hypothetical levels of W_k, which are roughly in the same range as the bandwidth of W_i.

[20] In this sense, Parsonage and Neuburger (1992) and Van der Pol and Cairns (2000) provide empirical evidence that it is feasible to assume a subjective discount rate of zero for the discounting of future health benefits.

the standard deviation in selected periods, if we set W_i such that the optimal $B_i = 30,000$ can be seen in Table 2 in Appendix B. Since the expected tontine return as well as the tontine volatility are driven by the individual death probability, both increase as i becomes older. Aged 62 in $t = 1$, person i can expect to receive a first-year tontine return of 345.15 EUR which amounts to 1.15% of the initial tontine investment of $B_i = 30,000$ EUR. In $t = 10$, his expected return is roughly 1.6 times higher than in the first year but still amounts to only 1.86% of the initial investment. The payments thus increase slowly in the early retirement years because of the slow increase of death probabilities in the early years. After 20 years, at the age of 81, the return is already 4.8 times as large as in the first year and the single payment in this year amounts to 5.51% of the initial investment. For very high ages, the payments increase tremendously: at the age of 91, in $t = 30$, the tontine return is almost 16.3 times as large as in the first year and yields 18.73%. Every year of further survival then yields even steeper increasing returns, being 47.69% of the initial investment at the age of 101 in $t = 40$, and finally 100% at the maximum attainable age of 105 in $t = 44$. Neglecting interest rate effects, person i can expect to recoup his initial investment in year 26 at the age of 87. Since the standard deviation of the tontine returns depends on the individual mortality, volatility increases similarly with age. In comparison, an immediate fair annuity with investment of 30,000 EUR for person i would yield an income of yearly 1458.96 EUR. Therefore, the investor could recoup his initial investment already after 21 years at the age of 82. This is because the annuity provides stable payments whereas tontine payments increase with age. This first result indicates that a tontine might be the preferred product for someone who expects to live long. Figure 3 shows the yearly expected payout patterns of a tontine and an annuity with investment volume normalized to unity. The important question then is which combination of both instruments is best suited to meet the old-age liquidity-need?
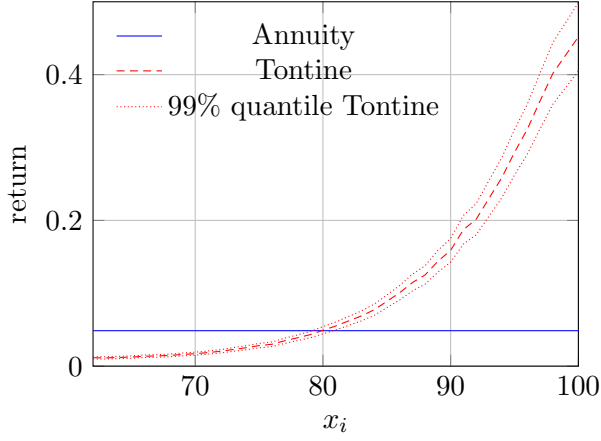
Figure 3: Normalized Payout of Annuity vs Tontine

Parameter	Notation	Value
Forecast horizon (in years)	T	100
Maximum attainable age	Ω	105
Fraction of female newborns	ω^f	48.68 %
Fraction of male newborns	ω^m	51.32 %
Lower boundary age at tontine entrance	$\underline{x}$	62
Upper boundary age at tontine entrance	$\overline{x}$	100
Size of the tontine	N	10,000
Monte Carlo Paths	M	10,000
Subjective discount factor	δ	1
CPT value function parameters	a, b	0.88
CPT loss sensitivity factor	λ	2.25
CPT w^+ parameter	γ	0.61
CPT w^- parameter	κ	0.69

Table 1: MCPT parameters Base Case

Based on these considerations, we determine the CPT utility $CPT(Z_{i,t})$ in each period for different levels of W_i for member i. Figure 4a shows the CPT utility of member i for an initial wealth $W_i = 610,000$ EUR in each point in time t which is the expected contribution of the CPT utility ton the aggregated MCPT utility. Since survival probabilities decline with age, the impact of each CPT utility declines with age and finally converges toward zero. If available funds are higher (lower) than the age-increasing liquidity-need, which we incorporate as time-changing reference point, funds are valued as gains (losses). We first consider the case where person i completely annuitizes his initial wealth (solid blue line). As annuity payments are constant, an increasing liquidity-need $D_{i,t}$ causes declining CPT utility. In early years, demand can be met. As the liquidity-need rises, it exceeds the available funds and CPT utility decreases and finally becomes negative. At the same time, as age increases, the declining survival probability causes a a lower CPT utility which reduces the impact of late periods on MCPT utility. For the very late years, the huge decline of survival probabilities outweighs the decreasing CPT utility. Finally,
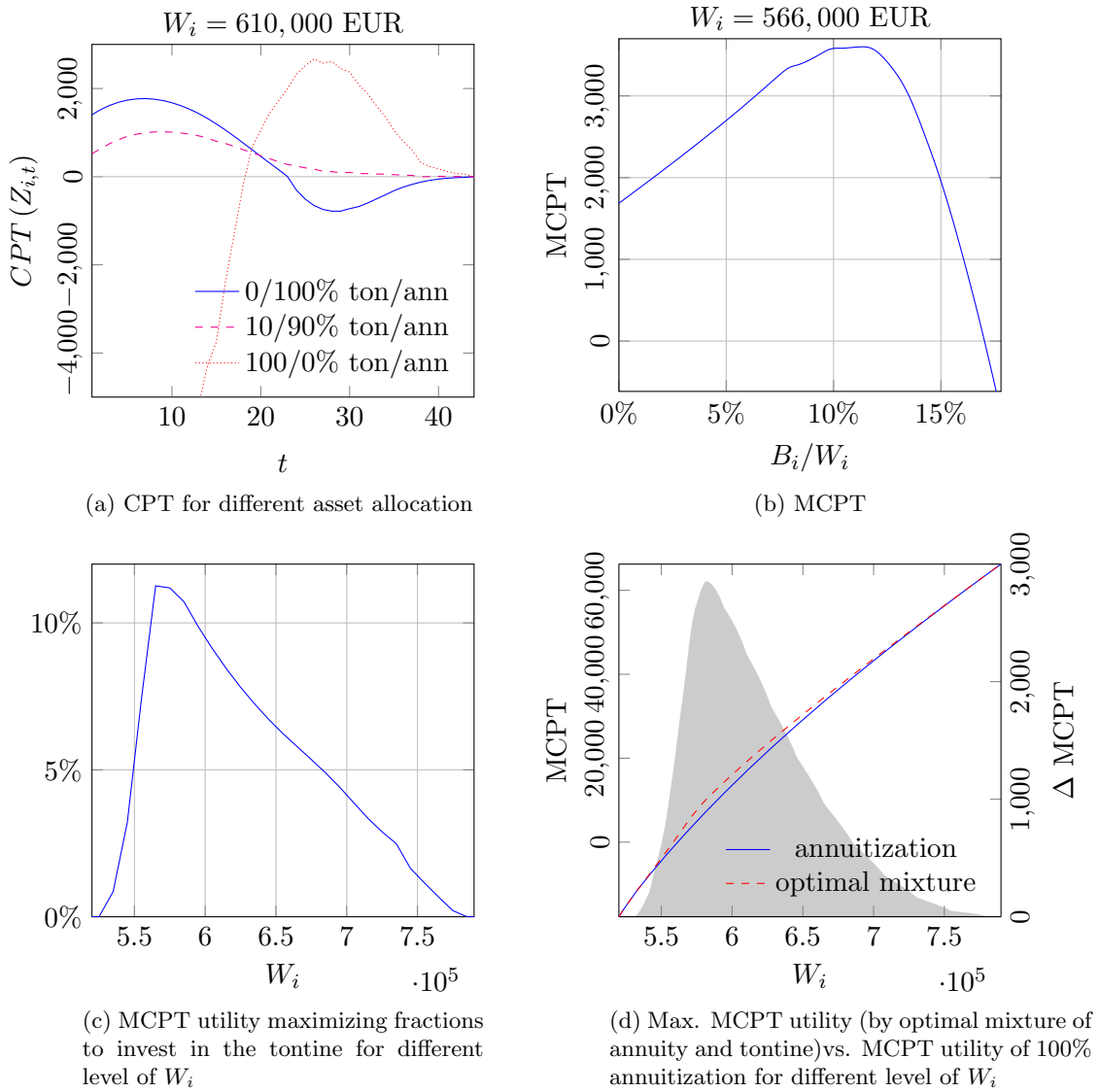
$W_i = 610,000$ EUR

$CPT\,(Z_{i,t})$

— 0/100% ton/ann
-- 10/90% ton/ann
⋯ 100/0% ton/ann

(a) CPT for different asset allocation

$W_i = 566,000$ EUR

MCPT

B_i/W_i

(b) MCPT

(c) MCPT utility maximizing fractions to invest in the tontine for different level of W_i

(d) Max. MCPT utility (by optimal mixture of annuity and tontine)vs. MCPT utility of 100% annuitization for different level of W_i

— annuitization
-- optimal mixture

Figure 4: Base Case

the impact approaches zero. Second, we consider the complete tontinization of initial wealth (densely dotted red line). Since tontine payments are driven by mortality, payments are very low in the early years and increase in age, thus the liquidity-need cannot be met in early ages and can easily satisfy $D_{i,t}$ in later years. Tontine and annuity payments proceed adversely and a portfolio of both can help to generate payout patterns which enable the increasing liquidity demand to be financed appropriately. The dashed magenta line shows the CPT utilities of a payout pattern of a portfolio consisting of 10% tontine and 90% annuity. While still being able to satisfy the demand at the early ages, it is also able to provide sufficient funds in the later years. The sum of the CPT utilities yields the MCPT utility, and we are searching for those combinations of tontine and annuity investments for different levels of W_i, which maximize MCPT utility.

Figure 4b exemplarily shows the MCPT utility for different fractions of $W_i = 566,000$ EUR

invested being in the tontine (B_i). The remaining fraction $W_i - B_i$ is annuitized. Starting from a situation of complete annuitization, i can increase his MCPT utility by investing a positive fraction in the tontine, and finally maximizes his MCPT utility if he invests 11.396% of W_i in the tontine. An optimal fraction exists because of two counteracting effects: A higher investment in the tontine increases the later years' CPT utilities more than it decreases the early years' CPT utilities. Up to an optimal point, the increase in CPT utility in the late years outweighs the decrease in CPT utility in early years. Beyond this optimal point, the decrease in CPT utility in early years outweighs the increase in CPT utility in the late years, yielding a declining MCPT utility. These effects are resulting from the fact that up to the age of 80, the annuity provides a higher return than the tontine, while beyond the age of 80, the tontine outperforms the annuity. Therefore, one unit of additional investment in the tontine decreases CPT utilities until the age of 80 and increases CPT utilities beyond the age of 80, finally yielding an optimal, MCPT utility maximizing tontine investment level.

Figure 4c shows the optimal, MCPT utility maximizing fractions to be invested in the tontine for different levels of initial wealth W_i. If $W_i < 532,000$ EUR, it is not optimal to invest anything in the tontine. This is because even for complete annuitization, annuity payments are so low that the CPT utility losses in early years, caused by investing in the tontine, are large and cannot be offset by the CPT utility gains in later years, caused by increasing tontine payouts. As W_i increases, the optimal fraction to invest in the tontine increases very sharply up to $W_i = 566,000$, EUR and decreases thereafter. In this wealth region the reduction in early years CPT utilities due to shifting from annuity to tontine investment can be overcompensated by the increase in late years' CPT utilities. This is because the marginal CPT utility in early years is lower for higher W_i, and therefore more wealth can be shifted from the annuity to the tontine investment. The optimal tontine fraction decreases beyond the peak at $W_i = 566,000$ because for higher W_i, marginal CPT utility decreases for late years' consumption and less wealth in relative terms is needed to increase late years' CPT utilities. In other words, the CPT utilities in early years' do not decline much, while late years' consumption can still be financed with the additional tontine payments. For $W_i > 779,000$ EUR it is again optimal not to invest at all in the tontine. This is because at this wealth level, the annuity payments are sufficient to satisfy the liquidity-need in early as well as in later years. An investment in the tontine thus would reduce early consumption possibilities and therefore reduce early years CPT utility, while the gain from later consumption would be very small because later years' liquidity-need can already be met by the annuity payments. Therefore, the tontine would take away funds in early years in which survival

prospects are high and therefore negatively impact utility. In turn, the tontine would provide funds in states when the additional tontine payments are not needed because funds from the annuity payments are already sufficient. In addition, these funds hardly contribute to MCPT utility because of low survival prospects at high ages.

Figure 4d shows the optimal MCPT utility for different levels of W_i compared to the MCPT utility under complete annuitization. For $W_i < 532,000$ EUR and $W_i > 779,000$ EUR, complete annuitization provides the highest MCPT utility. As seen before, in these domains tontine investment reduces the MCPT utility. Therefore the optimal decision and complete annuitization correspond. For $W_i \geq 532,000$ EUR and $W_i \leq 779,000$ EUR, the highest MCPT utility can be achieved by investing an optimal fraction in the tontine. The highest MCPT utility increase can be generated at $W_i = 582,000$ EUR, while the utility increase decelerates for lower and higher amounts of W_i. This can be seen in the gray shaded area, which corresponds with the scale on the right hand side of the figure.

5.2   Variation: Expected Utility Theory (EUT) Calibration

If we set the parameters of the value function of the CPT to $a = b = 0.5$, we receive a square root utility. In addition, setting the parameters of the weighting function to $\gamma = \kappa = 1$ yields actual occurrence probabilities instead of their subjective perceptions. Appendix A.4 shows the resulting model, which is an expected utility calibration with objective probability weights. As Figure 5 shows, the expected utility-maximizing fractions of tontine investment are generally similar compared to the base case setting. It is striking that at $W_i = 590,995.07$ EUR the optimal fraction to invest in the tontine immediately jumps from 0 to 10,80% and decreases thereafter, until it finally reaches 0 again at $W_i = 950,158.80$ EUR. Compared to the base case, investment in the tontine is optimal for higher W_i than in the base case. Furthermore, the maximum tontine investment is slightly lower. Utilizing the EUT calibration, we can explain the immediate jump in the optimal fraction of tontine investment, which appears similar to the base case.

Figure 6a shows the expected utility values for $W_i = 570,000$ EUR for different levels of tontine investment. The highest expected utility can be achieved if no investment in the tontine takes place. If the tontine investment increases, the expected utility first decreases and at roughly 12%, there is a little peak with a local maximum where expected utility slightly increases, but finally decreases thereafter again. As shown in figure 6b, for a 10,000 EUR higher initial wealth $W_i = 580,000$ EUR, two changes in the shape of the expected utility curve can be
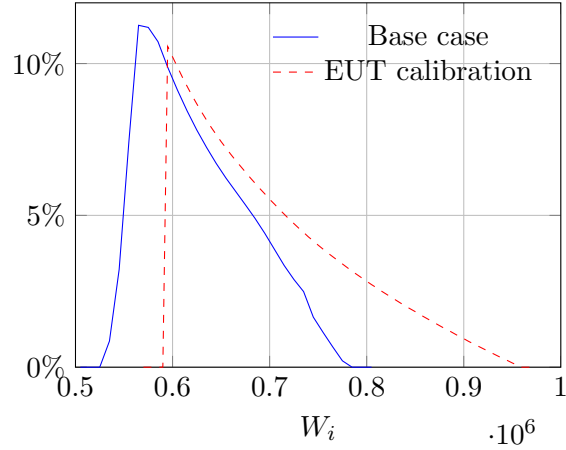
Figure 5: Expected Utility Theory calibration maximizing fractions to invest in the tontine for different levels of W_i
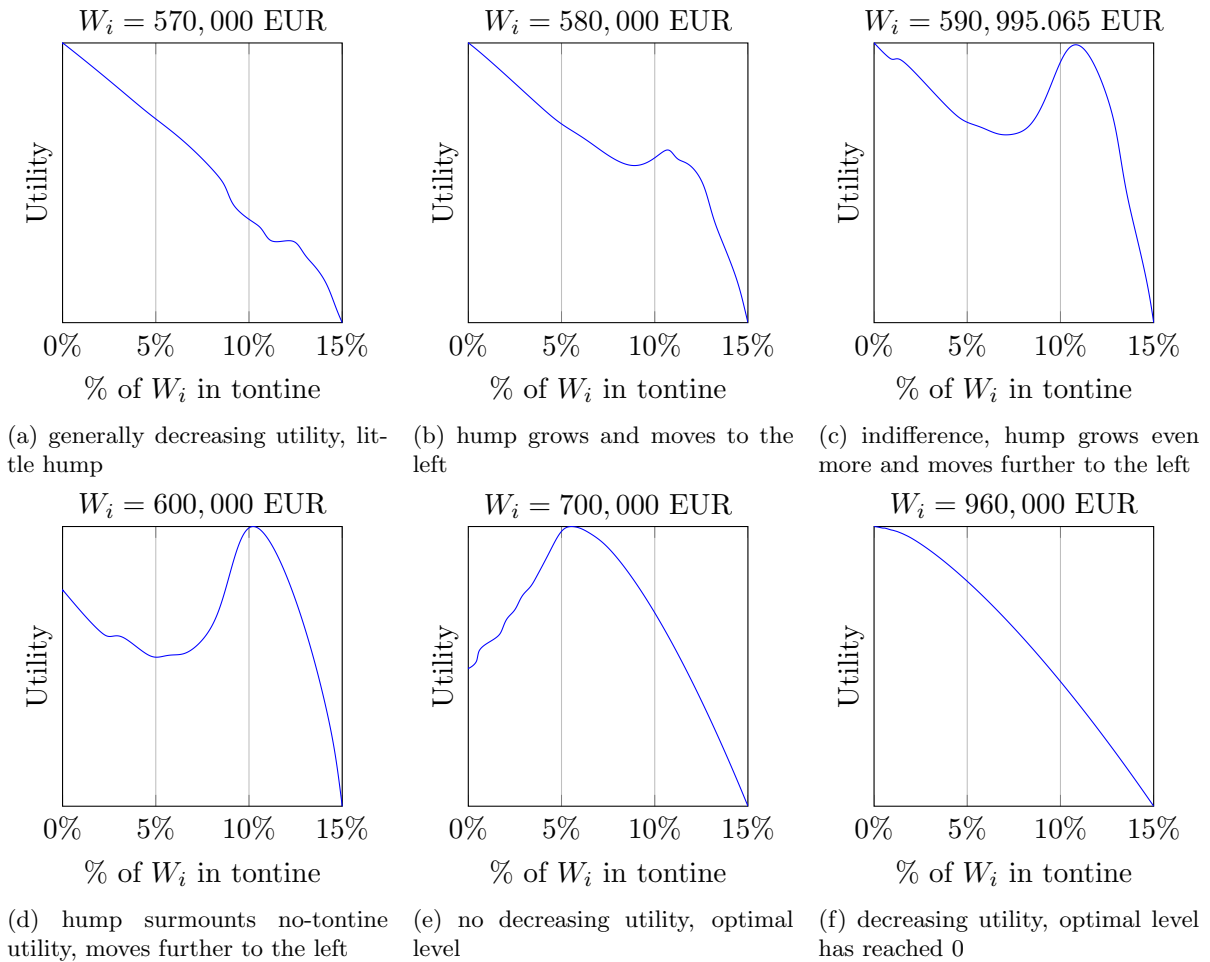

(a) generally decreasing utility, little hump

(b) hump grows and moves to the left

(c) indifference, hump grows even more and moves further to the left

(d) hump surmounts no-tontine utility, moves further to the left

(e) no decreasing utility, optimal level

(f) decreasing utility, optimal level has reached 0

Figure 6: Expected Utility for different levels of W_i and increasing fractions of tontine investment relative to W_i for square-root utility calibration of the base case

observed: first the hump increases, meaning that the positive influence on expected utility of tontine investments increases, and second, the hump moves a bit to the left compared to the previous wealth level, meaning that the local maximum occurs for lower fractions of tontine investment, compared to the situation presented in 6a. As initial wealth reaches the threshold value $W_i = 590,995.065$ EUR (Figure 6c), the hump is as large as that the expected utility with 10.82% tontine investment equals the expected utility without tontine investment. Therefore, the individual is indifferent between no tontine investment and 10.82% tontine investment. For a tontine investment between 0 and 10.82%, the expected utility is lower compared to the maximum expected utility. For tontine investments larger than 10.82%, the expected utility decreases as well. As W_i further increases, the peak further moves to the left and surmounts the expected utility without tontine investment (Figure 6d). Gradually, the local minimum between no tontine investment and optimal tontine investment disappears (Figure 6d). Finally, as W_i is very large, the slope around the local maximum is very flat and finally disappears when the expected utility maximizing fraction to invest in the tontine hits 0 again (Figure 6f).

5.3 Variation: Tontine Size

For an increased tontine size of $N = 100,000$ (compared to $N = 10,000$ in the base case), the volatility of the tontine payments declines (Table 2 in Appendix B). As presented in Figure 7, less volatile tontine payments cause that it is optimal to invest in the tontine for lower W_i than in the base case scenario. Similarly, for higher W_i, investing in the tontine remains beneficial with an increased pool size. This is because less volatile payments generally enhance CPT utilities. Therefore, it is optimal for both a lower and a higher W_i to invest in the tontine compared to $N = 10,000$.
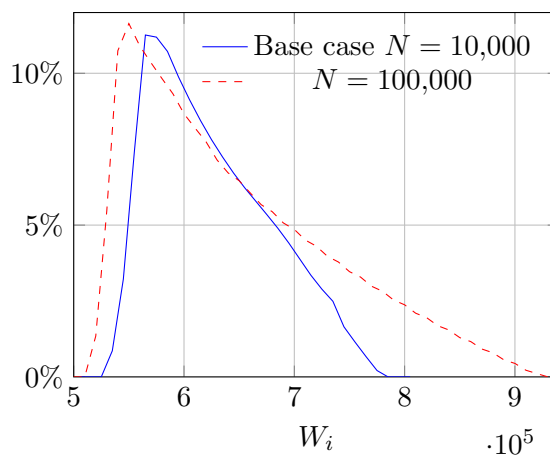

Figure 7: MCPT utility maximizing fractions to invest in the tontine for different levels of W_i for $N = 100,000$

5.4 Variation: Stochastic Liquidity Need

If we assume a stochastic liquidity-need, the fact whether cash flows lead to gains or losses with respect to the liquidity-need is affected by the volatility of the tontine payments as well as by the volatility of the liquidity-need. A stochastic liquidity-need increases the overall volatility and therefore CPT utilities decline. First, we set the variance of the liquidity-need at $\sigma^2_{\mu_{i,t}} = \sigma^2_{D_{i,t}} = 2,500^2$. As a consequence, the wealth level at which it becomes optimal to invest in the tontine increases compared to the base case (see Figure 8). Furthermore, it remains optimal to invest in the tontine for higher W_i compared to the base case. The reason for this lies in the fact that in early years, the more volatile nature of gains and losses makes it desirable to hold more funds to be able to fulfill the liquidity-need. Every unit taken away from the annuity in the early years causes a huge decline in early years CPT utilities. Therefore, it is optimal only for a higher W_i to invest in the tontine. The opposite effect applies to high W_i. The volatile liquidity-need brings about situations of high liquidity-need in which payments from the tontine can support its coverage. Therefore, it is optimal for a higher W_i to hold some fraction of the tontine. Furthermore, the optimal level of tontine investment is lower compared to the base case because the tontine investment itself adds another layer of volatility to the payments, which decreases utility. As we further increase the volatility of the liquidity-need to $\sigma^2_{\mu_{i,t}} = \sigma^2_{D_{i,t}} = 5,000^2$, we can observe a boost in all three effects. A higher W_i is required to start investing in the tontine in order to lower the risk of experiencing a utility-harming drop far below the liquidity-need. As the level of W_i is relatively high, the tontine investment loses its efficiency compared to the resulting annuity payments, yielding a lower optimal fraction to be invested in the tontine. Nevertheless, due to the higher volatility, it is still optimal for a higher W_i to invest in the tontine in order to be able to absorb shocks of the liquidity-need.
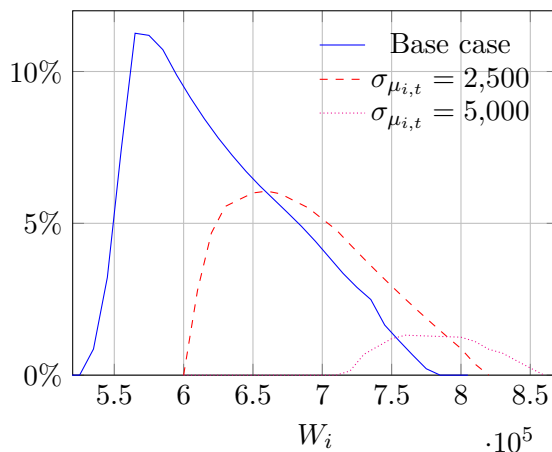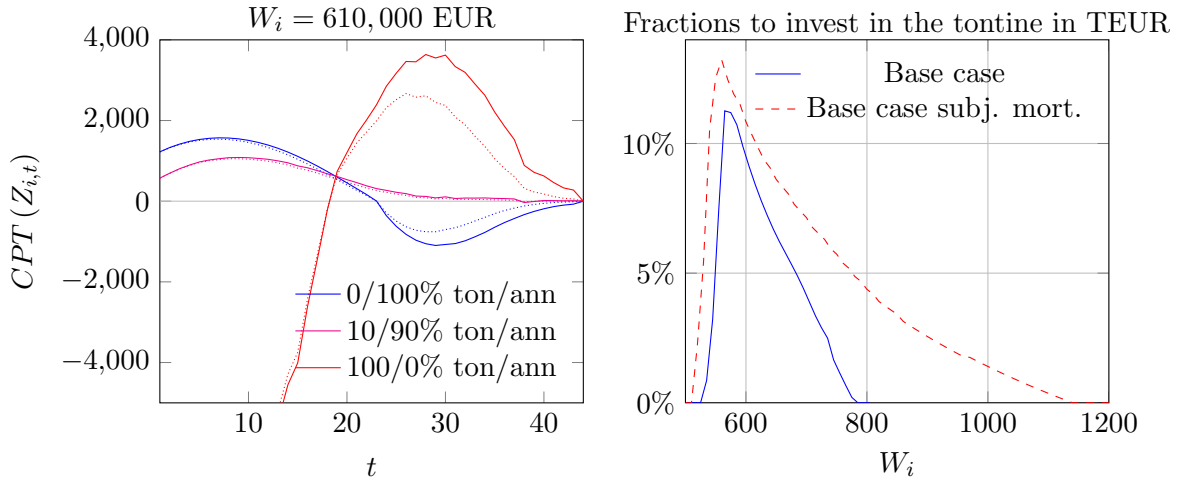
Figure 8: MCPT utility maximizing fractions to invest in the tontine for different levels of W_i for stochastic liquidity-need

5.5 Variation: Subjective Mortality

If we adjust the mortality to $d = 0.8$, the individual expects to live longer than average. This means that future periods have a greater impact on MCPT utility because survival probabilities decline less fast. Therefore, later years CPT utilities are higher compared to the base case. This situation is presented in figure 9a, where the dotted lines represent the CPT utility paths for the base case and the solid lines represent the CPT utility paths for the subjective, improved mortality. As a result, positive and negative subjective CPT utilities both have a higher impact on total MCPT utility compared to the base case, indicating that it might be more favorable to invest a higher fraction in the tontine because it is more likely to experience the later years' CPT utilities. Figure 9b shows that it is optimal to invest in the tontine for lower W_i because, by investing in the tontine later years' CPT utilities gain more relevance and are higher although early years' CPT utilities are reduced. By investing more intensely in the tontine, overall MCPT utility can be increased. Furthermore, it is optimal to invest in the tontine up to a higher W_i, compared to the base case. This is because marginal CPT utility in later years increases as survival probabilities increase. Early years CPT utility losses can be overcompensated by later years CPT utility gains. Furthermore, later years' CPT utility losses also have a higher impact on the MCPT and, therefore, a tontine investment in higher W_i regions can help to mitigate the otherwise resulting underfunding problem. In addition, it is optimal to invest a higher fraction in the tontine for all W_i's for which it is optimal to invest in the base case. This is due to the increased probability of experiencing CPT utilities in the late years.

28

(a) CPT for different asset allocation, actual vs subjective mortality ($d = 0.8$)

(b) MCPT utility maximizing fractions to invest in the tontine for different levels of W_i for actual and subjective mortality ($d = 0.8$)

Figure 9: Subjective Mortality Beliefs

5.6 Variation: Changing Liquidity Need

In this section we change the shape of the liquidity-need. First, we parallel shift the standard retirement smile curve up by 10,000 EUR. Second, we assume an exponential growth of the standard retirement smile by $D_t^{exp} = 1.01^t D_t$. Since the standard retirement smile represents the average liquidity-need unconditioned on the health status, an exponential growth can be interpreted as the liquidity-need conditional on bad health. Figures 10a and 10b show the resulting liquidity-need curves and the optimal fractions to invest in the tontine for the different liquidity-need curves. If we assume a parallel, upward shift of the liquidity-need curve by 10,000 EUR, two characteristics of the optimal investment choice can be observed. First, the optimal investment pattern shifts to the right, which means that it is optimal to invest a fraction in the tontine only for higher initial wealth endowment W_i. Second, the optimal fractions to invest in the tontine are lower compared to the base case. The reason for these two properties lies in the increasing liquidity-need in every period. For a relatively low W_i, it is not optimal to invest in the tontine because the loss in CPT utilities in the early years due to a reduction of annuitized wealth exceeds the CPT utility gains in later years. Only if there is sufficient initial wealth it is optimal to invest in the tontine. The maximal tontine investment in the parallel shift case is lower compared to the peak in the base case. This is a consequence of the rightward shift of the optimal tontine investment curve. A lower fraction of the tontine is necessary to be invested in the tontine simply because, on an absolute level, more initial wealth is available. As a consequence, a lower fraction invested in the tontine is able to generate enough income to provide higher late years CPT utilities than the reduction in early years consumption decreases

(a) Variations of the liquidity-need curve

(b) MCPT utility maximizing fractions to invest in the tontine for different levels of W_i for different shapes of Liquidity need
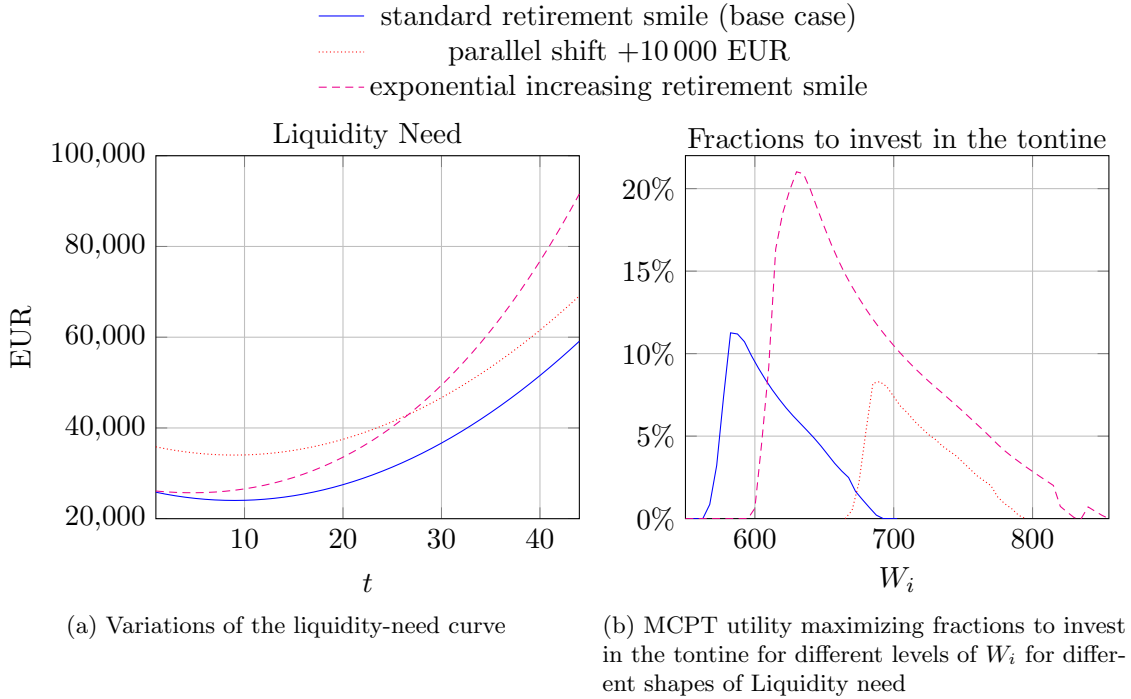
Figure 10: Variations in Liquidity Need Shape

overall MCPT utility. If we assume an exponentially increasing retirement smile, investment in the tontine starts for a higher W_i compared to the base case and below the parallel shift case at approximately $W_i = 600,000$ EUR. Furthermore, the peak of the optimal amount to be invested in the tontine is almost twice as large compared to the base case. Optimal positive fractions of tontine investment persist longer for high W_i. This is because in the early years the liquidity-need in the exponential case is relatively close to the base case and disproportionately increases with age, compared to the base case. Therefore, the CPT utility decrease in early years is relatively low when investing some fraction in the tontine, while the CPT utility gains of the tontine investment in the late years are very high. Thus, with large amounts invested in the tontine, the early years' CPT utilities do not suffer much, while later years' CPT utilities benefit strongly. As a consequence, larger amounts to be invested in the tontine are optimal to satisfy the liquidity-need best. To sum up, the tontine is most powerful if the liquidity-need is low in the early years and high in the later years of retirement.

6 Summary and Conclusion

The changing social, financial and regulatory framework, such as an increasingly aging society, the current low-interest environment, as well as the implementation of risk-based capital standards in the insurance industry, lead to the search for new product forms for private pension provision. These product forms should reduce or avoid investment guarantees and risks stem-

ming from longevity, still provide reliable insurance benefits and at the same time reflect in the payout pattern the increasing financial resources required for very high ages. We propose the traditional tontine to serve as such "product innovation", especially in combination with a traditional life annuity.

To assess the effects of tontine investments on policyholders' welfare, we develop a model by which individual old-age liquidity-need and payouts stemming from annuity and tontine investment can be evaluated and result in an optimal retirement planning decision, based on individual preferences, characteristics and subjective mortality beliefs.

To show the effects of a tontine investment on retirement planning, we model the development of the changing population structure for the next 10 decades in Germany. Based on the changing mortality dynamics, we describe a fair revolving tontine. To assess its advantages and disadvantages compared to a traditional life annuity, we derive a targeted consumption level from empirical data and compare the tontine payout structure with the payout structure of a traditional annuity with regard to the ability of meeting the desired consumption.

Our results reveal that a portfolio of annuity and tontine can provide the highest expected MCPT utility. While the annuity pays a stable, constant pension, the tontine provides volatile, age-increasing payouts. The results of our analyses prove to be sensitive with respect to the initial wealth endowment and the subjective expectation about the remaining lifetime of an individual. For very low and very high endowments, complete annuitization is optimal, whereas for medium endowments of initial wealth, it might be optimal to invest some fraction in the tontine, depending on individual circumstances. Taking these circumstances into account, our results indicate that, from a policyholder perspective, a tontine can be a beneficial supplement to existing retirement planning solutions.

Future research could incorporate investment risk and analyze its effects on the optimal tontine investment decision. The level of investment risk might significantly change the optimal allocation of retirement wealth. In this context, the integration of reinvestment opportunities of tontine and annuity returns might yield an additional determining factor for the optimal retirement planning decision. Another interesting area for further research is the analysis from an insurer's perspective. It will be interesting to analyze whether the supply of tontines can reduce the insurer's capital requirements, or reduce safety loadings included in annuity prices, since the tontine involves no longevity risk for the provider, and might partially substitute annuities.

References

Ågren, M. (2006). Prospect theory and higher moments. Technical report, Working Paper, Department of Economics, Uppsala University.

Banks, J., Blundell, R., and Tanner, S. (1998). Is there a retirement-savings puzzle? *The American Economic Review*, 88(4):769–788.

Bernheim, B. D., Skinner, J., and Weinberg, S. (2001). What accounts for the variation in retirement wealth among u.s. households? *The American Economic Review*, 91(4):832–857.

Bissonnette, L., Hurd, M. D., and Michaud, P.-C. (2014). Individual survival curves comparing subjective and observed mortality risks.

Blanchet, D. (2013). Estimating the true cost of retirement. `https://corporate.morningstar.com/ib/documents/MethodologyDocuments/ResearchPapers/Blanchett_True-Cost-of-Retirement.pdf`.

Brockwell, P. J. and Davis, R. A. (2013). *Time series: theory and methods.* Springer Science & Business Media.

Bundesanstalt für Finanzdienstleistungsaufsicht (2014). *Statistik der Bundesanstalt für Finanzdienstleistungsaufsicht, Erstversicherungsunternehmen und Pensionsfonds.* BaFin, Stabsstelle Presse-u. Öffentlichkeitsarbeit.

Cairns, A. J., Blake, D., and Dowd, K. (2008). Modelling and management of mortality risk: a review. *Scandinavian Actuarial Journal*, 2008(2-3):79–113.

Forman, J. B. and Sabin, M. J. (2014). Tontine pensions: A solution to the state and local pension underfunding crisis. *University of Pennsylvania Law Review*, 163.

Graham, R. L., Grotschel, M., Lovász, L., and Wilf, H. S. (1997). Handbook of combinatorics. *Mathematical Intelligencer*, 19(2):65.

Gruenberg, E. M. (1977). The failures of success. *The Milbank Memorial Fund Quarterly. Health and Society*, 55(1):3–24.

Haider, S. J. and Stephens Jr, M. (2007). Is there a retirement-consumption puzzle? evidence using subjective retirement expectations. *The Review of Economics and Statistics*, 89(2):247–264.

Hamermesh, D. S. (1984). Consumption during retirement: The missing link in the life cycle. *The Review of Economics and Statistics*, 66(1):1–7.

Hens, T. and Rieger, M. O. (2010). *Financial economics: A concise introduction to classical and behavioral finance.* Springer Science & Business Media.

Hoermann, G. and Ruß, J. (2008). Enhanced annuities and the impact of individual underwriting on an insurer's profit situation. *Insurance: Mathematics and Economics*, 43(1):150–157.

Hong, Y. (2011). On computing the distribution function for the sum of independent and nonidentical random indicators. *Dep. Statit., Virginia Tech, Blacksburg, VA, USA, Tech. Rep. 11_2.*

Hong, Y. and Meeker, W. Q. (2010). Field-failure and warranty prediction based on auxiliary use-rate information. *Technometrics*, 52(2):148–159.

Hong, Y., Meeker, W. Q., and McCalley, J. D. (2009). Prediction of remaining life of power transformers based on left truncated and right censored lifetime data. *The Annals of Applied Statistics*, 3(2):857–879.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Kochskämper, Susanna und Pimpertz, J. (2015). Herausforderungen an die Pflegeinfrastruktur. *IW-Trends Vierteljahresschrift zur empirischen Wirtschaftsforschung*, 42(3).

Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting us mortality. *Journal of the American statistical association*, 87(419):659–671.

Manes, A. (1932). Versicherungswesen, iii. *Bd., Personenversicherung, Teubner, Leipzig und Berlin.*

Mariger, R. P. (1987). A life-cycle consumption model with liquidity constraints: Theory and empirical results. *Econometrica*, 55(3):533–557.

McKeever, K. (2009). Short history of tontines, a. *Fordham J. Corp. & Fin. L.*, 15:491.

Milevsky, M. (2015). *King William's tontine: why the retirement annuity of the future should resemble its past.* Cambridge University Press.

Milevsky, M. A. (2006). *The calculus of retirement income: Financial models for pension annuities and life insurance.* Cambridge University Press.

Milevsky, M. A. and Salisbury, T. S. (2015). Optimal retirement income tontines. *Insurance: Mathematics and Economics*.

Parsonage, M. and Neuburger, H. (1992). Discounting and health benefits. *Health economics*, 1(1):71–76.

Renshaw, A. and Haberman, S. (2008). On simulation-based approaches to risk measurement in mortality with specific reference to poisson lee–carter modelling. *Insurance: Mathematics and Economics*, 42(2):797–816.

Renshaw, A. E. and Haberman, S. (2003). Lee–carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, 33(2):255–272.

Ruß, J. and Schelling, S. (2015). Prospect theory and the demand for cliquet-style guarantees. *Working Paper*.

Sabin, M. J. (2010). Fair tontine annuity. *Social Science Research Network Working Paper Series*.

Sabin, M. J. (2011). A fast bipartite algorithm for fair tontines. *Social Science Research Network Working Paper Series*.

Schmidt, U. (2015). Insurance demand under prospect theory: A graphical analysis. *Journal of Risk and Insurance*.

Standard Life (2013). The retirement smile. `http://ukgroup.standardlife.com/content/news/new_articles/2013/260613NewDrawdownTransferOption.xml`.

Statistisches Bundesamt (2012). *Geburten, Sterbefälle, Eheschließungen*. Auszug aus Wirtschaft und Statistik. Statistisches Bundesamt.

Statistisches Bundesamt (2015). Bevölkerung Deutschlands bis 2060: 13. koordinierte Bevölkerungsvorausberechnung. `https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/VorausberechnungBevoelkerung/BevoelkerungDeutschland2060Presse5124204159004.pdf?__blob=publicationFile`.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323.

Van der Pol, M. M. and Cairns, J. A. (2000). Negative and zero time preference for health. *Health Economics*, 9(2):171–175.

Volkova, A. Y. (1996). A refinement of the central limit theorem for sums of independent random indicators. *Theory of Probability & Its Applications*, 40(4):791–794.

Willets, R. (2004). The cohort effect: insights and explanations. *British Actuarial Journal*, 10(04):833–877.

Worldbank (2015). Databank. Life expectancy at birth. `http://data.worldbank.org/indicator/SP.DYN.LE00.IN`.

Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *The Review of Economic Studies*, 32(2):137–150.

A   Appendix

A.1   Tontine Volatility

The variance of the return conditional that i survives in t is

$$Var\left[r_{i,t}|A_{i,t}^{\mathsf{c}}\right] = E\left[r_{i,t}^2|A_{i,t}^{\mathsf{c}}\right] - \left(E\left[r_{i,t}|A_{i,t}^{\mathsf{c}}\right]\right)^2$$

$$= E\left[\left(\sum_{\substack{k=1\\k\neq i}}^{N} a_{i,k,t}B_k\mathbf{1}_{\{A_{k,t}\}}\right)^2\right] - (q_{i,t}B_i)^2 .$$

If we assume constant $\bar{a}_{i,k,t}$'s, the variance can be expressed as

$$Var\left[r_{i,t}|A_{i,t}^{\mathsf{c}}\right] = q_{i,t}^2 B_i^2 \frac{\displaystyle\sum_{\substack{k=1\\k\neq i}}^{N} B_k^2 E\left[\mathbf{1}_{\{A_{k,t}\}}^2\right]}{\displaystyle\sum_{\substack{k=1\\k\neq i}}^{N} B_k^2 q_{k,t}^2 + \sum_{\substack{k=1\\k\neq i}}^{N}\sum_{\substack{k=1\\j\neq i}}^{N} B_k B_j q_{k,t} q_{j,t}}$$

$$+ q_{i,t}^2 B_i^2 \frac{\displaystyle\sum_{\substack{k=1\\k\neq i}}^{N}\sum_{\substack{k=1\\j\neq i}}^{N} B_k B_j E\left[\mathbf{1}_{\{A_{k,t}\}}\right] E\left[\mathbf{1}_{\{A_{j,t}\}}\right]}{\displaystyle\sum_{\substack{k=1\\k\neq i}}^{N} B_k^2 q_{k,t}^2 + \sum_{\substack{k=1\\k\neq i}}^{N}\sum_{\substack{k=1\\j\neq i}}^{N} B_k B_j q_{k,t} q_{j,t}} - q_{i,t}^2 B_i^2 .$$

For $N \to \infty$

$$Var\left[r_{i,t+1}|A_{i,t}^{\mathsf{c}}\right] = q_{i,t}^2 B_i^2 - q_{i,t}^2 B_i^2 = 0$$

because $E\left[\mathbf{1}_{\{A_{k,t}\}}\right] = q_{k,t}$. Nevertheless, the tontine we employ is large, but still risky[21]. For a fixed tontine size N, the tontine payouts are still volatile and follow a Poisson binomial distribution in each t which we approximate by a normal distribution[22] $\mathcal{N}\left(\mu_{i,t}, \sigma_{i,t}\right)$ for large N where $\mu_{i,t} = q_{i,t} B_i$ and

$$\sigma_{i,t} = \frac{q_{i,t}Bi}{\sqrt{M-1}}\sqrt{\sum_{m=1}^{M}\left(\frac{\displaystyle\sum_{\substack{k=1\\k\neq i}}^{N} B_k\mathbf{1}_{\{A_{k,t}\}}^m}{\displaystyle\sum_{\substack{k=1\\k\neq i}}^{N} B_k q_{k,t}} - 1\right)^2} .$$

with M Monte Carlo simulation paths.

---

[21] At a tontine size of $N = 100,000,000$ members, the volatility would be negligible. Of course, a tontine of this size is not realistic. Therefore we design a tontine of a size which might be practically realizable and therefore comprises of significant volatility.

[22] see for example Volkova (1996), Hong et al. (2009), Hong and Meeker (2010) and Hong (2011).

A.2 Normal Distribution of Gains and Losses

The CDF of the normally distributed gains and losses is

$$F_{\mathcal{N}_{i,t+\tau}}(u) = \frac{1}{\sigma_{i,t+\tau}\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}\left(\frac{s-\mu_{i,t+\tau}}{\sigma_{i,t+\tau}}\right)^2} ds. \tag{31}$$

The PDF of the normally distributed gains and losses is

$$f_{\mathcal{N}_{i,t+\tau}}(u) = \frac{1}{\sigma_{i,t+\tau}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{z-\mu_{i,t+\tau}}{\sigma_{i,t+\tau}}\right)^2}. \tag{32}$$

A.3 Modeling Subjective Mortality

Hoermann and Ruß (2008) propose a gamma distribution for modeling $D \sim \Gamma(\alpha, \beta, \gamma)$ with density function

$$f_{(\alpha,\beta,\gamma)}^{\Gamma}(d) = \frac{1}{\Gamma(\alpha)\,\beta^{\alpha}}(d-\gamma)^{\alpha-1}e^{-\frac{d-\gamma}{\beta}},$$

expected value

$$E(D) = \alpha\beta + \gamma$$

and Variance

$$Var(D) = \alpha\beta^2$$

for $d \geq \gamma$, $\gamma \in \mathbb{R}$ and $\alpha, \beta > 0$.

A.4 Expected Utility Theory Calibration of CPT

Instead of using subjective utility valuation by MCPT, we estimate lifetime utility of person i, who is x years old in t, by expected utility theory where we use a square root utility function

$$EU(i) = \sum_{t}^{T-x} \delta^{\tau} \ EU(Z_{i,t}). \tag{33}$$

This is basically the prospect theory where the parameters of the value function $a = b = 0,5$ yielding a square root utility, and the parameters of the weighting function $\gamma = \kappa = 1$ yielding actual weighting of the states instead of subjective perceptions. Therefore, instantaneous expected utility in t is

$$EU(Z_{i,t}) =_{\tau} p_x \int_{-\infty}^{\infty} v(z)\, f_{i,t}(z)\, dz \tag{34}$$

where

$$v\left(z\right) = \begin{cases} \sqrt{z} & z \geq 0 \\ -\sqrt{|z|} & z < 0. \end{cases} \tag{35}$$

The PDF of available funds for i in t results from equation (32), with its moments coming from equation (25) and equation (17). Finally we numerically maximize equation (33) subject to the optimal level of tontine investment B_i.

$$\max_{B_i} EU\left(i\right)$$
$$\text{s. t. } B_i \leq W_i, B_i \geq 0 \tag{36}$$

B Tables

t	1	10	20	30	40	44
$\mu_{i,t}$	345.15	559.72	1,652.04	5,619.14	14,307.30	30,000
$\sigma_{i,t}^{N=10,000}$	19.05	25.30	63.63	227.70	572.00	1169.77
$\sigma_{i,t}^{N=100,000}$	5.93	8.74	23.00	78.43	187.74	370.75

Table 2: Properties of normally distributed tontine returns for $B_i = 30,000$ in t in EUR for tontine size $N = 100,000$ vs. $N = 10,000$