Annals of Actuarial Science (2020), **14**, pp. 302–315
doi:10.1017/S1748499520000214

PAPER

Optimal portfolio choice with tontines under systematic longevity risk

Irina Gemmo[1*], Ralph Rogalla[2] and Jan-Hendrik Weinert[3]

[1]Chair of Integrative Risk Management and Economics, D-MTEC, ETH Zurich, Zurich, Switzerland; [2]Greenberg School of Risk Management, Insurance and Actuarial Science, Tobin College of Business, St John's University, New York City, NY, USA; [3]Viridium Group, Neu-Isenburg, Germany
*Corresponding author. E-mail: igemmo@ethz.ch

(Received 11 June 2020; accepted 11 June 2020; first published online 13 July 2020)

Abstract

We derive optimal portfolio choice patterns in retirement (ages 66–105) for a constant relative risk aversion utility maximising investor facing risky capital market returns, stochastic mortality risk, and income-reducing health shocks. Beyond the usual stocks and bonds, the individual can invest his assets in tontines. Tontines are cost-efficient financial contracts providing age-increasing, but volatile cash flows, generated through the pooling of mortality without guarantees, which can help to match increasing financing needs at old ages. We find that a tontine invested in the risk-free asset dominates stock investments for older investors without a bequest motive. However, with a bequest motive, it is optimal to replace the tontine investment over time with traditional financial assets. Our results indicate that early in retirement, a tontine is only an attractive investment option, if the tontine funds are invested in a risky asset. In this case, they crowd out stocks and risk-free bonds in the optimal portfolios of younger investors. Over time, the average optimal portfolio weight of tontines decreases. Introducing systematic mortality risks noticeably reduces the peak allocation to tontines.

Keywords: Tontines; Mortality; Systematic longevity risk; Retirement planning

JEL Classification: D14; E21; G22; J10.

1. Introduction

Around the world, many societies experience ongoing demographic change caused by declining birth rates and a simultaneous increase in life spans. According to the World Bank (2015), between 1960 and 2015, worldwide life expectancy at birth has increased from 52.5 to above 71.7 years. These developments put a tremendous strain on many statutory pension systems of the traditional pay-as-you-go variety, as the number of beneficiaries increases, while the number of contributors decreases. Consequently, funded pension products and private retirement provisions are gaining in importance.

At the same time, households face increasing financial needs in retirement. De Nardi *et al.* (2010), among many others, document that medical expenditure quickly increases as individuals age. The same holds for costs associated with long-term care or those related to other services needed to maintain the standard of living. To the extent that these costs are not fully covered by insurance, households ought to ensure that they hold sufficient assets at later ages to finance these expenses, or that their retirement income stream increases over time.

One approach to creating such an age-increasing income stream would be to invest in products that are based on the principle of tontines. Tontines are financial products with survival-contingent payouts originally developed in the 1650s by Lorenzo de Tonti to secure long-term funding for the French state. In their original form, each tontine owner received a lifelong annual pension in exchange for an initial lump sum payment. The shares of deceased tontine members were spread among the survivors, until the last survivor received everyone's pension payments. Contingent on survival, a tontine thus offers an age-increasing payout structure (though without guarantees), relying on the pooling of mortality between policyholders. This characteristic makes the tontine concept appear interesting against the backdrop of increasing financial needs in old age, since already small investment amounts can generate substantial payouts later in life. Accordingly, tontines have recently gained increasing attention in the academic literature and in real life.

McKeever (2009), Milevsky (2015), and Li & Rothschild (2019) review the historical development of tontines. Sabin (2010), Milevsky & Salisbury (2015), and Milevsky & Salisbury (2016) study actuarially fair and optimal payout structures of tontines, and Chen *et al.* (2019) analyse the effects of combining a tontine and an annuity into a unified product. Weinert (2017a) estimates the cost of a tontine compared to traditional life insurance products from an economic as well as a regulatory perspective. Weinert (2017b) analyses the implications of providing tontine members with the option to cancel the contract and derives the fair surrender value. Weinert & Gründl (2016) study the suitability of tontines with their age-increasing payout structure as complements to traditional retirement products against the background of the aforementioned demographic challenges. That paper, however, only analyses the fundamental effects of tontinisation in a simplified framework without capital markets. Outside the academic literature, Hayashi (2020) reports that tontines are experiencing increasing interest among the aging population in Japan.

In this paper, we take a more holistic approach and study the impact of incorporating tontines into a life-cycle framework of consumption and portfolio choice for a retired individual. In doing so, we add to the increasing literature on the role of life-contingent financial products in optimal household portfolios (Horneff *et al.*, 2010; Maurer *et al.*, 2013; Hubener *et al.*, 2014; Horneff *et al.*, 2015). Specifically, we derive optimal portfolio choice patterns for a constant relative risk aversion (CRRA) utility maximising household facing capital market risks, systematic mortality risks, and costs of maintaining a constant standard of living that increase with age in case of frailty. Given the high dimensionality of our optimisation problem, we employ a simulation/regression approach, specifically the one discussed in Denault & Simonato (2017) and Denault *et al.* (2017), which we further describe below.

If the tontine funds are invested risk-free, the tontine has the characteristics of a risky investment with a time-changing risk-return profile and a lower bound guaranteed return of the risk-free rate, which in every period competes against the risk-return profile of a stock investment. We find that a tontine invested in the risk-free asset dominates stock investments for investors in their early 90s and older without a bequest motive. However, with a bequest motive, it is optimal to replace the tontine investment over time by a risk-free investment and stocks. Our results indicate that, at the beginning of retirement, a tontine is only an attractive investment option, if the tontine funds are invested in a risky asset. In this case, they are crowding out stocks and risk-free bonds in the optimal portfolios of investors until the early 80s, before their relevance decreases towards the end of the life-cycle.

The remainder of this paper is structured as follows: In section 2, we describe our life-cycle model for the retirement phase and its components. In section 3, we present the optimal asset allocation patterns that we derive numerically by solving our model for a variety of parameter sets and utility specifications. Finally, section 4 concludes.

2. The Model

2.1 Utility

We model the retirement phase of a representative individual in discrete time from age 66 ($t = 1$) to age 105 ($T = 40$). Our subject's utility drawn from consuming a single, non-durable

https://doi.org/10.1017/S1748499520000214 Published online by Cambridge University Press

consumption good is described by a time-separable CRRA utility function. The individual seeks to maximise lifetime utility, which – in recursive form – is specified as

$$V_t = u(C_t) + \beta E_t \left[p_t V_{t+1} + \left(1 - p_t\right) b \cdot u(D_{t+1}) \right] \tag{1}$$

with

$$u(*) = \begin{cases} log(*) & \text{for } \rho = 1 \\ \frac{(*)^{1-\rho}}{1-\rho} & \text{for } \rho \neq 1 \end{cases}$$

C_t denotes consumption at time t, D_t is the level of bequest at time t, ρ is the coefficient of relative risk aversion, b is the strength of bequest motive, and $\beta > 0$ is the time preference discount factor. p_t is the survival probability from time t to $t+1$, which we assume to be known at time t, but which evolves stochastically over time as described below. In the final period of life at the maximum attainable age T, $p_T = 0$ and therefore, $V_T = u(C_T) + \beta E_T \left[b \cdot u(D_{T+1}) \right]$.

2.2 Capital market

The individual has access to capital markets through investing in risk-free bonds, risky stocks (denoted by $\bullet$), and risky tontines (denoted by $\circ$). The one-period real gross bond return is constant over time and described by the accumulation factor R_f. The stochastic real gross stock return from time $t-1$ to time t is denoted by the accumulation factor $R_t^\bullet$, which is assumed to be serially independent and identically log-normally distributed with an expected value of $E(R_t^\bullet) = (1 + \mu^\bullet)$ and a volatility (i.e. standard deviation) of $\sigma^\bullet$.

In addition to bonds and stocks, the individual can – at each point in time – invest into revolving 1-year tontines, which allow members to decide every year anew how much money they want to commit over the next period. This contrasts with other tontine designs where funds are committed for life. The pool of tontinists, for which we assume that our investor is a representative member, is closed and exists until there is only one survivor remaining. Pool members are assumed to be homogeneous in terms of their personal and economic characteristics, that is, they have the same age, sex, preferences, income, and wealth profiles. Assuming that our investor is representative for all members of the tontine pool has two implications. First, if our representative individual decides to commit funds to the tontine in a given period, all members of the tontine pool will do so. Moreover, in every period, each member of the tontine pool commits an equal amount of funds to the tontine.

The initial 1-year tontine is set up at retirement in $t = 1$ and consists of N_1 members. Over the first year, the funds invested into the tontine accumulate with the tontine's annual gross investment return R_2. At the end of the first year, the accumulated tontine funds are paid out in full, evenly distributed amongst the N_2 surviving participants, and the 1-year tontine renews for those N_2 survivors. This process continues as long as $N_t \geq 2$. In case $N_t = 1$, there is no further opportunity to pool mortality risks, which is a defining characteristic of the tontine. Instead, the tontine collapses into a purely financial product with returns equal to those of the underlying asset. In that case, we consider the tontine to have ceased to exist.

Formally, the tontine's total gross return from time $t-1$ to time t is

$$R_t^\circ = \begin{cases} \frac{N_{t-1}}{N_t} R_t & \text{if alive in } t \\ 0 & \text{if dead in } t \end{cases} \tag{2}$$

https://doi.org/10.1017/S1748499520000214 Published online by Cambridge University Press

where $R_t \in \{R_f, R_t^\bullet\}$[1]. As such, the tontine return is risky by nature, since it depends on the random number of surviving tontinists in the pool. If the tontine funds are invested in bonds, the tontinists' mortality risk will be the only source of risk in the tontine return. If the underlying investment is risky, the tontine return is subject to mortality risk as well as to the underlying market risk.

For CRRA investors, Milevsky & Salisbury (2015) show that the optimal/close to optimal (depending on the level of risk aversion) tontine design is one that provides individual tontinists with a payout structure that resembles that of a constant annuity. This can be achieved through a *natural tontine*, which they define as a tontine with total payouts that decrease proportionally to the survival rates, such that the decrease in the number of participants does not result in an exponential increase in individual payouts.

If, however, the tontine fractions in the investors' portfolios could be adjusted flexibly over time, it might not be necessary to adjust the tontine payout scheme itself. Instead, investors could hold shares in a more traditional tontine design, such as the one described in equation (2) and studied here, and adapt their portfolios in order to smooth consumption, for example, by reducing the tontine fraction over time. This could provide investors with more flexibility, for example, in case of increasing liquidity needs in old age due to health expenditures.

2.3 Retirement income and health state

At the beginning of every period, the household receives an exogenous, health-state-dependent retirement income Y_t. This income is assumed to be constant ($Y_t = \bar{Y}$) as long as the individual remains healthy. In contrast, if the individual becomes frail, income is assumed to decrease over time (and to potentially become negative) to reflect an age-increasing liquidity need Y_t^- associated with frailty. The individual's health state is described by a two-state Markov chain, for which we assume that frailty is an absorbing state. With π_f representing the probability of a healthy individual becoming frail, the transition matrix from state 1 (healthy) to state 2 (frail) is given by $P = \begin{pmatrix} 1-\pi_f & \pi_f \\ 0 & 1 \end{pmatrix}$. Therefore, the health-state-dependent retirement income is $Y_t = \bar{Y} - \mathbb{1}_P \cdot Y_t^-$, where $\mathbb{1}_P$ is one if the household is frail and zero otherwise.

Naturally, the most obvious approach to considering frailty-related health expenses would be to directly model an increased consumption need. The reason we refrain from doing so and rather model decreasing retirement income is to reduce computational effort in determining the optimal policies. The net result of these two approaches is the same. One could argue that – since securing one's health and standard of living in case of frailty is of fundamental importance – health expenses in case of frailty are prioritised over the consumption of other goods and services. Hence, health expenses are immediately deducted from retirement income and decisions with respect to consumption and portfolio choice are only made subsequently based on the remaining budget.

2.4 Wealth accumulation

At the beginning of every period, the individual can spread the wealth on hand W_t across bonds B_t, stocks S_t, tontines Υ_t, and consumption C_t. The budget constraint is

$$W_t = B_t + S_t + \Upsilon_t + C_t \tag{3}$$

Individual disposable wealth on hand in $t+1$ is

$$W_{t+1} = B_t R_f + S_t R_{t+1}^\bullet + \Upsilon_t R_{t+1}^\circ + Y_{t+1} \tag{4}$$

[1] While in this study we assume that tontine funds are either fully invested into a risk-free bond or into stocks, the tontine design studied here allows for flexibility in the asset allocation. In particular, the choice of asset maturities does not depend on the projected maximum lifetime of the tontine pool, as the investment horizon is always one period, due to the revolving nature of the product.

$B_tR_f + S_tR^\bullet_{t+1} + \Upsilon_tR^\circ_{t+1}$ describes the value of financial wealth in $t+1$ and Y_{t+1} is the retirement income. Short selling is not allowed, thus

$$B_t, S_t, \Upsilon_t \geq 0 \tag{5}$$

Since, in case of death, the tontine investment goes to the tontine pool and not to the heirs, the bequest in $t+1$ is

$$D_{t+1} = B_tR_f + S_tR^\bullet_{t+1} \tag{6}$$

2.5 Mortality dynamics

To incorporate systematic mortality risks, we rely on the two-factor approach of Cairns *et al.* (2006). Specifically, we model the stochastic dynamics of the logits of the time-varying conditional 1-year mortality rates (q^t_x) for an individual aged x in year t according to

$$\text{logit}\left(q^t_x\right) = \log\left(\frac{q^t_x}{1-q^t_x}\right) = \kappa^t_1 + \kappa^t_2\,(x-\bar{x}) \tag{7}$$

where $\bar{x} = (x_u - x_l + 1)^{-1}\sum_{x=x_l}^{x_u} x$ is the mean of the range of ages considered to be fitted, κ^t_1 is the "level" of mortality, which usually has a downward trend reflecting generally improving mortality rates over time, and κ^t_2 is the "slope" coefficient, which has a gradual upward drift reflecting the fact that, historically, mortality at high ages has improved at a slower rate than at younger ages. The column vector $\kappa^t = (\kappa^t_1, \kappa^t_2)'$ is assumed to follow a two-dimensional random walk with drift $\kappa^{t+1} = \kappa^t + \tau + \chi \cdot Z_{t+1}$, where τ is the drift of κ^t, χ is the lower triangular Cholesky matrix of the κ^t-covariance matrix Σ, and Z_{t+1} is a bivariate standard normal shock. Furthermore, we assume that the mortality rate at the end of our projection horizon, that is at age 105, is $q^T_{105} = 1$.

2.6 Calibration

For the base case calibration of the model, we set the risk aversion parameter to $\rho = 1$ (Aït-Sahalia & Hurd, 2016). Following Weinert & Gründl (2016), we set the subjective discount factor to $\beta = 0.98$. We refrain from an accumulation phase and begin the analysis at retirement age 66 ($t = 1$) with an initial wealth endowment of $W_1 = 150{,}000$ EUR. Using Ordinary Least Squares (OLS) regression, we calibrate the described Cairns *et al.* (2006) mortality model to mortality data for US males aged 65–105 over the period 1950–2014 provided by the Human Mortality Database (HMD). We obtain the following point estimates:

$$\hat{\kappa}^0 = \begin{pmatrix}-2.4106763\\ 0.1050414\end{pmatrix}, \hat{\tau} = \begin{pmatrix}-0.0090878\\ 0.0003335\end{pmatrix}, \text{and } \hat{\Sigma} = \begin{pmatrix}0.0005446 & 0.0000162\\ 0.0000162 & 0.0000008\end{pmatrix}$$

We assume a maximum attainable age of 105 ($T = 40$). The individual receives a constant yearly retirement income of $\bar{Y} = 24{,}000$ EUR. However, with probability $\pi_f = 0.03$, the individual becomes frail and faces an increased liquidity need, which depends on the age of the individual. In this case, the liquidity need is 12,000 EUR at age 66 ($t = 1$) and linearly increases up to 30,000 EUR at age of 105 ($T = 40$). That is, $Y^-_t = \frac{15}{26}t + 12{,}000$. Once the individual faces a higher liquidity need, this state will remain until death. Therefore, the health-state-dependent retirement income $Y_t = \bar{Y} - \mathbb{1}_P \cdot Y^-_t = 24{,}000 - \mathbb{1}_P \cdot (\frac{15}{26}t + 12{,}000)$. The frailty probability, the resulting age-dependent liquidity needs, as well as the initial wealth endowment are estimated based on the German Socio-Economic Panel (SOEP) in the period 1984–2013. The risk-free gross bond return is $R_f = 1.02$, the expected stock return is $\mu^\bullet = 0.06$, and the volatility of stock return is $\sigma^\bullet = 0.18$ as in Horneff *et al.* (2010). In the base case scenario, we assume that the tontine funds are invested into the risk-free asset, that is, $R_t = R_f = 1.02$. The initial size of the tontine is $N_1 = 10{,}000$. Furthermore, we set the bequest parameter to $b = 1$. We assume no correlation between stock

https://doi.org/10.1017/S1748499520000214 Published online by Cambridge University Press

Table 1. Model calibration

Variable	Description	Value	Source
$\bar{\omega}$	Entry age	66	Horneff *et al.* (2010)
W_1	Initial wealth endowment	150, 000 €	Estimation based on SOEP (1984 to 2013)
$\bar{Y}$	Retirement income	24, 000 €	Estimation based on SOEP (1984 to 2013)
π_f	Frailty probability	0.03	Estimation based on SOEP (1984 to 2013)
β	Subjective discount factor	0.98	Weinert & Gründl (2016)
b	Bequest motive	0; 1	Cocco *et al.* (2005)
ρ	Degree of risk aversion	1	Aït-Sahalia & Hurd (2016)
N_1	Initial tontine size	100; 1, 000; 10, 000	
R_f	Risk-free gross bond return	1.02	Horneff *et al.* (2010)
$\mu^{\bullet}$	Expected stock return	0.06	Horneff *et al.* (2010)
$\sigma^{\bullet}$	Volatility of stock return	0.18	Horneff *et al.* (2010)
$\hat{\kappa}^0$	Initial mortality model parameters	(−2.4106763; 0.1050414)	Calibrated based on HMD for US males aged 65–105 (1950–2014)
$\hat{\tau}$	Drift of κ^t	(−0.0090878; 0.0003335)	Calibrated based on HMD for US males aged 65–105 (1950–2014)
$\hat{\Sigma}$	κ^t-covariance matrix	$\begin{pmatrix} 0.0005446 & 0.0000162 \\ 0.0000162 & 0.0000008 \end{pmatrix}$	Calibrated based on HMD data for US males aged 65–105 over 1950–2014

returns and income shocks. We use 10, 000 simulated life-cycle trajectories of stock and tontine returns, health state, mortality model parameters, and tontine pool size to determine expected lifetime utilities and optimal policies (see the description of the numerical strategy below).

In subsequent variations of the base case, we consider a scenario in which the investor does not have a bequest motive ($b = 0$) and one in which tontine funds are invested in stocks (i.e. $R_t = R_t^{\bullet}$). Further, we analyse the impact of alternative mortality dynamics by disabling the random mortality shocks (*trending mortality*; $\chi = 0$) and then by additionally disabling the mortality drift (*constant mortality*; $\chi, \tau = 0$). Finally, we study the impact of alternative initial tontine pool sizes ($N_1 = 100$ and $N_1 = 1, 000$). Table 1 summarises the parameters for our model calibrations.

2.7 Numerical strategy

We seek to maximise equation (1) with respect to consumption C_t subject to equations (3)–(6). Traditionally, the workhorse of solving such discrete-time portfolio choice models is dynamic optimisation via backward induction over a discretised, potentially high-dimensional, state space comprising all state variables. While this approach is straightforward, it suffers from the curse of dimensionality, as the computational effort to find optimal solutions increases exponentially with the number of state variables. To ease the computational burden related to a greater number of state variables, alternative solution strategies have been proposed more recently. Among those are various simulation/regression approaches (Brandt *et al.*, 2005; Koijen *et al.*, 2007, 2010). Given the high dimensionality of our optimisation problem with six state variables (age, wealth, health, remaining tontine population, and two mortality factors), we also employ a simulation/regression approach, specifically the one discussed in Denault & Simonato (2017) and Denault *et al.* (2017). In what follows, we briefly discuss the main aspects of this solution technique. For a detailed step-by-step description of the algorithm, including examples, we refer to Denault *et al.* (2017, p. 380 ff).

Generally, the simulation/regression approach employs standard dynamic optimisation techniques. Key element in reducing the computational effort associated with solving the model, however, is the differential treatment of endogenous and exogenous state variables. Endogenous state variables, in our case wealth, are discretised following the traditional approach. The interrelation between the value function and the exogenous state variables, on the other hand, is captured by specifying future utility as a parametric function of policy and exogenous state variables. To this end, I (in our case $10,000$) life-cycle trajectories of our exogenous state variables are simulated using the stochastic dynamics described above, and a sample set of M policies (in our case all asset portfolios with weight steps of 10%, excluding those that contain any asset weights of zero) is defined. Then, for each (discretised) value of the endogenous variable at time t, the utility at time $t+1$ for each of the $I \cdot M$ combinations of sample policies and exogenous state variable realisations is determined. These future utilities are then regressed onto a basis of policy and exogenous state variables (in our case, a constant, the 3 asset weights, their squares and cross-products, the 4 exogenous state variables, and the 12 cross-products between asset weights and exogenous state variables) to derive the parametric function. Using this functional relation between policies, (exogenous) state variables, and future utility, we then determine for each state variable trajectory the utility maximising asset allocation[2].

3. Results

3.1 Tontine characteristics

A central characteristic of the tontine is the age-increasing return it generates through the mortality credit, that is, the redistribution of invested capital of deceased tontine members to the survivors in the pool. As per our assumption, the tontine does not take on new members beyond those who participated initially, and hence, the pool of participants decreases over time with members deceasing, as shown in Figure 1.

Given this decrease in the size of the pool, surviving members hold increasingly larger shares of the tontine, and each additional death results in a larger relative payout to subsequent survivors. This is exemplified by Table 2, which presents expected values and standard deviations of the annual mortality-related tontine return (conditional on survival of our investor, but not conditional on the existence of the tontine), that is $\frac{R_t^\circ}{R_f} - 1$, at selected ages for different initial tontine sizes and mortality regimes. For our baseline scenario with an initial pool size of $N_1 = 10,000$ and under the stochastic mortality scenario, the expected mortality return at age 70, that is, the return over the period from age 69 to age 70, is 1.6% and monotonously increases to 60.82% at age 105. This increase in return is accompanied by a corresponding increase in the volatility of the mortality credit from 0.14% at age 70 to 21.08% at age 105, because the variability of mortality realisations increases with decreasing pool size. Comparing these results to those under the constant mortality

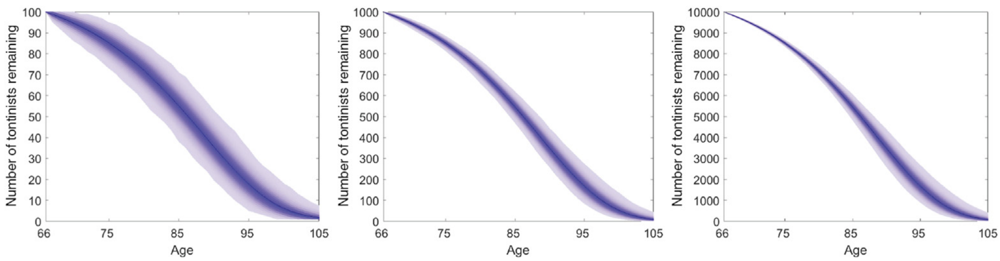

Figure 1. Simulated development of the number of remaining tontine members over time. Initial number of tontine members at age 66: $N_1 = 100$, $N_1 = 1,000$, and $N_1 = 10,000$, respectively. The graphs show the 0.1–99.9 percentiles simulated paths. Darker areas represent higher probability mass.

[2] Note that determining the optimal asset allocation via the parametric function enforces non-anticipativity.

https://doi.org/10.1017/S1748499520000214 Published online by Cambridge University Press

Table 2. Expectation and standard deviation of mortality credit (conditional on survival) at selected ages depending on initial tontine size N_1 and mortality dynamics

	Initial tontine size at the age of 66	Mean/standard deviation of mortality credit (in %) at age 70	80	90	100	101	102	103	104	105
					Stochastic mortality					
Expectation	$N_1 = 100$	1.61	4.16	11.53	33.59	36.95	37.53	36.85	32.60	28.61
	$N_1 = 1,000$	1.61	4.15	11.49	34.09	38.70	43.05	47.88	54.04	59.42
	$N_1 = 10,000$	1.60	4.15	11.51	34.39	38.54	43.12	48.38	54.19	60.82
St. dev.	$N_1 = 100$	1.32	2.50	6.52	37.09	45.31	48.46	52.60	50.35	50.66
	$N_1 = 1,000$	0.42	0.82	2.39	12.78	16.04	21.66	27.65	37.33	50.26
	$N_1 = 10,000$	0.14	0.38	1.57	7.58	8.98	10.93	13.25	16.40	21.08
					Trending mortality					
Expectation	$N_1 = 100$	1.61	4.16	11.39	33.54	37.45	38.82	39.88	35.30	29.91
	$N_1 = 1,000$	1.61	4.14	11.41	33.52	38.00	42.13	47.26	53.24	59.18
	$N_1 = 10,000$	1.60	4.14	11.42	33.76	37.81	42.18	47.27	53.05	59.58
St. dev.	$N_1 = 100$	1.32	2.47	6.14	35.40	43.57	49.73	54.79	53.43	51.85
	$N_1 = 1,000$	0.42	0.78	1.92	9.63	12.09	15.77	21.24	30.46	41.03
	$N_1 = 10,000$	0.13	0.25	0.60	2.99	3.74	4.89	6.40	8.59	11.69
					Constant mortality					
Expectation	$N_1 = 100$	1.69	4.82	13.67	36.00	37.70	37.19	32.95	27.11	20.62
	$N_1 = 1,000$	1.68	4.78	13.64	39.19	43.24	48.00	54.01	59.76	65.25
	$N_1 = 10,000$	1.67	4.78	13.67	39.04	43.36	48.13	53.52	59.27	65.65
St. dev.	$N_1 = 100$	1.34	2.72	7.41	43.07	50.49	52.62	52.10	48.76	43.41
	$N_1 = 1,000$	0.42	0.84	2.30	13.46	17.47	23.02	32.00	46.49	64.54
	$N_1 = 10,000$	0.13	0.27	0.73	4.19	5.32	7.06	9.33	12.62	17.37

scenario, we observe that under constant mortality the expected mortality credit is higher (1.67% at age 70; 65.65% at age 105) and the volatility is lower than in our base case (0.13% at age 70; 17.37% at age 105). This is due to the fact that under the stochastic mortality scenario, mortality rates exhibit a downward trend, which reduces the mortality credit in expectation, while they also randomly fluctuate around this trend, which induces higher return variability.

In case the initial number of tontinists is only $N_1 = 100$, the increase in the volatility of the mortality credit over time is even more pronounced than in the base case, due to the reasons discussed above. Under stochastic (constant) mortality, volatility now increases from 1.32% (1.34%) at age 70 to 50.66% (43.41%) at age 105. By contrast, the return expectation now exhibits a hump-shaped profile over time. Under stochastic mortality, expected returns peak at 37.53% at age 102, before they drop to 28.61% at age 105. Under constant mortality, the return peak of 37.7% is realised at age 101, before expected returns drop to 20.62% at age 105. Given the small initial pool size, a tontine investor faces a considerable risk that N_t will decrease to 1 and, hence, that the tontine ceases to exist prior to the end of his planning horizon, as discussed in Section 2.2. In these cases, the mortality credit is zero, which negatively affects expected return outcomes.

To shed more light onto this aspect, Table 3 presents simulated probabilities of only a single tontinist surviving at advanced ages for alternative initial tontine pool sizes and mortality dynamics. For our baseline scenario of $N_1 = 10,000$ initial tontinists, none of our 10,000 simulation runs

https://doi.org/10.1017/S1748499520000214 Published online by Cambridge University Press

Table 3. Probability at selected ages of tontine ceasing to exist depending on initial tontine size N_1 and mortality dynamics

Initial tontine size at the age of 66	$Pr\ (N_t = 1)$ (in %) at age 96	97	98	99	100	101	102	103	104	105
Stochastic mortality										
$N_1 = 100$	0.00	0.00	0.08	0.40	1.61	5.18	11.75	21.74	34.79	49.27
$N_1 = 1,000$	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.38	2.16
$N_1 = 10,000$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trending mortality										
$N_1 = 100$	0.00	0.00	0.03	0.16	0.63	2.77	7.72	17.85	32.79	49.76
$N_1 = 1,000$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.06
$N_1 = 10,000$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Constant mortality										
$N_1 = 100$	0.00	0.00	0.20	1.06	3.83	10.73	22.01	37.76	54.06	69.45
$N_1 = 1,000$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.28	3.00
$N_1 = 10,000$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

resulted in the tontine ceasing to exist before the end of the projection horizon, independent of the mortality dynamic. For a small initial tontine size of $N_1 = 100$, on the other hand, we observe tontine closures as early as age 98 under all mortality scenarios. At age 100, 1.61% of our simulation runs under the stochastic mortality scenario result in the number of tontinists dropping to one, while this is the case in 3.83% of our sample paths under the constant mortality scenario. At our final age of 105, the tontine has ceased to exist in almost 50% of the simulations under non-constant mortality, while under constant mortality this rate stands at almost 70%[3].

3.2 Optimal asset allocation profiles in retirement

First, we consider our base case calibration, which is characterised by a large tontine ($N_1 = 10,000$), tontine funds invested into the risk-free asset ($R_t = R_f = 1.02$), and by stochastic mortality dynamics. Furthermore, the investor has a bequest motive ($b = 1$) and a risk aversion parameter for consumption and bequest of $\rho = 1$. At the age of 66, he is endowed with initial wealth $W_1 = 150,000$ EUR.

Figure 2 shows the 5-year moving average of optimal portfolio composition over time, that is, the composition of savings. At early retirement ages, the investor optimally allocates 100% of his investment into stocks. Due to the relatively high survival probability at the beginning of retirement, the expected mortality credit of the tontine is relatively low compared to the expected stock return. Consequently, given the investor's low risk aversion, stocks dominate the other two assets. With higher age, the expected mortality credit and, consequently, the expected tontine return increase. Starting in the mid-80s, the investor optimally substitutes part of the stocks with an investment in the tontine and in bonds. The average allocation into tontines peaks in the investor's late 80s at around 25%. Beyond that age, the bequest motive becomes more relevant, and the investor then starts to reduce the share invested in the tontine. Since the investment in the tontine goes to the other tontine members in the event of death, and not to the heirs, and the probability of dying increases with increasing age, the tontine investment is reduced further and

[3] Using the probabilities listed in Table 3 and the unconditional expected mortality credits in Table 2, we can determine the expected mortality credits conditional on the tontine not having ceased to exist. For example, in case $N_1 = 100$ and under stochastic mortality, the simulated expected mortality credit earned at age 105, conditional on the tontine still being active at age 104, is $\frac{0.2861}{(1-0.4927)} = 0.5640$. Hence, conditional on the tontine still being active, the expected mortality profile is not hump-shaped but increasing in age.

https://doi.org/10.1017/S1748499520000214 Published online by Cambridge University Press

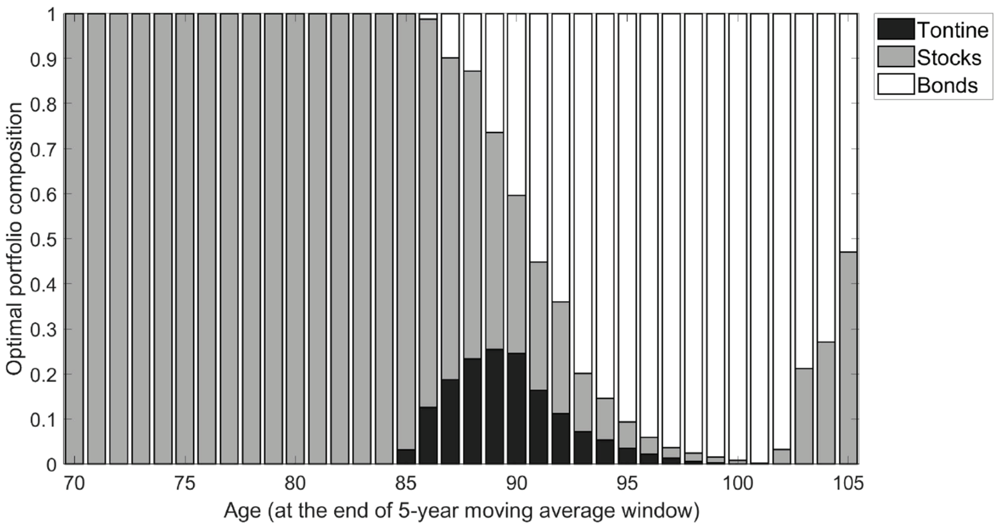

Figure 2. Optimal portfolio composition: base case. Five-year moving average of optimal portfolio composition. Large initial tontine size $N_1 = 10,000$; tontine funds invested into the risk-free asset ($R_t = R_f$); bequest motive $b = 1$; CRRA risk aversion parameter $\rho = 1$; initial wealth $W_1 = 150,000$ EUR; stochastic mortality scenario. The x-axis shows the age at the end of each overlapping 5-year moving average window.

further. At this stage, the risk-free bonds are slowly crowding out the risky investments. By the age of 100, all financial wealth is optimally invested into risk-free bonds. Towards the end of the life-cycle, with the average retirement income decreasing due to deteriorating health, the investor has to increasingly draw down his financial wealth to finance consumption. At the same time, retaining sufficient wealth to satisfy the bequest motive becomes more and more relevant. So, in order to maintain a stable consumption profile while still providing for bequest, the investor again increases his stocks holdings in order to cash in on the equity risk premium.

Figure 3 shows a variation of the base case without bequest motive ($b = 0$). As in the base case, it is optimal to hold 100% of stocks until the mid-80s, when stocks are partially substituted by tontines and risk-free bonds. In contrast to the more modest role tontines play in base case portfolios, the average share invested into tontines increases to about 72% by the investor's late 80s. In the absence of a bequest motive, the two risky alternatives, stocks and tontines, are only evaluated by their return distributions for when the investor is alive. With an increasing expected mortality credit, the tontine outperforms the stocks and crowds them out entirely. Since the tontine dominates the stocks at older ages, the investor's portfolio only consists of tontine shares and bonds. Despite the investor's relatively low risk aversion and lack of interest in leaving a bequest, the tontine share decreases towards the end of the life-cycle. This is due to the high tontine return volatility and the short remaining time horizon, which in combination makes it more difficult to smooth consumption over time. In the absence of a bequest motive, we do not see an increase in stock investments towards the end of the life-cycle.

Figure 4 describes the same situation as in the base case, with the difference that the tontine funds are now invested into the risky asset rather than the risk-free asset. That is, the tontine return is composed of the mortality credit and the return of the risky underlying. As such, the tontine dominates both the riskless and the risky asset resulting in a portfolio of 100% tontine from the beginning of retirement until the age of mid-80s. With increasing age, the expected mortality credit increases, but with the decreasing survival probability, the bequest motive becomes more relevant. Hence, the tontine becomes less attractive with age, since tontine holdings are not bequeathable. From his mid-80s, an investor with bequest motive of $b = 1$ adds stocks and bonds to his financial portfolio. The optimal relative stake in tontines decreases, which allows the

https://doi.org/10.1017/S1748499520000214 Published online by Cambridge University Press

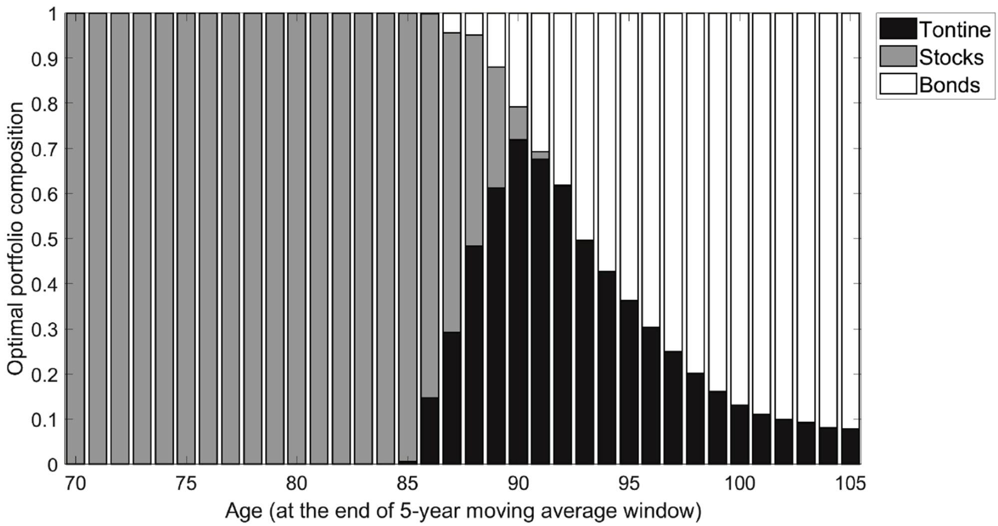

Figure 3. Optimal portfolio composition: no bequest. Five-year moving average of optimal portfolio composition. Large initial tontine size $N_1 = 10,000$; tontine funds invested into the risk-free asset ($R_t = R_f$); no bequest motive $b = 0$; CRRA risk aversion parameter $\rho = 1$; initial wealth $W_1 = 150,000$ EUR; stochastic mortality scenario. The x-axis shows the age at the end of each overlapping 5-year moving average window.

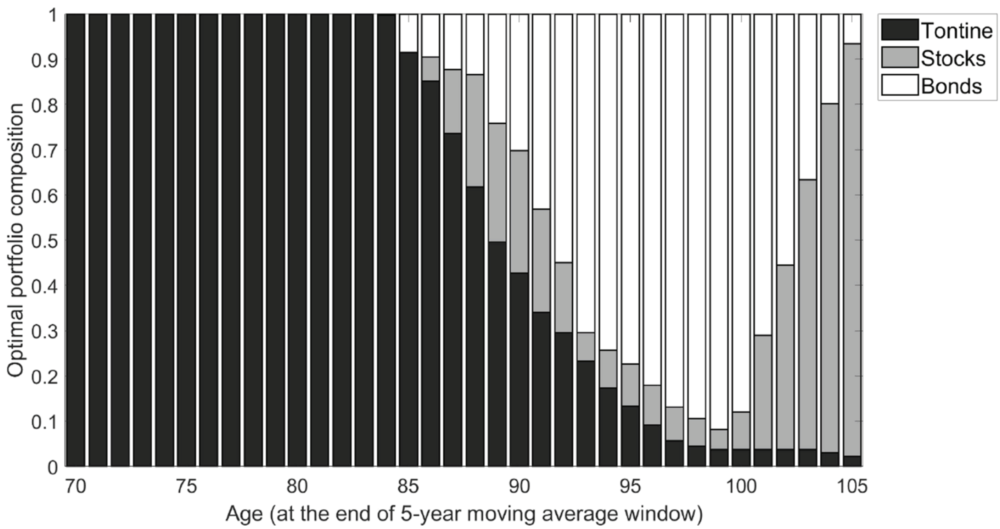

Figure 4. Optimal portfolio composition: tontine invested in stocks. Five-year moving average of optimal portfolio composition. Large initial tontine size $N_1 = 10,000$; tontine funds invested into the risky asset ($R_t = R_t^\bullet$); bequest motive $b = 1$; CRRA risk aversion parameter $\rho = 1$; initial wealth $W_1 = 150,000$ EUR; stochastic mortality scenario. The x-axis shows the age at the end of each overlapping 5-year moving average window.

investor to better smooth consumption. Similar to the base case, the share of risk-free bonds in the portfolio is increasing from the mid-80s to the late 90s and towards the end of the life-cycle, the investor increases his share in risky assets again. Interestingly, this increase in risky assets is much stronger than in the base case. With 15 years worth of additional mortality credits earned early in retirement, compared to the base case, the investor has, on average, generated a substantially higher amount of wealth, such that potentially even more adverse capital market developments would not jeopardise his ability to finance desired consumption and bequest. Therefore, it is not

https://doi.org/10.1017/S1748499520000214 Published online by Cambridge University Press

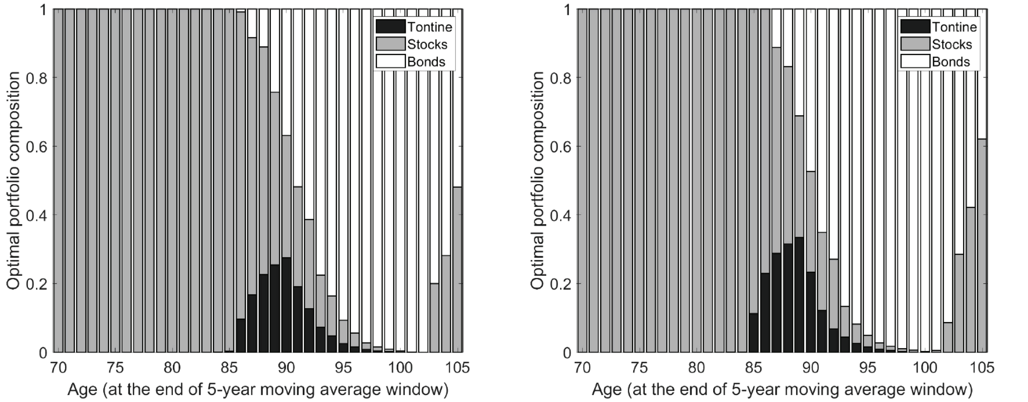

Figure 5. Optimal portfolio composition: alternative mortality dynamics. Five-year moving average of optimal portfolio composition. Left panel: trending mortality ($\chi = 0$), right panel: constant mortality ($\chi, \tau = 0$). Large initial tontine size $N_1 = 10,000$; tontine funds invested into the risk-free asset ($R_t = R_f$); bequest motive $b = 1$; CRRA risk aversion parameter $\rho = 1$; initial wealth $W_1 = 150,000$ EUR; stochastic mortality scenario. The x-axis shows the age at the end of each overlapping 5-year moving average window.

Table 4. Change in tontine investment for alternative tontine sizes. Differences (in percentage points) in 5-year moving averages of optimal tontine portfolio weight between alternative scenarios with lower tontine size and the base case calibration. Δ for $N_1 = 1,000$: difference between tontine investments for an initial tontine size of $N_1 = 1,000$ versus the base case; Δ for $N_1 = 100$: difference between tontine investments for an initial tontine size of $N_1 = 100$ versus the base case. Base case calibration: $N_1 = 10,000$, $\rho = 1$

Age at the end of 5-year moving average period	70	75	80	85	90	95	100	105
Δ for $N_1 = 1,000$	0.00	0.01	−24.78	7.58	1.41	0.06	0.00	0.00
Δ for $N_1 = 100$	−20.00	0.00	−40.00	−21.74	16.81	0.21	0.02	0.00

necessary for him to hedge his position by holding a substantial position in riskless bonds. This is supported by the fact that – despite the presence of a bequest motive – the investor continues to hold a small fraction of his wealth in non-bequeathable tontines to potentially benefit from the high mortality credits than late in the life-cycle.

Figure 5 illustrates the scenarios in which mortality dynamics are altered. The panel on the left depicts the 5-year moving average of the optimal tontine share in the investment portfolio by age in case of trending mortality, that is, $\chi = 0$. Removing stochasticity from the mortality model but retaining the generally downward-sloping mortality trend only has a minor impact on the optimal portfolio allocation. This is in contrast to the outcome if both random mortality shocks and deterministic mortality trends are removed. This case of constant mortality, that is, $\chi, \tau = 0$ is presented in the panel on the right. Here, the overall asset allocation pattern is similar to the base case. The optimal share invested in the tontine in the second half of the 80s, however, peaks at about 33%, about 8 percentage points (or about one third) higher than in the base case scenario. This can be attributed to the fact that under constant mortality the tontine has a better risk-return trade-off as discussed in Table 2.

Table 4 describes the change in the 5-year moving averages of the optimal portfolio share invested in tontines when moving from the base case to a scenario with tontines that have a smaller number of initial members. In case $N_1 = 1,000$ (row "Δ for $N_1 = 1,000$" in Table 4), allocations to big- and medium-sized tontines do not differ significantly early and late in retirement. Notable differences can only be observed in the late 70s and early 80s, when investors with access to big tontines first invest on average about 25 percentage points more in a tontine than investors with access to medium-sized tontines. By the mid-80s, the relation has flipped, and investors with

https://doi.org/10.1017/S1748499520000214 Published online by Cambridge University Press

access to the medium size tontine hold a higher fraction of their wealth in tontines relative to those with access to big tontines, about 7.6 percentage points, on average. In case the initial tontine size is only $N_1 = 100$ (row "Δ for $N_1 = 100$" in Table 4), we observe a similar general pattern in that allocations to big tontines exceed those to small tontines earlier in the life-cycle while they fall short later, particularly in late 80s; however, the differences between that calibration and the base case are more pronounced.

4. Conclusion

We derive optimal portfolio choice patterns in retirement (ages 66 to 105) for a CRRA utility maximising investor facing risky capital market returns, stochastic mortality risk, and income-reducing health shocks. Beyond the usual stocks and bonds, the individual can invest his assets in tontines. Tontines are cost-efficient financial contracts providing age-increasing, but volatile cash flows, generated through the pooling of mortality without guarantees, which can help to match increasing financing needs at old ages. We construct a 1-year revolving tontine to allow for a flexible payout structure that is determined by the amount of tontine investment.

For our base case scenario of a log-investor with bequest motive, in which the tontine funds are invested in a risk-free asset, we find an optimal 5-year moving average portfolio share invested in a tontine of up to 25%. In this case, an investment in the tontine seems most attractive for an investor in his 80s and 90s, when the mortality risk in the tontine pool is high enough to generate an attractive tontine return but the investor's survival probability is also still high enough for a low relevance of the bequest motive. At earlier ages, the mortality risk in the tontine pool, and hence the mortality credit, is relatively low. Therefore, the stock investment dominates the tontine. At later ages, towards the end of the lifetime, the tontine share in the optimal portfolio is crowded out by other assets, since tontine investments are not bequeathable.

In the absence of a bequest motive, the average optimal fraction invested into tontines rises up to about 72%. With an increasing expected mortality credit, the tontine crowds out the risky stock investment due to tontine's superior risk-return profile. Since the tontine dominates the stocks at older ages, the investor's portfolio only consists of tontine shares and bonds until the end of the lifetime. That is, even without a bequest motive, investors will not hold all their assets in tontines late in life to help smooth consumption over time.

If the tontine funds are invested in risky stocks rather than risk-less bonds, the tontine return is composed of the mortality credit and the return of the risky underlying. In this scenario, we find that the tontine dominates both, the riskless and the risky asset, resulting in a portfolio of 100% tontine from the beginning of retirement until the age of mid-80s. In the presence of a bequest motive, an investor starts adding stocks and bonds to his financial portfolio starting in his mid-80s. Despite the presence of a bequest motive, tontines are not crowded out entirely until the very last year of the life-cycle. The optimal relative stake in tontines is decreasing, which allows the investor to smooth consumption due to an increasing expected mortality credit.

Compared to a "constant mortality" scenario, in which mortality tables do not change over time, introducing systematic longevity risks reduces the peak of the average tontine share from about 33% to about 25%, as mortality credit expectations decrease, while mortality credit volatility increases.

Acknowledgements. The authors are grateful for helpful advice and suggestions by David Blake, Gal Wettstein, Peter Hieber, one anonymous referee, and participants at meetings of the American Economic Association (2020), the Chair of Integrative Risk Management and Economics at ETH Zurich (2020), the Department of Actuarial Science at HEC Lausanne (2019), the American Risk and Insurance Association (2019), the German Insurance Science Association (2019), and the European Group of Risk and Insurance Economists (2018), as well as the 14th International Longevity Risk and Capital Markets Solutions Conference (2018). Irina Gemmo and Jan-Hendrik Weinert gratefully acknowledge financial support from the International Center for Insurance Regulation (ICIR) at Goethe-University Frankfurt and thank the W. R. Berkley Corporation for supporting their work with Berkley Fellowships at the Greenberg School of Risk Management, Insurance, and

Actuarial Science at St. John's University. The findings, views, and interpretations expressed herein are those of the authors and should not be attributed to Viridium Insurance Group.

References

Aït-Sahalia, Y. & **Hurd, T.R.** (2016). Portfolio choice in markets with contagion. *Journal of Financial Econometrics*, **14**(1), 1–28.

Brandt, M.W., Goyal, A., Santa-Clara, P. & **Stroud, J.R.** (2005). A simulation approach to dynamic portfolio choice with an application to learning about return predictability. *The Review of Financial Studies*, **18**(3), 831–873.

Cairns, A.J., Blake, D. & **Dowd, K.** (2006). A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance*, **73**(4), 687–718.

Chen, A., Hieber, P. & **Klein, J.K.** (2019). Tonuity: a novel individual-oriented retirement plan. *ASTIN Bulletin: The Journal of the IAA*, **49**(1), 5–30.

Cocco, J.F., Gomes, F.J. & **Maenhout, P.J.** (2005). Consumption and portfolio choice over the life cycle. *Review of Financial Studies*, **18**(2), 491–533.

De Nardi, M., French, E. & **Jones, J.B.** (2010). Why do the elderly save? The role of medical expenses. *Journal of Political Economy*, **118**(1), 39–75.

Denault, M., Delage, E. & **Simonato, J.-G.** (2017). Dynamic portfolio choice: a simulation-and-regression approach. *Optimization and Engineering*, **18**(2), 369–406.

Denault, M. & **Simonato, J.-G.** (2017). Dynamic portfolio choices by simulation-and-regression: Revisiting the issue of value function vs portfolio weight recursions. *Computers & Operations Research*, **79**, 174–189.

Hayashi, Y. (2020). Japan's aging population breathes new life into a centuries-old investment idea. *Wall Street Journal*, 25-Feb-2020. Available online at the address https://www.wsj.com/articles/japans-aging-population-breathes-new-life-into-a-centuries-old-investment-idea-11582628400 [accessed 09-June-2020].

Horneff, V., Maurer, R., Mitchell, O.S. & **Rogalla, R.** (2015). Optimal life cycle portfolio choice with variable annuities offering liquidity and investment downside protection. *Insurance: Mathematics and Economics*, **63**, 91–107.

Horneff, W., Maurer, R. & **Rogalla, R.** (2010). Dynamic portfolio choice with deferred annuities. *Journal of Banking & Finance*, **34**(11), 2652–2664.

Hubener, A., Maurer, R. & **Rogalla, R.** (2014). Optimal portfolio choice with annuities and life insurance for retired couples. *Review of Finance*, **18**(1), 147–188.

Koijen, R.S., Nijman, T.E. & **Werker, B.J.** (2007). Appendix to: when can life cycle investors benefit from time-varying bond risk premia? Available online at the address https://netspar.nl/assets/uploads/nijmanwerker06017app.pdf

Koijen, R.S., Nijman, T.E. & **Werker, B.J.** (2010). When can life cycle investors benefit from time-varying bond risk premia? *The Review of Financial Studies*, **23**(2), 741–780.

Li, Y. & **Rothschild, C.** (2019). Selection and redistribution in the Irish tontines of 1773, 1775, and 1777. *Journal of Risk and Insurance*, forthcoming. https://doi.org/10.1111/jori.12274

Maurer, R., Mitchell, O.S., Rogalla, R. & **Kartashov, V.** (2013). Lifecycle portfolio choice with systematic longevity risk and variable investment-linked deferred annuities. *Journal of Risk and Insurance*, **80**(3), 649–676.

McKeever, K. (2009). Short history of tontines, *Fordham Journal of Corporate and Financial Law*, **15**, 491.

Milevsky, M.A. (2015). *King William's Tontine: Why the Retirement Annuity of the Future Should Resemble Its Past*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139879316

Milevsky, M.A. & **Salisbury, T.S.** (2015). Optimal retirement income tontines. *Insurance: Mathematics and Economics*, **64**, 91–105.

Milevsky, M.A. & **Salisbury, T.S.** (2016). Equitable retirement income tontines: mixing cohorts without discriminating. *ASTIN Bulletin: The Journal of the IAA*, **46**(3), 571–604.

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

Weinert, J.-H. (2017a). *Comparing the Cost of a Tontine with a Tontine Replicating Annuity*. ICIR Working Paper Series (31/2017).

Weinert, J.-H. (2017b). *The Fair Surrender Value of a Tontine*. ICIR Working Paper Series (26/2017).

Weinert, J.-H. & **Gründl, H.** (2016). *The Modern Tontine: An Innovative Instrument for Longevity risk Management in an Aging Society*. ICIR Working Paper Series (22/2016).

World Bank (2015). Databank. Life expectancy at birth. Available online at the address http://data.worldbank.org/indicator/SP.DYN.LE00.IN

Cite this article: Gemmo I, Rogalla R and Weinert J-H (2020). Optimal portfolio choice with tontines under systematic longevity risk. *Annals of Actuarial Science* **14**, 302–315. https://doi.org/10.1017/S1748499520000214