Insurance: Mathematics and Economics 99 (2021) 63–78

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Variable annuities: Market incompleteness and policyholder behavior

Thorsten Moenig

Department of Risk, Insurance, and Healthcare Management; Temple University, 1801 Liacouras Walk, Alter Hall 611, Philadelphia, PA 19122, United States of America

ARTICLE INFO

Article history:
Received May 2020
Received in revised form March 2021
Accepted 12 March 2021
Available online 24 March 2021

JEL classification:
G13, G22

Keywords:
Variable annuities
Optimal policyholder behavior
Market incompleteness
Lifecycle utility model
Guaranteed minimum withdrawal benefits

ABSTRACT

Variable annuities (VAs) are popular personal savings and investment vehicles with long-term guarantees. They include various exercise-dependent features, and the pricing, valuation and hedging of the guarantees depend critically on the investors' decision making. I study whether the optimal exercise behavior of a VA investor is affected by market incompleteness, which arises since the VA generates payout profiles that *intersect* financial risk and idiosyncratic mortality risk and can thus not be (fully) replicated with traditional financial and insurance products.

Implementing a generic VA policy with a withdrawal guarantee (GMWB) that reflects typical US product characteristics, I find that the optimal withdrawal behavior derived from a lifecycle utility model is closely approximated by the corresponding (after-tax) value maximization strategy. That is, from the perspective of the investor, the market around VA policies is sufficiently complete to justify the common approach of assessing optimal VA policyholder behavior with a risk-neutral valuation method.

1. Introduction

Policyholder behavior is an important risk factor for life insurance companies offering contracts that include exercise-dependent features, but is still not well understood (Campbell et al., 2014; Bauer et al., 2017). This is particularly relevant for U.S. variable annuities (VAs), which are personal savings and investment vehicles that currently account for approximately $100 billion in annual sales and around $2 trillion in net assets—that is, nearly a quarter of the U.S. insurance industry's total assets. VAs are equipped with long-term financial guarantees, and the value of these guarantees – as well as the insurers' hedging effectiveness – often depend not only on investment performance and longevity but also critically on the investors' (i.e., policyholders') behavior regarding e.g. withdrawals, lapses/surrenders, and investment allocation.

In this study, I explore whether market incompleteness has a material impact on the investors' optimal VA exercise behavior. VAs intersect the financial risk and the idiosyncratic mortality risk faced by the policyholder. Their payout depends on the performance of the underlying investment, but the guarantee offers downside protection – depending on the guarantee type – either in case of the policyholder's premature death or in case of her longevity. In contrast, traditional financial securities offer exposure to (or protection from) financial risks only, while traditional life insurance or life annuity products allow the investor to manage only her idiosyncratic mortality risk. However, no combination of stocks and life insurance/annuity can replicate the payout structure of a VA guarantee and produce a payment (only) in the joint event where the policyholder dies prematurely *and* her investment has underperformed.

Asset pricing theory suggests that if the market around VA products were complete, then the investor could always replicate any potential VA payout profile through other available assets. Her optimal exercise strategy would then be to maximize the risk-neutral expected value of the VA payouts, and to use the other market assets to construct her desired consumption profile based on her preferences (Cox and Huang, 1989). That is, maximizing her utility of intertemporal consumption or terminal wealth would always produce identical exercise behavior as value-maximization, regardless of her preferences. In an incomplete market, on the other hand, the investor's desired consumption profile may not be attainable without the additional flexibility provided by the VA. This may lead her to adjust her VA exercise behavior: in order to gain utility, she has to give up some of the VA's value.

That is, VAs offer unique payout structures that cannot be replicated by traditional financial or insurance products, and this incompleteness of the market surrounding VA products can impact the investors' optimal VA exercise behavior and thus the value of the VA guarantee. I assess the practical relevance of this effect using a Guaranteed Minimum Withdrawal Benefit (GMWB) rider as a prominent example of VA guarantees. GMWBs allow the policyholder to withdraw her investment amount over the course

E-mail address: moenig@temple.edu.

https://doi.org/10.1016/j.insmatheco.2021.03.007
0167-6687/

of the policy in small annual installments, as long as she remains alive and the policy remains active. They were the most popular VA guarantee in the mid-2000s (Sell, 2006).

In the incomplete market setting, I embed the VA+GMWB product in a lifecycle utility (LC) model where the investor chooses to maximize her utility from annual consumption in an intertemporal setting. I contrast this with the investor's optimal withdrawal behavior derived under the "subjective risk-neutral valuation" (SRNV) framework of Moenig and Bauer (2016), which is the complete-market equivalent of the LC model, and find that the two approaches produce very similar behavior, provided that (in the LC model) the investor has access to a separate investment account. In particular, the difference in the insurer's surplus from the guarantee corresponds to a change of only around 3 basis points (4%) in the guarantee fee under the baseline scenario, with minor variations of the impact depending on the investor's preference parameters, financial situation, and VA contract specifications.

The gap is further reduced if either mortality or the investor's bequest motive are eliminated. This suggests that the residual effect can be attributed largely to the market incompleteness with respect to the investor's idiosyncratic mortality risk, and would be further diminished if the model also granted the investor access to traditional life insurance and annuity products in order to meet her bequest motive and mitigate her financial longevity risk. However, even without that, the market surrounding VA products appears sufficiently complete so that the investor's optimal exercise behavior can be reasonably approximated by a risk-neutral valuation approach.

This valuation method is consistent with the vast majority of studies on policyholder behavior in the VA literature. Among many others, this includes the work of Ulm (2006) on the real option to transfer funds between VA sub-accounts; the optimal GMWB withdrawal behavior assessed e.g. by Milevsky and Salisbury (2006) and Dai et al. (2008); the "universal pricing framework" for VAs by Bauer et al. (2008); (Marshall et al., 2010)'s valuation of a Guaranteed Minimum Income Benefit (GMIB) rider; (Huang et al., 2014) who study the optimal time to begin withdrawals (and thereby lock in the guaranteed annual amount) under a lifetime withdrawal guarantee (GLWB[1]); the research on optimal policy lapses by MacKay et al. (2017) and Piscopo and Rüede (2018); and so-called 1035 (VA policy) exchanges (Moenig and Zhu, 2018). The interested reader may consult (Bauer et al., 2017) for additional references.

Underlying these studies is the common assumption of a complete and frictionless market which allows the respective authors to assess the investors' optimal exercise strategies using arbitrage pricing methods. This approach has the advantage that it simultaneously maximizes the insurer's liabilities from the product, thus reflecting the company's worst-case scenario. This is appealing from a risk management perspective, but often fails to explain competitive market prices (Bauer et al., 2017). Recent studies have shown that asymmetric market frictions – such as tax benefits for the investor (Moenig and Bauer, 2016) or acquisition expenses for the insurer (Moenig and Zhu, 2018) – drive a wedge between the respective valuations of the policyholder and the insurer.[2] In particular, Moenig and Bauer (2016) demonstrate the relevance of these frictions for optimal GMWB withdrawal behavior. Their SRNV model combines a standard risk-neutral valuation approach with the preferential tax treatment of a U.S. VA policy (relative to a potential replicating portfolio). The authors show that only by taking into account the proper tax treatment can optimal withdrawal behavior and fair GMWB fee rates be reconciled with empirical observations from the U.S. VA market.

In line with other recent studies (e.g. Ulm, 2020; Goudenège et al., 2019), I therefore incorporate taxes into my model frameworks and show that the driving factors of policyholder behavior in VAs are in fact the value of the guarantee as well as the relevant market frictions (here, taxes) faced by the investor, rather than the potential incompleteness of the market. Therefore, my findings support the assumption of a complete market setting that underlies the aforementioned studies. However, the vindication is only partial since the studies generally do not account for the product's preferential tax treatment, which can be relevant for the investors' optimal behavior, at least in certain situations.[3]

Assessing the investor's optimal withdrawal behavior requires the numerical implementation of a dynamic optimization problem, which poses significant computational challenges. This is particularly relevant for the LC model which entails additional state and control variables. This made it necessary to implement the models in the Black–Scholes framework—where the interest rate is constant and the asset process follows a Geometric Brownian Motion, similar to e.g. Milevsky and Salisbury (2006), Bauer et al. (2008), and Dai et al. (2008). More recent studies, however, have analyzed withdrawal guarantees under stochastic interest rates and volatility (Peng et al., 2012; Donnelly et al., 2014; Shevchenko and Luo, 2017; Gudkov et al., 2019), a regime-switching process (Lin et al., 2009; Ignatieva et al., 2018), a jump diffusion process (Chen et al., 2008), and a general Lévy framework (Bacinello et al., 2016). They all produce substantially different valuation results (and, of course, hedging strategies) than the corresponding Black–Scholes model.

Fortunately, the SRNV approach requires less than 0.1% of the computation time of the LC model, and is therefore more suitable to include these alternative financial market specifications in order to yield a more accurate valuation of the guarantee and a better prediction of policyholder behavior. The approach can also be used to study alternative contract specifications (e.g. step-ups); tax structures (e.g. a VA in a *non-qualified* retirement savings account); stochastic mortality (though Gudkov et al., 2019 suggest that this may not be a critical factor); investor decisions (e.g. policy (1035) exchanges); as well as the potential impact of other behavioral factors that do not appear "rational" even in a standard utility-based framework, such as anchoring, disposition effect, framing, hindsight bias, and inertia. In addition, the SRNV model can be amended to account for exogenous expenditure or liquidity shocks, which may cause a change in withdrawal behavior if they are not separately insurable. I refer the interested reader to Costabile et al. (2020) who present a GMWB valuation framework that incorporates exogenous withdrawal shocks, to Carpenter (1998) who analyzes the exercise of employee stock options using a risk-neutral valuation model with an exogenous withdrawal state, and to Mahayni and Schneider (2012) who study the investor's optimal allocation strategy inside the VA in the presence of background risk.

While this study focuses on the *policyholders'* market incompleteness and assumes that VA providers operate in a complete market environment, Delong (2014), Blanchet-Scalliet et al. (2015), and Chevalier et al. (2016) assess the problem of pricing VA guarantees if the *insurer* faces an incomplete financial market,

1 A GLWB rider entitles the policyholder to withdraw a guaranteed amount each year for the remainder of her lifetime, without an aggregate withdrawal limit.

2 Forsyth and Vetzal (2014) consider the pricing of a GLWB rider, whereby the optimization problem of the investor – which may be driven by preferences and tax considerations – is separate from but feeds into the pricing and hedging strategy of the insurer which relies on arbitrage-pricing methods.

3 On the other hand, Moenig and Zhu (2018) show that an investor's optimal decision to exchange an existing VA policy for a new one is *not* impacted by neither market incompleteness nor tax considerations.

e.g. as a result of jumps in the asset return process or due to basis risk, that is when the hedging instruments used by the insurer have different returns than the mutual fund(s) underlying the VA guarantee.

The investor's decision process in the lifecycle model entails annual choices regarding the amount to withdraw from (or whether to surrender) the VA+GMWB policy, how much to consume during the year, and how to allocate her remaining wealth in the "outside" investment account. I find that this outside account is a critical component in completing the market: if the investor does not have the opportunity to save and invest her withdrawals – or to use other assets to smooth her consumption over the lifecycle and across the potential realizations of her financial and mortality risk – then her optimal withdrawal behavior from the VA looks substantially different. Therefore, the present study cautions against the assessment of optimal policyholder behavior in an incomplete market setting without accounting for the investor's outside investment options (e.g. Gao and Ulm, 2012, 2015).

This study is most closely related to Horneff et al. (2015) who also embed a VA+GMWB policy in a lifecycle utility framework. While I focus on the valuation of the guarantee and the potential determinants of the investor's exercise behavior, their emphasis is on the changing demand for the GMWB rider and other assets over the lifecycle. In particular, the authors find that withdrawal guarantees offer a meaningful welfare gain – including for younger age groups – in a world where investors also have access to stock and bond investments (but no life insurance or annuities). This is somewhat in contrast to my conclusion that the VA market is reasonably complete, in which case the presence of the GMWB rider would not impact the investor's welfare substantially.

Similar lifecycle approaches have been applied to study VAs with GLWBs. For instance, Steinorth and Mitchell (2015) determine optimal withdrawal behavior for a risk-averse investor, with a focus on the value of the guarantee to the investor. More broadly, lifecycle utility models have been used to assess the benefits of VAs (and related products) in a retirement investment portfolio (see e.g. Brown et al. (2001), Horneff et al. (2010), and Maurer et al. (2013), among others), to study the optimal asset allocation during the payout phase of a VA policy (Charupat and Milevsky, 2002), and to determine optimal partial withdrawals for a VA policy that does not include a withdrawal guarantee (Gao and Ulm, 2015).

The insights provided by the present study should be useful for life insurers and regulators. Since assuming liability-maximizing withdrawal behavior is likely an unfeasible pricing strategy for competitive VA providers, understanding the relevant sources of (optimal) policyholder behavior is critical for the pricing and hedging of the guarantees. In particular, the insight that – aside from marginal tax rates – the investor's unobservable characteristics (i.e., financial situation and preferences) are largely irrelevant to her optimal behavior and the GMWB value should come as good news.

The remainder of the paper is organized as follows: I begin with a brief review of asset pricing theory on market incompleteness and its insights for the case of a VA policy with guarantees. To assess the practical relevance of these insights, I then implement and analyze a VA+GMWB policy in a dynamic optimization model with intertemporal choices: Section 3 describes the model frameworks to determine the investor's optimal withdrawal behavior (in both a lifecycle and a risk-neutral world) as well as the insurer's valuation. This includes a discussion of its numerical implementation and parameterization. I then present the numerical results, followed by conclusions.

2. Theory: Market incompleteness and variable annuities

VAs offer investors exposure to the upside risk of the stock market, while at the same time protecting them against the downside risk, but generally only in case of *either* their survival *or* their death. Thus, for investors, VAs operate at the intersection of their financial risk and their idiosyncratic mortality risk. The level of incompleteness faced by VA investors is therefore based on their ability – using other available assets – to protect themselves against financial risk as well as their mortality or longevity risk.

Consider then a simple world that contains two risk variables, each with two possible future outcomes: financial risk is reflected by either a good economy (G) or a bad economy (B), and idiosyncratic mortality risk is captured by the investor being either alive (A) or dead (D). A (future) "state of the world" is therefore any combination (i, j) with $i \in \{G, B\}$ and $j \in \{A, D\}$.

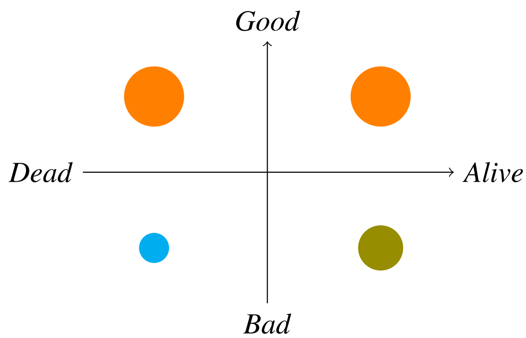
An *Arrow–Debreu security* for state (i, j) pays the numeraire of 1 if state (i, j) occurs, and pays 0 in any other state. Standard asset pricing theory defines a complete market as one where Arrow–Debreu securities exist for every state and are tradable in arbitrary quantities. A complete market allows the investor to allocate her funds F optimally across the different states based on her preferences (such as her level of risk aversion and the strength of her bequest motive), since the existence of the Arrow–Debreu securities ensures that any desired payout structure can be attained. For simplicity, I assume here that there are no market frictions, and let $p_{i,j} \in (0, 1)$ denote the (unique) price of an Arrow–Debreu security for state (i, j).

To relate option exercise with market incompleteness, suppose that – in addition to funds F – the agent also owns an asset Y that promises to pay $Y_{i,j}$ in states (i, j). Moreover, suppose the asset contains an option that – if exercised by the agent – shifts some payouts from one state to another, producing a payout profile $Y'_{i,j}$. Does the agent's optimal exercise decision depend on her preferences? The answer depends on whether the market is complete (without asset Y).

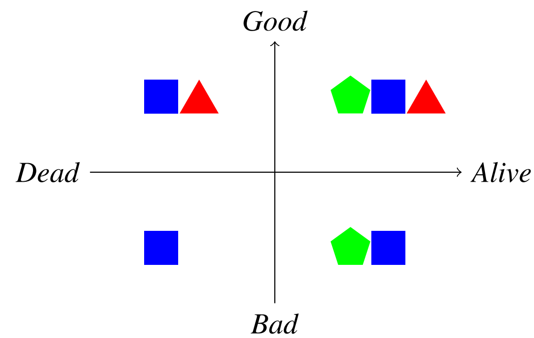
If the market is *complete*, then the agent can sell the state-contingent payouts $Y_{i,j}$ for $V_Y := \sum_{i,j} p_{i,j} Y_{i,j}$, and she can then allocate her total funds $F + V_Y$ across the four states according to her preferences. She can do the same if she exercises the option and receives $V'_Y := \sum_{i,j} p_{i,j} Y'_{i,j}$. Therefore, the agent should exercise the option if and only if $V'_Y > V_Y$ because that provides her with the most funds.[4] That is, in a complete market, the policyholder of a VA contract should optimally exercise the embedded options based on this principle of value maximization. In particular, this would produce the same optimal behavior as if she was maximizing (e.g.) the expected utility of her consumption, irrespective of her preferences.

On the other hand, if the market is *incomplete*, the potential payout profiles $Y_{i,j}$ and $Y'_{i,j}$ cannot necessarily be replicated with other assets. Therefore, no objective monetary "value" can be assigned to Asset Y with and without option exercise. The optimal exercise decision then depends on the agent's preferences over payouts in the different states and the payout profile of asset Y with and without the exercised option, respectively. Thus, in an incomplete market setting, the investor's optimal exercise behavior is generally derived from some form of utility maximization, and the extent to which she chooses to deviate from the value-maximizing behavior depends critically on the level of market incompleteness, that is how well she can replicate the potential option payouts with her other available assets.

4 More practically, she would simply embed the $Y_{i,j}$ in the total payouts that she would choose based on available funds $F + V_Y$, and only sell parts of the asset payout if necessary.



(a) VA + Guarantee



(b) Other Securities

Fig. 1. Relative Investment Payouts in Simple 4-State World. Notes: The figures reflect the potential payouts that an investor can receive from different investments. In panel (a), the size of the circle indicates the magnitude of the payout from a VA that includes a living benefit guarantee. In panel (b), the blue square reflects payouts from a risk-free asset; the red triangle reflects payouts from a stock investment; and the green pentagon reflects payouts from a traditional life annuity.

The impact of market incompleteness on optimal policyholder behavior for VAs therefore depends on what other assets the investor has access to or, more precisely, to what degree these assets can replicate the potential VA payout profiles. VAs with guarantees offer a relatively unique payout structure that links the investor's financial and mortality risk. The more common assets in this space are risk-free investments such as treasury bonds or bank accounts, stocks, and traditional life insurance and life annuity products.

To assess the impact of these assets on the level of incompleteness of the VA market, consider again our simple, four-state world. Fig. 1 illustrates the possible payout structures from both the VA (with a living benefit guarantee) and from a portfolio consisting of the other assets. In particular, as panel (a) of Fig. 1 indicates, the VA is primarily an investment vehicle whose returns are tied to the state of the economy overall. Therefore, payouts are larger in the "Good Economy" states. In the "Bad Economy" states, the VA account has lost value. However, due to the living benefit guarantee, the insurer has to make up for the investor's losses if she survives. Therefore, the investor receives a larger payout in the "Bad/Alive" state than in the "Bad/Dead" state. Fig. 1(b) reflects the payouts from the more common investment options: risk-free investments (blue squares) provide the same payout in all states; stocks (red triangles) offer an additional payout if the economy is doing well; and a traditional life annuity (green triangle) makes fixed payouts if and only if the investor is alive.

Comparing the two panels of Fig. 1, one can see that no combination of the assets from panel (b) can fully replicate the payout profile of a VA with guarantee. That is, even with unrestricted access to bank accounts, stocks, and life annuities, the market around VAs is still incomplete. This is because the VA policy provides a separate payout in the *joint* state where the economy is "Bad" and the investor is "Alive", while a stock or a traditional life annuity can discriminate payouts only across one of the two dimensions. To attain specific payouts in each of the four joint states, the investor requires a third asset that intersects the financial and mortality risk. VAs with guarantees are such assets and can thus help complete the market for investors looking to protect themselves financially against both economic and mortality risk.

However, one can also see from Fig. 1 that the assets from panel (b) help complete the VA market, possibly to a large degree. With risk-free asset and stocks only (blue squares and red triangles), the investor can replicate the *overall* discrepancy in VA payouts between the "Good Economy" and the "Bad Economy" states. With risk-free asset and life annuities (blue squares and green pentagons), the investor can replicate the *overall* discrepancy in VA payouts between the "Dead" and the "Alive" states. And if given access to all three assets (i.e., risk-free asset, stocks, and life annuities) simultaneously, the investor could replicate the VA payout even more closely.

This begs the question to what extent the VA market is complete for a typical investor. That is, to what degree will her optimal decision making differ from value maximization? To answer that, I implement a dynamic optimization problem to assess the optimal behavior of an investor who owns a VA policy with a GMWB rider, as a popular and representative example of a VA guarantee. In particular, I compare the optimal withdrawal behavior – through its impact on the insurer's valuation of the guarantee – under a risk-neutral approach in a complete market with that under a lifecycle utility model. The latter reflects the (more realistic) setting of a VA investor facing an incomplete market environment in which restrictions to outside investment options do not allow for a perfect replication of the potential VA payouts.

Fig. 1 suggests that the discrepancy in VA payouts appears to be less significant across the mortality states than across the economic states. That is, access to the stock market likely does more to complete the VA market than access to traditional life annuities. In addition, stock markets for individual investors tend to be more liquid than the market for life annuities. Due to computational restrictions in the LC model, I therefore give the investor access to stocks but not to traditional life insurance products and life annuities. My numerical results ultimately justify this setup, as I find that the access to stocks sufficiently completes the VA market for the investor for all practical purposes. In the following section, I present the model framework and the numerical implementation procedure. Thereafter, I present and discuss the results.

3. Modeling optimal withdrawal behavior for VA+GMWBs

VA+GMWB products available in the U.S. market contain a diverse set of features. They differ by fee rates and how these are assessed; guarantee types (return-of-premium vs. step-up, as well the frequency of such a step-up); the guaranteed annual withdrawal amount; surrender charges; fund (re-)allocation options; potential fee waivers in case of non-withdrawals; and purchase payment credits; among others. They can be purchased both inside and outside of a qualified retirement savings account. I model a generic VA+GMWB policy that reflects the core product features and decision criteria, but make a few necessary simplifications for computational feasibility. I believe that these simplifications are immaterial to the key insights of the analysis.

I model the optimal behavior of an investor who has just purchased a VA+GMWB policy inside a qualified retirement savings plan.[5] The investor also owns a (non-qualified) savings/investment account. She derives utility from consumption on an annual basis. Therefore, while alive, the investor chooses how much to withdraw from the VA policy, how much to consume, and how to allocate her funds in the outside investment account. In the incomplete-market setting, these decisions are made jointly at the beginning of each policy year and in order to maximize the discounted expected utility of current and future consumption. In a complete market, the investor's outside account and her consumption decisions are irrelevant to her optimal withdrawal behavior; therefore, only the investor's withdrawal decision will be incorporated into the model.

In this section, I will first characterize the financial market and the investor's mortality risk. Next, I describe the VA+GMWB product, before stating the investor's optimization problem in a lifecycle model as well as its risk-neutral equivalent. I proceed with the insurer's valuation method, briefly describe the numerical implementation procedure, and finally present the values used to parameterize the model.

3.1. Financial & mortality risk

The financial market consists of a single risky asset, with time-t price S_t, and a risk-free asset that compounds (continuously) at annual rate r. The risky asset does not pay dividends, follows the Black–Scholes model with constant volatility $\sigma > 0$ and drift parameter $\mu \geq 0$, and thus evolves according to the stochastic differential equation

$$\frac{dS_t}{S_t} = \mu\, dt + \sigma\, dZ_t\,, \quad S_0 > 0, \tag{1}$$

where $(Z_t)_{t>0}$ is a standard Brownian motion.

The investor also faces idiosyncratic mortality risk. Following standard actuarial notation, q_t denotes the conditional probability that an (alive) individual age t dies in the following year, that is prior to age $t+1$, and ${}_tp_x$ denotes the probability that an x-year old survives the next t years.

Consistent with the actuarial literature, I define $\mathbb{P}$ as the product measure of the *physical* probability measure of the financial market and the physical measure of the investor's mortality rates, and $\mathbb{Q}$ as the product measure of the *risk-neutral* measure for financial risk and the physical measure for idiosyncratic mortality risk.

3.2. VA product features

At age x, the investor purchases a T-year VA+GMWB policy with a single premium M. The money is placed into a mutual fund that invests in a combination of the two available assets. VA providers generally limit the fund's potential equity exposure in order to reduce the value of – and improve their hedging effectiveness for – the guarantee, which acts as downside protection on the investment. To reflect this, I assume that the fund maintains a constant equity exposure, whereby a share $\rho \in (0, 1]$ is invested in the risky asset, with the remainder allocated to the risk-free asset.

Let A_t denote the VA account value at time t. Since U.S. VAs do not contain sales charges, we have the initial condition $A_0 = M$. Instead, the insurer uses a periodic fee to recover its acquisition expenses and to cover its administrative costs. Unlike other studies, I include this VA base fee in our model because – as (Chen et al., 2008) and Sun et al. (2018) show – it has a substantial impact on the value of the GMWB rider. In addition, the insurer charges a guarantee fee for the GMWB rider. These fees are assessed continuously at annual rates φ^{base} and φ^{guar}, respectively, in proportion to the concurrent VA account value, and are taken directly out of the VA account. That is, no additional capital injections by the investor are necessary. At times of non-withdrawals, the VA account value thus evolves as follows:

$$\frac{dA_t}{A_t} = \left(\mu^{\text{VA}} - \varphi^{\text{base}} - \varphi^{\text{guar}}\right) dt + \sigma^{\text{VA}}\, dZ_t\,, \tag{2}$$

with $\mu^{\text{VA}} = \rho\mu + (1-\rho)r$ and $\sigma^{\text{VA}} = \rho\sigma$.

The VA policy includes a return-of-premium GMWB rider[6] that allows the policyholder to withdraw the initial investment M in annual installments of up to g, while the policy remains active and irrespective of the VA's financial performance.

Thereby, withdrawals are taken directly out of the VA account, if sufficient funds are available. That is, if the policyholder withdraws an amount w_t at time t, and if A_{t^-} denotes the VA account value just prior to the withdrawal, then the VA account value immediately after the withdrawal is given by

$$A_{t^+} = \left(A_{t^-} - w_t\right)^+\,, \tag{3}$$

where $(a)^+ \equiv \max\{a, 0\}$. If the VA account is not sufficient to cover the withdrawal, then the insurer pays the difference, that is

$$X_t^W := \left(w_t - A_{t^-}\right)^+\,, \tag{4}$$

out of its general fund. Between withdrawal dates t^+ and $(t+1)^-$, the account value evolves according to Eq. (2), so that:

$$A_{(t+1)^-} = A_{t^+} \cdot e^{\mu^{\text{VA}} - \varphi^{\text{base}} - \varphi^{\text{guar}} - \frac{\left(\sigma^{\text{VA}}\right)^2}{2} + \sigma^{\text{VA}} \cdot \Delta Z_t}\,, \tag{5}$$

where $\Delta Z_t = Z_{t+1} - Z_t$ follows a standard normal distribution.

The remaining guaranteed withdrawal amount – whose time-t value I denote by G_t, for $t = 1, 2, \ldots, T$ – is adjusted as well upon a withdrawal, namely as follows:

$$G_{t+1} = \begin{cases} \left(G_t - w_t\right)^+ & \text{, if } w_t \leq g \\ \left(\min\left\{G_t - w_t\,,\, A_{t^+}\right\}\right)^+ & \text{, if } w_t > g\,, \end{cases} \tag{6}$$

with $G_1 = M$. That is, in the case of a regular withdrawal, the remaining guaranteed amount is reduced by the amount withdrawn, while withdrawals in excess of the annual guaranteed amount g result in a potentially more significant reduction of the remaining guaranteed withdrawal amount. This adjustment formula (or a similar version) is common practice and serves to discourage excess withdrawals.

In early policy years, excess withdrawals are also subject to a surrender charge at rate s_t that is applied in proportion to the excess amount withdrawn at time t. Moreover, the U.S. federal tax authority imposes a 10% early withdrawal penalty on all withdrawals prior to age 59.5. That is, upon a time-t withdrawal of w_t, the policyholder receives a (pre-tax) payout of

$$w_t^{\text{gross}} = w_t - (w_t - g)^+ \cdot s_t - 0.1 \cdot \left[w_t - (w_t - g)^+ \cdot s_t\right] \cdot \mathbb{1}_{x+t<59.5}\,.$$

[5] The use of a *qualified* account – in contrast to Moenig and Bauer (2016) – drastically reduces the computation time as it does not require the VA's remaining "tax base" as an additional state variable.

[6] A step-up feature can be implemented without much complication in a risk-neutral setting, by expressing the account value and the remaining guaranteed amount in relations to the "baseline" guaranteed amount, and updating all state variables accordingly upon a step-up decision. However, in a lifecycle model with an outside investment account, this relationship does not extend to the outside account, so that here the model would have to include not only the outside account value but also the baseline guaranteed amount as relevant state variables. To maintain computational feasibility, I therefore implement only return-of-premium riders, for which the baseline guaranteed amount remains constant at the initial investment amount.

If the policyholder dies in policy year t, the beneficiaries receive the VA account value at the end of the year of death, that is

$$DB_t = A_{t^-} ,$$

and the policy terminates. If the policy is still active at maturity, the policyholder receives the VA account value A_T – or the annual guaranteed amount, if larger – in a lump-sum payment:

$$w_T = \max\{A_{T^-}, \min\{g, G_T\}\} .$$

The VA is located *inside* a qualified retirement savings account, and I model taxation accordingly based on current U.S. law: all contributions to such a qualified account are pre-tax, no taxes are assessed while the funds remain in the account, and all withdrawals from the account are taxed as ordinary income. If the investor has a constant marginal income tax rate of τ, the after-tax withdrawal amount is thus given by

$$w_t^{\text{net}} = w_t^{\text{gross}} \cdot (1-\tau) .$$

The income tax will be similarly applied to the payout at maturity ($w_T^{\text{gross}}(1-\tau)$) and the death benefit payout ($DB_t(1-\tau)$).[7]

3.3. Utility-based optimization

Outside investment account

The investor also maintains a separate, non-qualified, investment account. I denote the time-t value of this "outside" account by O_t. At policy anniversary dates, that is for $t = 0, 1, \dots, T-1$, the policyholder chooses the ratio $\nu_t \in [0, 1]$ that indicates the account's equity investment (in the market's risky asset) over policy year $t+1$. The remainder is invested in the risk-free asset. All earnings over the year are taxed at time $t+1$ at the (constant) marginal capital gains tax rate κ. That is, between policy anniversary dates, the outside account evolves as follows:

$$O_{(t+1)^-} = O_{t^+} \cdot \left[\left(\nu_t \cdot \frac{S_{t+1}}{S_t} + (1-\nu_t) \cdot e^r \right) - \kappa \cdot \left(\nu_t \cdot \frac{S_{t+1}}{S_t} + (1-\nu_t) \cdot e^r - 1 \right)^+ \right] . \tag{7}$$

On policy anniversary date t, the investor receives income I, withdraws w_t from the VA, and consumes an amount $c_t > 0$ for the year. The outside account value is therefore updated according to:

$$O_{t^+} = O_{t^-} + I + w_t^{\text{net}} - c_t . \tag{8}$$

Both O_{t^-} and O_{t^+} are after-tax account values, so that the investor's overall bequest amount in case of her premature death in policy year t is given by

$$b_t := DB_t \cdot (1-\tau) + O_{t^-} . \tag{9}$$

Investor preferences

Following (e.g.) (Lockwood, 2012), the investor strives to maximize her expected utility of current and future consumption, using a subjective annual discount factor β. At the beginning of any given year, she gains utility $u_C(x)$ from consuming an (after-tax) amount x, where the utility function $u_C(.)$ exhibits constant relative risk aversion and thus has the functional form

$$u_C(x) = \frac{x^{1-\gamma}}{1-\gamma} , \tag{10}$$

with relative risk aversion coefficient γ.[8]

For bequests, I assume that the bequest amount is converted into a perpetuity (with annual payments to the beneficiaries, beginning immediately). Therefore, the utility to the investor from bequeathing an after-tax amount b_t is given by

$$u_B(b_t) = B \cdot u_C\left(I^{\text{ben.}} + b_t \cdot \left[1 - e^{-r}\right]\right) \cdot \frac{1}{1-\beta} , \tag{11}$$

where B reflects the strength of the investor's bequest motive, and $I^{\text{ben.}}$ is the future annual income of the beneficiary.

In the model's sensitivity analysis, I consider alternative specifications to the investor's utility from consumption and bequests, which I will describe at that time.

Investor's optimization problem

The investor makes annual decisions regarding the consumption amount c_t, a potential withdrawal from the VA w_t, and the equity allocation ratio ν_t in the outside account. The investor gains utility from consumption and bequests. Her objective at time t therefore is to maximize her discounted expected utility, denoted by U_t, of all potential consumption and bequests from time t onward. This value function depends on the state variables A_{t^-}, G_t, and O_{t^-}, which I summarize in the state vector $y_t = (A_{t^-}, G_t, O_{t^-})$.

The investor's optimization problem can thus be described recursively (for $t = 1, \dots, T-1$) by the following Bellman equation:

$$U_t(y_t) = \max_{c_t, w_t, \nu_t} u_C(c_t) + \beta \cdot \mathbb{E}_t^{\mathbb{P}} \left[q_{x+t} \cdot u_B(b_{t+1}) + (1-q_{x+t}) \cdot U_{t+1}(y_{t+1}) \right] , \tag{12}$$

subject to the updating conditions (3), (5), (6), (7), and (8); the utility functions defined by Eqs. (10) and (11); the bequest amount (9); the asset movement given by Eq. (1); and the choice variable constraints

$$\begin{aligned} 0 &\le c_t \le O_{t^-} + I + w_t^{\text{net}} \\ 0 &\le \nu_t \le 1, \quad \text{and} \\ 0 &\le w_t \le \max\{A_{t^-}, \min\{g, G_t\}\} . \end{aligned}$$

Terminal condition

If the investor is alive when the VA matures at time T, she converts her entire savings, that is

$$W_T := w_T(1-\tau) + O_{T^-} ,$$

into a whole life annuity, which provides her a constant pre-tax payment of $W_T/\ddot{a}_{x+T}$ at the beginning of each year – beginning immediately – for the remainder of her life. The term $\ddot{a}_{x+T}$ is used consistently with standard actuarial notation to indicate the expected present value of \$1 paid at the beginning of each year for the remainder of the lifetime of an individual age $(x+T)$, and is computed as

$$\ddot{a}_{x+T} = \sum_{t=0}^{\infty} {}_t p_{x+T} \, e^{-rt} .$$

The earnings portion of the annuity payout will be taxed as income, that is at rate τ. Here, W_T serves as the investor's tax

[7] I also implemented this model assuming full annuitization of the VA payouts at maturity, and found that this assumption did not alter the main findings. Furthermore, I assume here that the beneficiaries also choose the lump-sum VA payout (as opposed to potentially continuing the qualified retirement account, subject to the relevant minimal distribution rules), and that they have the same marginal income tax rate as the policyholder. In light of the overall findings, I believe these assumptions to be innocuous as well.

[8] Strictly speaking, this functional form applies only for $\gamma \neq 1$. The case of $\gamma = 1$ is represented by $u_C(x) = \ln(x)$.

base. U.S. tax law prescribes that annuity payments in excess of W_T/e_{x+T} are considered earnings and thus taxed as income, in our case at rate τ. Here, e_{x+T} denotes the investor's life expectancy at age $x+T$, as specified by the Internal Revenue Service in *Publication 590-B*. The investor's after-tax payout from the annuity is therefore

$$c_T^{\text{net}} := \frac{W_T}{\ddot{a}_{x+T}} - \tau \cdot \left(\frac{W_T}{\ddot{a}_{x+T}} - \frac{W_T}{e_{x+T}} \right)^+ .$$

In addition, the investor continues to receive outside income I. Her annual utility from consumption is therefore $u_C(c_T^{\text{net}} + I)$, so that her time-T utility – and thus the terminal condition of her optimization problem (12) – is given by

$$U_T(y_T) = u_C\left(c_T^{\text{net}} + I\right) \cdot \sum_{t=0}^{\infty} {}_t p_{x+T}\, \beta^t . \tag{13}$$

Valuation at time 0

The investor's starting conditions are

$$A_{0^+} = M ; \qquad G_1 = M ; \qquad O_{0^+} = W_0 , \tag{14}$$

where W_0 denotes the investor's outside-account wealth at the time of her VA purchase.

At time 0, the investor does not make a withdrawal from her newly purchased VA, and only chooses optimal consumption c_0 and outside investment allocation ν_0 in order to maximize her discounted expected utility, akin to Eq. (12). I define this initial utility as U_0, given by

$$U_0 := \max_{c_0, \nu_0} u_C(c_0) + \beta \cdot \mathbb{E}_0^{\mathbb{P}}\left[q_x \cdot u_B(b_1) + (1 - q_x) \cdot U_1(y_1)\right], \tag{15}$$

subject to the starting conditions (14) and the updating conditions (5) and (7), along with the aforementioned definitions and choice variable constraints.

3.4. Risk-neutral valuation with taxes

I adopt the SRNV framework of Moenig and Bauer (2016) to value the VA+GMWB policy in a risk-neutral setting, as a complete-market equivalent to the lifecycle model described above. The SRNV model assumes a complete market – for both financial risk and idiosyncratic mortality risk – but accounts for the VA's preferential tax treatment relative to its replicating portfolio.

In particular, consider a stochastic after-tax cash flow Y occurring at time $t+1$. The time-t subjective risk-neutral value of Y is then given as the amount of (after-tax) money X that must be placed at time t in a portfolio and then (re-) invested in such a way that it exactly replicates Y at the end of the year—*after* the current earnings of the replicating portfolio have been taxed accordingly. Here, the assumption of a complete market ensures that such a replicating strategy exists.

If the investor has a constant marginal capital gains tax rate κ at which the portfolio earnings are taxed annually, Moenig and Bauer (2016) show that the SRNV method produces a unique value X given implicitly by:

$$e^r X = \mathbb{E}^{\mathbb{Q}}[Y] + \frac{\kappa}{1-\kappa} \cdot \mathbb{E}^{\mathbb{Q}}\left[(Y - X)^+\right] . \tag{16}$$

To apply the SRNV formula to the objective of this study, let V_t denote the time-t "subjective" value of the VA+GMWB policy to the investor, assessed immediately prior to the investor's time-t withdrawal decision (similar to U_t in the utility-based model). V_t depends on the VA account value A_{t^-} and the remaining guaranteed amount G_t. That is, in this complete-market model the state vector is $y_t = (A_{t^-}, G_t)$. V_t is the sum of the current after-tax payment from the VA (that is, w_t^{net}) and the *continuation value* of the VA following the withdrawal. The latter can be viewed as the current after-tax value of the VA at the end of the year, that is the subjective risk-neutral value of the time-$(t+1)$ VA value, which is either the death benefit payment ($DB_{t+1}(1-\tau)$) in case of death, or V_{t+1} if the investor survives the next year. I denote the continuation value by $\tilde{V}_t$ and determine it implicitly as the value of X in Eq. (16) that corresponds to

$$Y = q_{x+t} \cdot DB_{t+1} \cdot (1 - \tau) + (1 - q_{x+t}) \cdot V_{t+1}(y_{t+1}) .$$

The investor's dynamic optimization problem (for $t = 1, \ldots, T-1$) is thus

$$V_t(y_t) = \max_{w_t} w_t^{\text{net}} + \tilde{V}_t , \tag{17}$$

subject to the updating conditions (3), (5), and (6); the asset movement given by Eq. (1); the choice variable constraint

$$0 \le w_t \le \max\{A_{t^-}, \min\{g, G_t\}\} ;$$

and the terminal condition

$$V_T(y_T) = w_T(1 - \tau) . \tag{18}$$

At time 0, no optimization is necessary, as the initial risk-neutral after-tax value of the VA+GMWB policy is given by

$$V_0 := \tilde{V}_0 , \tag{19}$$

using the first two starting conditions of (14).

3.5. Insurer's valuation

The expected amount of guarantee fees that the insurer collects between time t^+ and time $(t+1)^-$ (for $t = 0, 1, \ldots, T-1$) is approximated by

$$X_t^F := (w_t - g)^+ \cdot s_t + \frac{\varphi^{\text{guar}}}{\varphi^{\text{base}} + \varphi^{\text{guar}}} \cdot A_t^+ \cdot \left(1 - e^{-\varphi^{\text{base}} - \varphi^{\text{guar}}}\right) .$$

The first part of this equation reflects the excess withdrawal fees at time t, if applicable, while the second part approximates the GMWB rider fees for the upcoming policy year. Thereby, the impact of the aggregate fee rate on the VA account value is akin to that of a continuously paid dividend on a stock price (Milevsky and Posner, 2001). In exchange, the insurer pays X_t^W—defined in Eq. (4)—to cover the cost of the withdrawal guarantee, when needed.

The VA+GMWB rider can also be valued recursively from the insurer's perspective. To do so, I define V_t^F and V_t^W as the expected present values of all forthcoming fee payments to the insurer and withdrawal payments from the insurer, respectively. These values are assessed immediately before the time-t withdrawal decision is made, and depend on the current state vector y_t, in addition to the investor's optimal withdrawal decision w_t. Therefore:

$$\begin{aligned} V_t^F(y_t) &= X_t^F + e^{-r}(1 - q_{x+t})\,\mathbb{E}_t^{\mathbb{Q}}\left[V_{t+1}^F(y_{t+1})\right] \\ V_t^W(y_t) &= X_t^W + e^{-r}(1 - q_{x+t})\,\mathbb{E}_t^{\mathbb{Q}}\left[V_{t+1}^W(y_{t+1})\right] , \end{aligned} \tag{20}$$

subject to the updating conditions (3), (5), (6), (7), and (8), and the terminal conditions

$$\begin{aligned} V_T^F(y_T) &= 0 \\ V_T^W(y_T) &= (\min\{g, G_T\} - A_{T^-})^+ . \end{aligned} \tag{21}$$

The starting conditions for the insurer are given in (14), with the initial valuation statistics given by

$$\begin{aligned} V_0^F &= X_0^F + e^{-r}(1 - q_x)\,\mathbb{E}_0^{\mathbb{Q}}\left[V_1^F(y_1)\right] \\ V_0^W &= e^{-r}(1 - q_x)\,\mathbb{E}_0^{\mathbb{Q}}\left[V_1^W(y_1)\right] . \end{aligned} \tag{22}$$

If the investor optimizes within the SRNV framework, the above references to the outside investment account value O_t

can be dropped; in all other regards, the insurer's valuation methodology is identical.

Lastly, I define the insurer's initial surplus from the GMWB rider as

$$\Sigma_0 := V_0^F - V_0^W .$$

3.6. Numerical implementation

Since the investor's optimization problems do not have closed-form solutions, I implement them numerically using *recursive dynamic programming*. This entails the following steps:[9]

(i) Discretize the state space grid consisting of the variables $A_{t-} \in [0, \infty)$, $G_t \in [0, M]$, and $O_{t-} \in [0, \infty)$ in order to create a discrete, three-dimensional grid. Let $\mathbb{G}$ denote the set of all grid points.

(ii) Set terminal conditions: For each grid point $(A, G, O) \in \mathbb{G}$, compute $U_T(A, G, O)$ using Eq. (13), as well as $V_T^F(A, G, O)$ and $V_T^W(A, G, O)$ using Eq. (21).

(iii) Recursive optimization: For each $t = T-1, T-2, \ldots, 1$, and – within each time step t – for each grid point $(A, G, O) \in \mathbb{G}$:

- Solve the investor's optimization problem specified by Eq. (12), thereby computing w_t and $U_t(A, G, O)$. Approximate the distribution of asset returns (i.e., S_{t+1}/S_t) using Gauss–Hermite quadrature; and approximate the end-of-period utility values $U_{t+1}(.)$ with linear interpolation between the utility values computed on the grid points during the previous time step $(t + 1)$.
- Compute $V_t^F(A, G, O)$ and $V_t^W(A, G, O)$ using Eq. (20). Thereby, use the values of w_t, c_t, and ν_t that solve the investor's optimization problem at this grid point. Again, rely on Gauss–Hermite quadrature and linear interpolation over the time-$(t + 1)$ grid values in order to approximate asset return distribution and end-of-year continuation values, respectively.

(iv) Determine the initial values to investor—U_0, using Eq. (15)—and insurer—V_0^F and V_0^W, using Eq. (22)—based on the starting conditions (14).

Computational considerations

The numerical implementation of the utility-based optimization problem is computationally intensive, due to the three state variables and the three control variables as well as the long maturity of the VA policy. Therefore, it is crucial to select a state space grid that is sufficiently but not unnecessarily fine.

In addition, it can be beneficial to put reasonable restrictions on the choice variables. In particular, I restricted the withdrawal amount w_t to a few possible choices, including 0, the guaranteed annual amount $\min\{g, G_t\}$, and the VA account value A_t (which corresponds to a policy surrender); I searched for the optimal consumption amount c_t in multiples of $1,000 (with a larger step size for consumption in excess of $300,000); and I only allowed the investment allocation ν_t to be an integer percentage. I verified that all these restrictions have a negligible impact on the investor's optimal behavior and on the product valuation in these models.

[9] This algorithm applies to the lifecycle model. The implementation of the SRNV optimization follows the same procedure, except that all references to the outside investment account O_{t-} are dropped, so that the state space grid is two-dimensional. References to c_t and ν_t may also be ignored. Moreover, $U_t(.)$ is replaced by $V_t(.)$, and references to Eqs. (12), (13), and (15) are replaced by Eqs. (17), (18), and (19), respectively.

Even with these simplifications and using 32 cores on a 3.0 GHz workstation, the utility-based optimization problem – coded in *Matlab* with a relatively efficient parallelization – takes more than 4 h to complete. In contrast, under the same conditions, the SRNV model finishes in approximately 10 s.

3.7. Parameter choices

For ease of reference, all parameter choices for the numerical implementation of the above optimization models are provided in Table 1.

For the *baseline* scenario, I consider a 60-year old individual who has just purchased a $100,000 VA policy with a 7% return-of-premium GMWB rider. The VA base fee is 100 bps; it accounts for the insurer's expenses and administrative costs, and is approximately in line with what insurers charge for such products (although there is considerable heterogeneity, see Moenig and Bauer, 2016). The guarantee fee rate is 70 bps, which approximately covers the cost of the GMWB rider in this baseline scenario.

The money is placed in a mutual fund that maintains an equity exposure of 80%. This reflects the common practice by VA providers to restrict investment options (specifically, equity exposure) in cases where the policyholder elects a guarantee rider such as a GMWB (Moenig and Bauer, 2016). The policy matures after 20 years, and – as is typical for B-share VAs, the most commonly sold product type – has a 7-year, decreasing surrender fee schedule. The investor's mortality rates follow the Society of Actuaries' individual annuitants' *2012 IAM Basic Table – Female*.

The investor's tax parameters are taken from Moenig and Bauer (2016) (30% marginal income tax rate and 23% marginal capital gains tax rate), and the tax-related life expectancy at maturity is prescribed by *Publication 590-B* of the U.S. Internal Revenue Service.

The financial market parameters are based on the historical returns of the S&P 500 index and 3-month Treasury Bill rates between 1928 and 2018. Using geometric averages, over this time frame, the index grew at approximately 9% per year, with an annual volatility of nearly 20%, and the treasury bills accumulated at approximately 3.3% per year.[10] To reflect the recent low-interest rate climate, I use a 3% risk-free rate in the baseline scenario of our model.

Brown and Poterba (2006) infer from the *2001 Survey of Consumer Finances* that over 70% of all U.S. VAs at the time were owned by households in the highest decile of wealth, and nearly 90% were owned by the two highest deciles. In 2017, the 90th percentile of wealth among U.S. households was nearly $1,200,000, and the 80th percentile was approximately $500,000.[11] Based on this information, I assume that the investor's initial (outside) wealth is $1,000,000.

I further assume that the investor is (largely) retired and receives only a modest annual after-tax income of $25,000 (e.g. through Social Security retirement benefits, IRA distributions, etc.). In addition to her income, she lives off of her current wealth as well as her potential VA payouts.

As with the *shape* of the utility functions, I follow (Lockwood, 2012) for their *parameterization*. In particular, the investor is moderately risk-averse with a CRRA coefficient of 2.0, her subjective discount rate is identical to the market interest rate, and the beneficiary's annual income equals her own. Moreover, the

[10] Source: http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/histretSP.html.

[11] Source: https://dqydj.com/net-worth-brackets-wealth-brackets-one-percent/.

Table 1
Parameter choices for numerical implementation.

Parameter	Description	Baseline	Sensitivity
VA+GMWB contract specification			
M	VA single premium	100,000	
ρ	Share of equity allocation	80%	100%
φ^{base}	VA base fee rate	100 bps	
φ^{guar}	GMWB fee rate	70 bps	
g	Guaranteed annual withdrawal amount	7000	
s_t	Surrender / excess withdrawal fee rate	7%, 6%, …, 1%; 0 for $t \geq 7$	
T	Time to maturity (years)	20	25
Taxation			
τ	Marginal income tax rate	30%	35%
κ	Marginal capital gains tax rate	23%	25%
e_{x+T}	Life expectancy at maturity (years)	13.4	(adjusted)
Financial market			
r	Risk-free rate	3%	4%
σ	Volatility of risky asset	20%	25%
μ	Expected growth rate of risky asset	9%	(adjusted)
Investor characteristics & preferences			
x	Age at inception	60	70
q_{x+t}	Mortality rates	*2012 IAM Basic Table – Female*	0
W_0	Outside wealth at time 0	1,000,000	500,000; 100,000
I	Future annual income (post-tax)	25,000	50,000
γ	CRRA parameter	2.0	3.0
β	Subjective discount factor	e^{-r}	0.90
B	Strength of bequest motive	1.0	0.0
$I^{\text{ben.}}$	Income of beneficiary	25,000	50,000

strength of the investor's bequest motive is 1.0—that is, she values consumption by her beneficiaries as much as her own.

I relax these parameters – one at a time – in the sensitivity analysis, according to the specifications in the last column of Table 1.

4. Results

I implement the investor's optimization problems and the corresponding guarantee valuation of the VA provider, as described in the previous section. This section begins with a brief analysis of the baseline case, that is the parameterization described in the third column of Table 1, contrasting the findings from the lifecycle optimization with its risk-neutral equivalent, the SRNV approach. I then vary the relevant input parameters and characteristics that are exclusive to the lifecycle model. These results are shown in Table 2. Lastly, I consider variations to the VA policy as well as the investor's tax rates. These variations also impact the SRNV approach, and the corresponding valuation results are displayed in Table 3.

4.1. Baseline case

Results for the baseline parameterization are displayed in columns [1] and [2] of Table 2 for the SRNV and the lifecycle case, respectively.

I find that the risk-neutral after-tax value of the $100,000 VA policy is around $78,000. That is, the policy is worth around 78 cents for every dollar of investment. This exceeds the after-tax value of the investment if it were to be withdrawn from the qualified account at time 0, which is $(1-\tau)$, that is 70 cents. Therefore, the valuation method used in this study is consistent in the sense that the VA policy is a worthwhile investment, at least in comparison to the replicating portfolio. As a result, the optimal withdrawal behavior predicted by this model can be taken seriously.

The investor's withdrawals account for nearly $8,500 in liabilities from the GMWB rider (in risk-neutral expected present value terms). The guarantee fee of 70bps is (just) sufficient to cover these costs, resulting in a small surplus of $325 to the VA provider in the SRNV setting. Column [2] of Table 2 shows that the valuation results under the lifecycle optimization are very similar, with a slightly larger guarantee value and fewer fee payments, resulting in a surplus of $138.

For comparison, a one basis point increase in the guarantee fee corresponds to approximately a $30 increase in the GMWB value and a $60 increase in the insurer's surplus from the guarantee. That is, the risk-neutral optimization approximates the lifecycle model very closely, with the discrepancy in surplus ($187) accounting for less than 0.2% of the investment amount or around 3 basis points in the guarantee fee.

This proximity is also illustrated in Fig. 2: the investor's optimal withdrawal behavior is very similar under the lifecycle (dashed red line) and the SRNV (solid blue line) optimization approach. In particular, the investor optimally withdraws the guaranteed annual amount when the guarantee is in the money, that is when the account value is below the remaining guaranteed amount (in this figure, $G_t = 100$). If account value and guarantee base are close at time $t = 10$, the investor makes an excess withdrawal of around four times the guaranteed amount. This moves the guarantee deeper into the money, to a remaining guaranteed amount around $70,000, which ensures that the investor can withdraw the annual guaranteed amount in each of the remaining ten contract years.[12] Notably, optimal withdrawal behavior is

[12] In the baseline lifecycle model the policyholder has an additional motivation to withdraw nearly half the guaranteed amount if the account value A_t is approximately equal to the benefits base G_t. This pattern is present in most lifecycle settings (see Fig. 3), with the exception of the cases without bequest motive ($B = 0$) or when the beneficiary has a higher income ($I^{\text{ben.}}$). On the other hand, the "spike" occurs over an extended region under the (Ameriks et al., 2011) parameterization which entails a larger bequest motive. This suggests that the policyholder's motivation for the additional withdrawal in these circumstances might be due to her financial concern about her dependents/heirs. Ceteris paribus, if she is more concerned about her bequest, then she places less value on the GMWB rider, which is only valuable while she is alive. Therefore she might be willing so sacrifice some guarantee value (by making an excess withdrawal relative to the otherwise optimal amount) for the sake of having more money guaranteed to be available as a bequest (should she pass away in the next few years and considering that the VA account might lose in value due to the underlying equity investment, whereas she can invest the withdrawals risk-free in the outside account).

Table 2
Valuation results: baseline and sensitivity to preferences.

	[1]	[2]	[3]	[4]	[5]	[6]
	SRNV	Baseline	$W_0 = 500K$	$W_0 = 100K$	$I = 50K$	$\gamma = 3$
V_0 / U_0 $(\cdot 10^{-8})$	77,954	−31,512	−44,357	−68,166	−22,900	−0.2378
V_0^F	8772	8605	8528	8283	8556	8591
V_0^W	8447	8467	8471	8459	8464	8450
Σ_0	325	138	57	(176)	92	141
Δ Surplus		(187)	(268)	(501)	(232)	(183)
	[7]	[8]	[9]	[10]	[11]	[12]
	$\beta = 0.90$	$I^{\text{ben.}}$	$B = 0$	Ameriks	Inkmann	No O_t
U_0 $(\cdot 10^{-8})$	−11,207	−28,485	−22,329	−0.5516	$42{,}719 \cdot 10^8$	−80,658
V_0^F	8553	8711	8755	8085	8678	8092
V_0^W	8479	8461	8453	8238	8488	6390
Σ_0	74	250	302	(153)	190	1702
Δ Surplus	(251)	(75)	(23)	(478)	(134)	1377

Note: The table displays the valuation of the GMWB rider based on optimal withdrawal behavior under the specification at the top of each column [1] through [12]. Column [1] is based on the SRNV approach, while columns [2] through [12] are based on the lifecycle model, with column [2] as the baseline parameterization of Table 1. Columns [3] through [9] show sensitivity tests, varying one parameter at a time relative to the baseline assumptions (see Table 1). Columns [10] and [11] represent alternative model specifications, based on empirical calibrations, by Ameriks et al. (2011) and Inkmann et al. (2010), respectively. Column [12] assumes a lifecycle optimization without an outside investment account. V_0 is the risk-neutral after-tax value of the VA+GMWB product in the SRNV model, while U_0 represents the investor's time-0 overall utility from the lifecycle models. V_0^F is the time-0 expected value of the guarantee fee payments and V_0^W is the time-0 expected value of the GMWB payouts. Σ_0 is the insurer's surplus from the GMWB rider. These quantities are defined in Section 3. Lastly, Δ Surplus shows the deviation in surplus between the lifecycle model and the risk-neutral version, i.e. column [1].

Table 3
Valuation results: sensitivity to parameters that impact complete-market valuation.

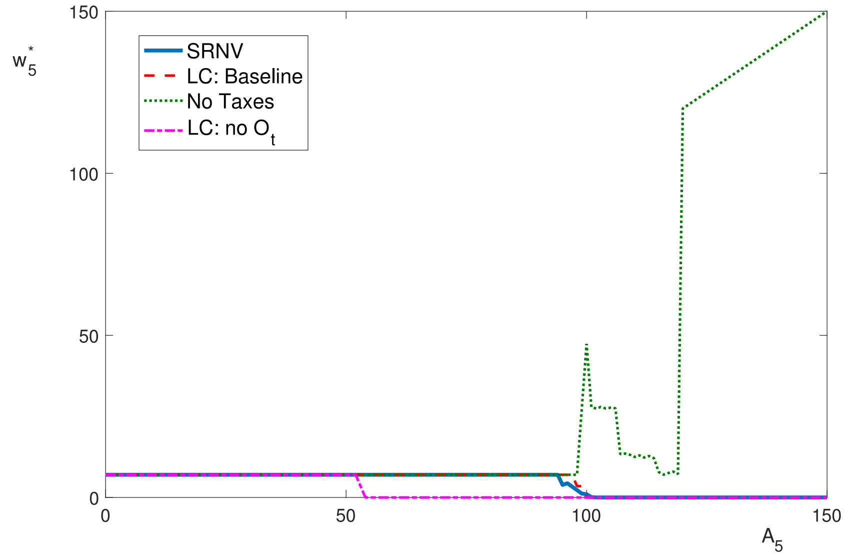
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
	Baseline	$x = 70$	$T = 25$	w/ GMDB	$\rho = 100\%$	$\sigma = 25\%$	$r = 4\%$	$\tau = 35\%$	$\kappa = 25\%$	$q_x = 0$	No Taxes
SRNV:											
V_0	77,954	76,670	79,082	78,661	84,559	84,559	78,235	72,386	80,212	78,836	98,177
V_0^F	8772	7877	10,512	8934	9184	9184	9379	8772	9051	9356	3542
V_0^W	8447	6768	8996	8359	11,067	11,067	5592	8447	8276	9618	6778
V_0^D				1334							
Σ_0	325	1109	1516	575	(1883)	(1883)	3787	325	775	(262)	(3236)
Lifecycle:											
U_0 $(\cdot 10^{-4})$	(3.1512)	(3.9730)	(3.6610)	(3.1493)	(3.1405)	(3.1378)	(2.3720)	(3.1616)	(3.1877)	(2.475)	(2.6830)
V_0^F	8605	7564	10,170	9112	9142	9154	9312	8607	8986	9276	3236
V_0^W	8467	6797	9029	8123	11,055	11,053	5620	8466	8276	9628	5687
V_0^D				1392							
Σ_0	138	767	1141	989	(1913)	(1899)	3692	141	710	(352)	(2,451)
Δ Surplus	(187)	(342)	(375)	415	(31)	(16)	(95)	(184)	(65)	(90)	786

Note: The table displays the valuation of the GMWB rider based on optimal withdrawal behavior under the specification at the top of each column, [1] through [11]. Within each column, results are shown for both the SRNV approach and the lifecycle model, and the bottom row (Δ Surplus) reflects the difference in the insurer's surplus between the two approaches. Specifications for the baseline parameterization (column [1]) and all sensitivity tests (columns [2], [3], and [5] through [10], varying one parameter at a time) are provided in Table 1. Column [4] considers the case where the VA also contains a return-of-premium death benefit guarantee (GMDB), and column [11] shows results under optimal behavior without tax considerations (i.e. $\tau = \kappa = 0$). V_0 is the risk-neutral after-tax value of the VA+GMWB product, while U_0 represents the investor's time-0 overall utility from the lifecycle model. V_0^F is the time-0 expected value of the guarantee fee payments and V_0^W is the time-0 expected value of the GMWB payouts. Σ_0 is the insurer's surplus from the GMWB rider. These quantities are defined in Section 3.

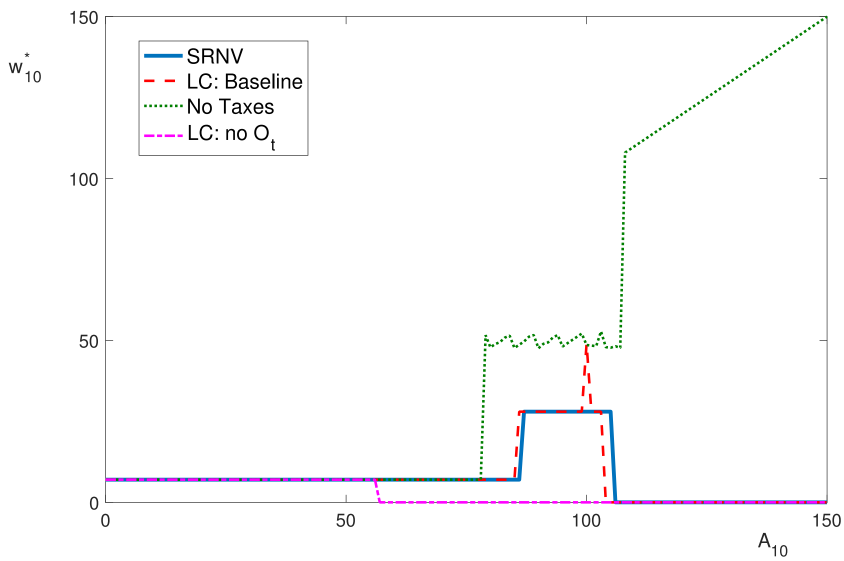
considerably different if the investor followed a *standard* risk-neutral valuation strategy, that is without taxes (dotted green line in Fig. 2). Most significantly, in this case the policyholder wants to surrender her policy when the guarantee moves "out of the money" (i.e., when $A_t > G_t$). This is particularly pronounced at time $t = 10$ when she no longer has to pay a surrender fee. With taxes, on the other hand, the surrender is costly because it triggers immediate tax payments and forces the policyholder to give up potential future tax benefits. These insights are similar to those of Moenig and Bauer (2016), who study optimal withdrawal behavior for a VA+GMWB policy in a *non-qualified* account.

4.2. Sensitivity to preferences and personal financial characteristics

Columns [3] to [9] of Table 2 show the GMWB valuation results under the lifecycle model for individual changes to the investor's preferences, income, and wealth. In all sensitivity cases, the results are quite close to those of the SRNV model (column



(a) $t = 5$.



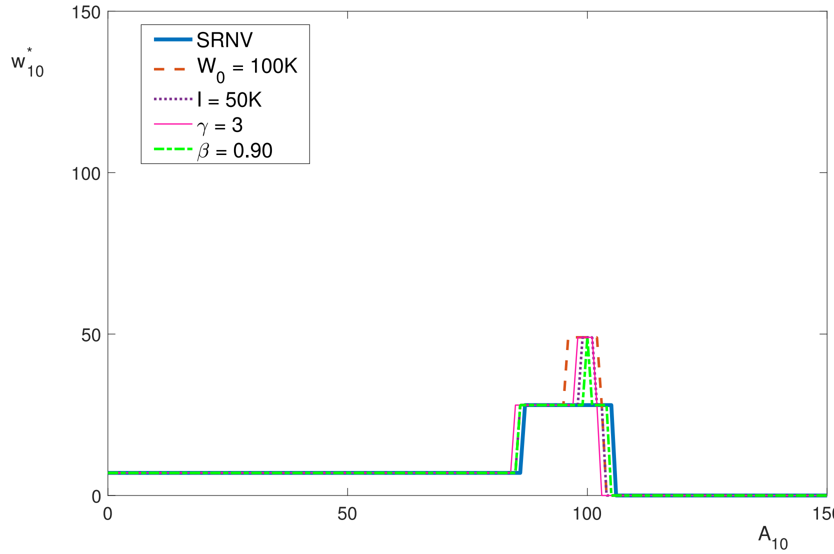
(b) $t = 10$.

Fig. 2. Optimal withdrawal strategy (baseline). Note: All units are in \$1,000. The figure shows the investor's optimal withdrawal strategy w_t^*—as a function of the VA account value A_t—at times $t = 5$ and $t = 10$, with remaining guaranteed amount and (where applicable) outside account value at the initial level, that is: $G_t = 100$, $O_t = 1{,}000$. The different lines represent different optimization models: the SRNV approach; the lifecycle model; a standard risk-neutral valuation without taxes; and a lifecycle model that does not include an outside account. All parameters are set as in the baseline specifications of Table 1.

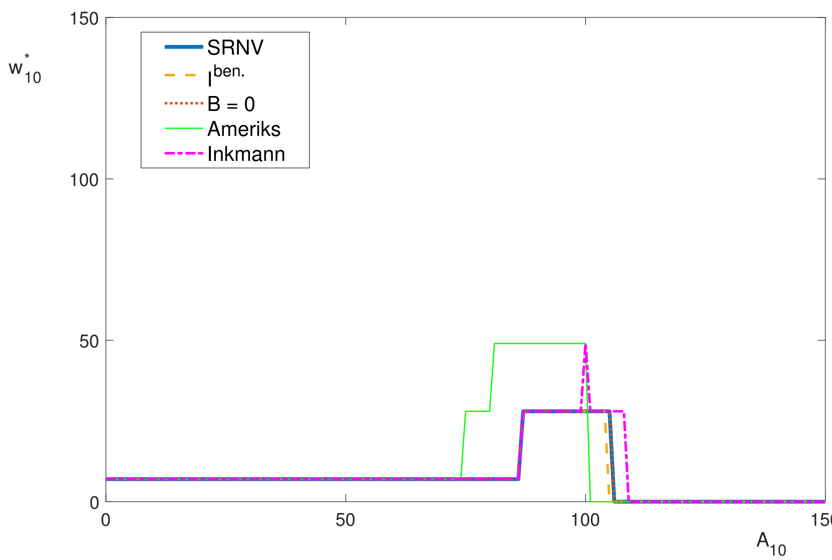
[1]) and the lifecycle baseline case (column [2]), with deviations in the provider's surplus generally around or below \$250, i.e. around four basis points in the guarantee fee.[13] This is a *minor* discrepancy, both in absolute terms and relative to the impact of other model parameters (see below). And even when the investor's outside wealth is reduced substantially so that a large part of her investments are tied up in the VA (see column [4] in Table 2 and the dashed orange line in Fig. 3(a)), her optimal behavior is still reasonably close to the risk-neutral case. Fig. 3 confirms these insights by demonstrating that it is indeed her optimal behavior that resembles the SRNV approach, and not just the guarantee values and surplus to the insurer. These findings therefore suggest that the market around VA+GMWB products is largely complete for a typical investor.

This is a consequence of the availability of an outside investment account. Column [12] of Table 2 illustrates the case where

[13] Interestingly, the lifecycle models tend to *reduce* the insurer's surplus relative to the SRNV approach. While this does not have to be the case (in theory or practice), it is certainly a plausible outcome. The presence of taxes creates market frictions that incentivize the policyholder to deviate from value-maximizing/surplus-minimizing behavior (which is displayed graphically in the dotted green lines of Fig. 2 and numerically in column [11] of Table 3). Adding *incompleteness* to the market means that the policyholder now has to also worry about smoothing her consumption (among other things), which takes away from her incentive to minimize tax payments. As a result, the policyholder may end up behaving in a way that is (slightly) less favorable for the insurer, as the above results suggest in the context of this study.



(a) Wealth, income & preferences.



(b) Bequests & alternative parameter settings.

Fig. 3. Optimal withdrawal strategy (sensitivity). Note: All units are in \$1,000. The figure shows the investor's optimal withdrawal strategy w_t^*—as a function of the VA account value A_t—at time $t = 10$, with remaining guaranteed amount and outside account value at the initial level, that is: $G_t = 100$, $O_t = 1{,}000$ (except for the dashed orange line in panel (a) where $O_t = 100$). The solid thick blue line reflects optimization under the SRNV approach, while the other lines are sensitivities of the baseline lifecycle model (with all parameters other than the one indicated being identical to the baseline specifications of Table 1). More details on the sensitivities can be found in Section 4.2.

the investor does not have access to such an account and is restricted to consuming her annual income plus any potential withdrawals. This changes her withdrawal incentives dramatically – see also the dash-dotted pink line in Fig. 2 – and has major implications for the valuation of the guarantee. In contrast, if the investor has access to an outside account, she evidently strives to maximize the (after-tax) value of her VA+GMWB policy and uses the outside account (rather than withdrawals from the VA) to smooth consumption over time and to align her overall equity exposure with her risk appetite. In particular, in the cases plotted in Fig. 2, she is less likely to make a withdrawal even when it may be financially optimal to do so because she does not currently need additional consumption (at the expense of future consumption and bequests) and the VA policy presents her only opportunity to save. Therefore, if one believes that the investor has access to outside savings and investment opportunities, it is crucial to account for them when modeling VA policyholder behavior in an incomplete-market setting.

Alternative preference specifications

As an additional robustness check, I implement two alternative preference specifications based on lifecycle models that have been *empirically calibrated* to individual retirement savings and bequest decisions.

Ameriks et al. (2011) assume largely the same setup as I do (EUT-CRRA) and rely on the same implementation procedure, except that the investor's bequest function is defined as

$$u_B(b_t) = \frac{B}{1-\gamma}\left(\psi + \frac{b_t}{B}\right)^{1-\gamma},$$

where B still indicates the strength of the investor's bequest motive, and the additional parameter ψ can be interpreted as the degree to which bequests are luxury goods. Using a baseline risk aversion coefficient of $\gamma = 3$, their calibration – accounting for both consumption data and survey responses – yields best-estimate model parameters $B = 47.6$ and $\psi = \$7280$.

Inkmann et al. (2010) study demand for the voluntary annuity market in the U.K. with a lifecycle model based on Epstein and Zin (1991) preferences. In particular, the investor's utility function is defined recursively as:

$$U_t(y_t) = \max_{c_t, w_t, v_t} \left\{ (1-\beta) c_t^{1-\frac{1}{\psi}} + \beta \left(\mathbb{E}_t^{\mathbb{P}} \left[q_{x+t} (b_{t+1})^{1-\gamma} B \right.\right.\right.$$
$$\left.\left.\left. + (1-q_{x+t}) U_{t+1}(y_{t+1})^{1-\gamma} \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\frac{1}{\psi}}} . \tag{23}$$

Here, ψ denotes the elasticity of intertemporal substitution. In the investor's optimization problem, Eq. (23) replaces Eq. (12). Moreover, the terminal condition (13) is replaced by

$$U_T(A_{T-}, G_T, O_{T-}) = \alpha_T \cdot \left(c_T^{\text{net}} + I \right) , \tag{24}$$

where α_t (for $t = T, T+1, \ldots, \omega - 1$) is defined recursively as

$$\alpha_t = \left\{ 1 - \beta + \beta (1 - q_{x+t})^{\frac{1-\frac{1}{\psi}}{1-\gamma}} (\alpha_{t+1})^{1-\frac{1}{\psi}} \right\}^{\frac{1}{1-\frac{1}{\psi}}} ,$$

ω denotes the investor's maximum attainable age (here: 120), and the terminal condition is $\alpha_\omega = 0$. The proof of Eq. (24) can be found in the Appendix. When calibrating the model to microeconomic data for U.K. stockholders, Inkmann et al. (2010) elicit the following parameters: $\beta = 0.99$, $\gamma = 5$, $\psi = 0.4$, and $B = 6.0$.

The valuation results under the Ameriks et al. (2011) and the Inkmann et al. (2010) specifications are shown in columns [10] and [11] of Table 2, respectively. The investor's optimal withdrawal strategies are graphed (for a representative case) in Fig. 3(b). These numbers and illustrations confirm the above insights that optimal withdrawal behavior – and thus the value of the GMWB rider – does not depend on the investor's unobservable preferences and can instead be modeled in a risk-neutral framework (albeit accounting for the product's tax benefits, when appropriate).

4.3. Sensitivity to contract features, financial market parameters, and taxes

Next, I consider variations to the parameters that may also impact the (subjective) risk-neutral valuation method. This includes VA contract specifications and financial market parameters, but also the investor's marginal tax rates. Results are displayed in columns [2] to [10] of Table 3, with the SRNV approach at the top and the corresponding lifecycle model – using the investor's baseline preference specifications from Table 1 – at the bottom.

I find that in all cases, the lifecycle results are quite close to their risk-neutral equivalent, with the difference in surplus amounting to at most \$415 out of a \$100,000 investment, which corresponds to around 7 basis points in the guarantee fee. Again, this is relatively minor, especially considering the sensitivity of the GMWB value and surplus to other model specifications. For instance, comparing the GMWB surplus of columns [4] and [5] to that of column [1] suggests that the \$415 corresponds approximately to the impact of a 1 percentage point increase in the volatility of the market asset, and a 12 basis point-increase in the risk-free rate. Moreover, the \$415 are also much less than the impact of purchase age and time to maturity (comparing the surplus from Columns [2] and [3] of Table 3 to Column [1]), and in practice insurers opt not to charge different fee rates based on these characteristics.

Qualitatively, the impact of the sensitivities on the GMWB value and surplus is consistent with intuition: the GMWB is less valuable if the investor is older (Column [2]) and thus less likely to survive long enough to take full advantage of the guaranteed withdrawals. With a longer time to maturity (Column [3]), the GMWB rider is more valuable, but the insurer also collects more fees. The same is true if the investor is guaranteed to survive to maturity of the VA (Column [10]). If the VA has more equity exposure (Column [5]), the mutual fund returns are more volatile which increases the value of the guarantee. As expected, this has a virtually identical impact as a corresponding direct increase in the volatility of the market asset (Column [6]). Under a larger risk-free rate (Column [7]), the VA account value grows faster on average (under measure $\mathbb{Q}$), which makes the guarantee less likely to be "in the money"; therefore, the investor makes fewer withdrawals, which allows the insurer to collect additional fees.

Next, I consider the impact of the investor's marginal tax rates. I find that a higher tax rate κ on earnings from the outside account (Column [9]) makes the VA more attractive, which marginally discourages withdrawals, leading to a lower guarantee value and more VA fee payments. However, the level of the marginal *income* tax rate τ (Column [8]) does not impact optimal withdrawal behavior: since all withdrawals from the VA are taxed *fully* at rate τ, the tax rate does not affect the optimal timing of the withdrawal. This is in contrast to the finding of Moenig and Bauer (2016): in their study the VA is placed in a non-qualified account, where withdrawals are only partially taxed, and in that case there can be an advantage to timing the withdrawal in order to reduce tax obligations. For reference, Column [11] of Table 3 presents valuation results if the investor's behavior is driven by value or utility maximization without considering the product's tax implications. As discussed above – based on the insight provided by Fig. 2 – the investor's withdrawal behavior is substantially different without taxes, resulting in frequent policy surrenders when the guarantee is out of the money. As a result, the GMWB rider does not appear to be profitable, even at relatively high fee rates, confirming the insights of Milevsky and Salisbury (2006), Chen et al. (2008), and Dai et al. (2008).

The sensitivity analysis also considers the case where a return-of-premium death benefit guarantee (GMDB) is included in the VA policy, as is generally the case in practice. The remaining guaranteed withdrawal amount G_t then also serves as the guaranteed death benefit. Therefore, in case of the policyholder's premature death in policy year t, the beneficiaries receive the larger of the VA account value and the remaining guaranteed amount, i.e. $DB_t = \max\{A_{t-}, G_t\}$, with the insurer paying

$$X_t^D := (G_t - A_{t-})^+ .$$

I let V_t^D denote the time-t risk-neutral expected present value of all forthcoming death benefit guarantee payments and update this function as follows:

$$V_t^D(y_t) = e^{-r} \left(q_{x+t} \, \mathbb{E}_t^{\mathbb{Q}} \left[X_{t+1}^D \right] + (1 - q_{x+t}) \, \mathbb{E}_t^{\mathbb{Q}} \left[V_{t+1}^D(y_{t+1}) \right] \right) ,$$

with terminal condition

$$V_T^D(y_T) = 0$$

and time-0 risk-neutral expected value

$$V_0^D = e^{-r} \left(q_x \, \mathbb{E}_0^{\mathbb{Q}} \left[X_1^D \right] + (1 - q_x) \, \mathbb{E}_0^{\mathbb{Q}} \left[V_1^D(y_1) \right] \right) .$$

The valuation results in the presence of the GMDB rider are shown in column [4] of Table 3, and are consistent with the earlier insights.

4.4. Revisiting market incompleteness

The above numerical results show that the investor's optimal withdrawal behavior in a lifecycle framework is reasonably well approximated by the optimal behavior in the corresponding risk-neutral environment (with taxes). That is, access to the stock market appears to largely – but not fully – complete the market around VA guarantees.

As discussed in Section 2, a VA investor faces two sources of risk: financial and mortality. Since she is risk averse, she wants to balance the respective utility she gains in case of a good economy and a bad economy, but also in case of her death and of her survival. In our model, the outside investment account allows the investor to mitigate her *financial* risk exposure, but the model does not offer her a corresponding mechanism available to mitigate her *mortality* risk. In fact, the only mortality-contingent financial instrument at the investor's disposal is the GMWB rider, which has value only in case of her survival. Therefore, if her value-maximizing withdrawal behavior creates an imbalance between the utility in the "death" state versus the "alive" state – relative to her risk preferences and the states' respective likelihoods – then her optimal withdrawal strategy in the lifecycle model diverts from that behavior in order to better align her utility between the "alive" and "death" states with her preferences.[14]

These insights are reflected in my numerical results: for instance, I find that the discrepancy between the two approaches is (slightly) larger when mortality rates are higher, that is for older investors and if the VA maturity is longer (see columns [2] and [3] in Table 3). The opposite occurs if the investor were guaranteed to survive to the maturity of the VA (see column [10] of Table 3). This reflects the situation where the investor only faces financial risk, and therefore the difference in surplus of \$90 ($<$ 0.1% of the VA policy's face value) can be viewed as the impact of the residual market incompleteness with respect to financial risk (along with possibly a small estimation error from the numerical implementation). This contributes approximately half to the \$187 in surplus difference in the baseline case, with the other half being attributable to the incompleteness of the market for idiosyncratic mortality risk. Similarly, if the investor does not have a bequest motive (see column [9] in Table 2), the impact of the mortality-related incompleteness vanishes and the lifecycle results coincide almost perfectly with the SRNV model. In contrast, the discrepancy in GMWB surplus is higher under the model specifications of Ameriks et al. (2011), who estimated a relatively high bequest motive (see column [10] in Table 2).

Since the model in this study does not grant the investor access to traditional life insurance and life annuity products, the optimal withdrawal behavior under the lifecycle model would be even closer to the risk-neutral case if the investor had access to those markets. Therefore, market incompleteness may be even less relevant for optimal VA policyholder behavior in practice than what the above results suggest.

5. Conclusion

I studied the determinants of an investor's optimal withdrawal strategy for a GMWB rider in a VA contract and found that optimal behavior is driven primarily by value maximization, though accounting for the product's tax benefits relative to more traditional investment accounts, where appropriate.

This is good news for VA providers (as well as investors and financial advisors), for two reasons. First, it is difficult for actuaries to properly parameterize a lifecycle model and to conduct appropriate sensitivity testing, since preference parameters are unobservable, difficult to elicit on a VA application form, and may vary considerably across potential investors. If behavior depended critically on those preferences, it would be difficult to price and especially to hedge the rider. Second, the numerical implementation of the lifecycle model requires a large amount of computing power, even for a simple guarantee structure and in a Black–Scholes financial world. A risk-neutral valuation model, on the other hand, can easily (from a computational perspective, anyway) handle more complicated GMWB structures, alternative tax environments, or financial models. I refer the interested reader to Moenig and Bauer (2016) for an implementation of real-world VA+GMWB products in the SRNV framework.

The findings in this study reveal a (small level of) market incompleteness with respect to the investor's idiosyncratic mortality risk. Access to the traditional life insurance and life annuity market would get the investor closer to a complete market. On the other hand, VA guarantees that are even more dependent on the investor's longevity – such as GLWBs, which not only deal with higher ages but also offer unlimited guaranteed withdrawals – should also be more susceptible to this incompleteness.[15] This may be a worthwhile direction for future research.

This study documents the (lack of) market incompleteness faced by VA investors. On the other hand, VA providers also face a potentially incomplete market since the assets used to hedge the VA liabilities differ from the mutual fund that the guarantee is written on, since the latter is generally illiquid and cannot be shorted. This form of *basis risk* can greatly impact the providers' hedging effectiveness (Sun et al., 2016) and has thus far received relatively little attention in the literature (see e.g. Delong, 2014; Trottier et al., 2018).

Lastly, I would like to see the theoretical predictions of this paper as well as the broader literature on optimal policyholder behavior in VAs complemented by empirical studies. These have been rather scarce, primarily due to the proprietary nature of policy-level data, with (Milevsky and Kyrychenko, 2008) and Knoller et al. (2016) being notable exceptions.

Acknowledgments

I am particularly grateful to Daniel Bauer for his guidance and feedback, as well as to Glenn Harrison, Ajay Subramanian, Eric Ulm, and Yongsheng Xu for their valuable input. I also thank four anonymous reviewers and participants at the 2011 ARIA Annual Meeting; the 46th Actuarial Research Conference; the 12th Symposium on Finance, Banking, and Insurance; the 2013 PARTY; the 2014 International Congress of Actuaries; as well as seminar participants at Georgia State University, Ulm University, the University of Manitoba, Manhattan College, the University of St. Thomas, the University of Wisconsin-Madison, the Twin Cities Actuarial Club, and Temple University for helpful comments. All remaining errors are mine.

Funding

This work was supported by the Society of Actuaries under a CAE research grant at Georgia State University.

14 This relates to the finding of Steinorth and Mitchell (2015) that optimal withdrawal behavior for GLWBs, also in a lifecycle context, is influenced by the longevity risk protection embedded in the guarantee.

15 See for example (Xiong et al., 2010), Blanchett (2012), and Pfau (2013) for related studies.

Appendix. Proof of Eq. (24)

The objective is to prove that the investor's time-T valuation of her post-VA investment strategy under Epstein-Zin preferences is specified by Eq. (24). I do so using backwards induction.

The terminal condition is $U_\omega = 0$, since there is no utility to be gained after death. This satisfies Eq. (24) with $\alpha_\omega = 0$. For the induction step, suppose that for $t \in \{T, \ldots, \omega - 1\}$ it is true that $U_{t+1} = \alpha_{t+1} c$. I now show that this implies that $U_t = \alpha_t c$.

Recall that upon maturity of the VA, the investor fully annuitizes her wealth, thus receiving guaranteed annual income $c := c_T^{\text{net}} + I$ for the remainder of her life. For $t = T, T+1, \ldots, \omega - 1$, her time-$t$ Epstein-Zin utility is therefore given by

$$\begin{aligned} U_t &= \left\{(1-\beta)c^{1-\frac{1}{\psi}} + \beta \left(\mathbb{E}_t^{\mathbb{P}}\left[q_{x+t} B \cdot 0 + (1-q_{x+t})(U_{t+1})^{1-\gamma}\right]\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right\}^{\frac{1}{1-\frac{1}{\psi}}} \\ &= \left\{(1-\beta)c^{1-\frac{1}{\psi}} + \beta \left((1-q_{x+t})(U_{t+1})^{1-\gamma}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right\}^{\frac{1}{1-\frac{1}{\psi}}} \\ &= \left\{(1-\beta)c^{1-\frac{1}{\psi}} + \beta (1-q_{x+t})^{\frac{1-\frac{1}{\psi}}{1-\gamma}} (\alpha_{t+1} c)^{1-\frac{1}{\psi}}\right\}^{\frac{1}{1-\frac{1}{\psi}}} \\ &= \left\{(1-\beta) + \beta (1-q_{x+t})^{\frac{1-\frac{1}{\psi}}{1-\gamma}} (\alpha_{t+1})^{1-\frac{1}{\psi}}\right\}^{\frac{1}{1-\frac{1}{\psi}}} c \\ &= \alpha_t\, c\,. \end{aligned}$$

This completes the induction step and thus the proof. □

References

Ameriks, J., Caplin, A., Laufer, S., Van Nieuwerburgh, S., 2011. The joy of giving or assisted living? Using strategic surveys to separate public care aversion from bequest motives. J. Finance 66 (2), 519–561.

Bacinello, A.R., Millossovich, P., Montealegre, A., 2016. The valuation of GMWB variable annuities under alternative fund distributions and policyholder behaviours. Scand. Actuar. J. 2016 (5), 446–465.

Bauer, D., Gao, J., Moenig, T., Ulm, E.R., Zhu, N., 2017. Policyholder exercise behavior in life insurance: The state of affairs. N. Am. Actuar. J. 21 (4), 485–501.

Bauer, D., Kling, A., Russ, J., 2008. A universal pricing framework for guaranteed minimum benefits in variable annuities. ASTIN Bull.:J. IAA 38 (2), 621–651.

Blanchet-Scalliet, C., Chevalier, E., Kharroubi, I., Lim, T., 2015. Max–min optimization problem for variable annuities pricing. Int. J. Theor. Appl. Finance 18 (08), 1550053.

Blanchett, D.M., 2012. Optimal portfolio allocations with GMWB annuities. J. Financ. Plan. 25 (10), 46–56.

Brown, J.R., Mitchell, O.S., Poterba, J.M., 2001. The role of real annuities and indexed bonds in an individual accounts retirement program. In: Risk Aspects of Investment-Based Social Security Reform. University of Chicago Press, pp. 321–370.

Brown, J.R., Poterba, J.M., 2006. Household ownership of variable annuities. Tax Policy Econ. 20, 163–191.

Campbell, J., Chan, M., Li, K., Lombardi, L., Lombardi, L., Purushotham, M., Rao, A., 2014. Modeling of Policyholder Behavior for Life Insurance and Annuity Products. Technical Report, Schaumburg: Society of Actuaries, Available online: https://www.soa.org/globalassets/assets/files/research/projects/research-2014-modeling-policy.pdf.

Carpenter, J.N., 1998. The exercise and valuation of executive stock options. J. Financ. Econ. 48 (2), 127–158.

Charupat, N., Milevsky, M.A., 2002. Optimal asset allocation in life annuities: a note. Insurance Math. Econom. 30 (2), 199–209.

Chen, Z., Vetzal, K., Forsyth, P.A., 2008. The effect of modelling parameters on the value of GMWB guarantees. Insurance Math. Econom. 43 (1), 165–173.

Chevalier, E., Lim, T., Romo Romero, R., 2016. Indifference fee rate for variable annuities. Appl. Math. Finance 23 (4), 278–308.

Costabile, M., Massabó, I., Russo, E., 2020. Evaluating variable annuities with GMWB when exogenous factors influence the policy-holder's withdrawals. Eur. J. Financ. 26 (2–3), 238–257.

Cox, J.C., Huang, C.-f., 1989. Optimal consumption and portfolio policies when asset prices follow a diffusion process. J. Econ. Theory 49 (1), 33–83.

Dai, M., Kuen Kwok, Y., Zong, J., 2008. Guaranteed minimum withdrawal benefit in variable annuities. Math. Financ.:Int. J. Math. Stat. Financ. Econ. 18 (4), 595–611.

Delong, Ł., 2014. Pricing and hedging of variable annuities with state-dependent fees. Insurance Math. Econom. 58, 24–33.

Donnelly, R., Jaimungal, S., Rubisov, D.H., 2014. Valuing guaranteed withdrawal benefits with stochastic interest rates and volatility. Quant. Finance 14 (2), 369–382.

Epstein, L.G., Zin, S.E., 1991. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. J. Polit. Econ. 99 (2), 263–286.

Forsyth, P., Vetzal, K., 2014. An optimal stochastic control framework for determining the cost of hedging of variable annuities. J. Econom. Dynam. Control 44, 29–53.

Gao, J., Ulm, E.R., 2012. Optimal consumption and allocation in variable annuities with guaranteed minimum death benefits. Insurance Math. Econom. 51 (3), 586–598.

Gao, J., Ulm, E.R., 2015. Optimal allocation and consumption with guaranteed minimum death benefits, external income and term life insurance. Insurance Math. Econom. 61, 87–98.

Goudenège, L., Molent, A., Zanette, A., 2019. The impact of taxation on GMWB contract in a stochastic interest rate framework. Working paper. Università Politecnica delle Marche.

Gudkov, N., Ignatieva, K., Ziveyi, J., 2019. Pricing of guaranteed minimum withdrawal benefits in variable annuities under stochastic volatility, stochastic interest rates and stochastic mortality via the componentwise splitting method. Quant. Finance 19 (3), 501–518.

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 Math. Econom. 63, 91–107.

Horneff, W.J., Maurer, R.H., Mitchell, O.S., Stamos, M.Z., 2010. Variable payout annuities and dynamic portfolio choice in retirement. J. Pension Econ. Financ. 9 (2), 163–183.

Huang, H., Milevsky, M.A., Salisbury, T.S., 2014. Optimal initiation of a GLWB in a variable annuity: No arbitrage approach. Insurance Math. Econom. 56, 102–111.

Ignatieva, K., Song, A., Ziveyi, J., 2018. Fourier space time-stepping algorithm for valuing guaranteed minimum withdrawal benefits in variable annuities under regime-switching and stochastic mortality. ASTIN Bull.: J. IAA 48 (1), 139–169.

Inkmann, J., Lopes, P., Michaelides, A., 2010. How deep is the annuity market participation puzzle? Rev. Financ. Stud. 24 (1), 279–319.

Knoller, C., Kraut, G., Schoenmaekers, P., 2016. On the propensity to surrender a variable annuity contract: an empirical analysis of dynamic policyholder behavior. J. Risk Insurance 83 (4), 979–1006.

Lin, X.S., Tan, K.S., Yang, H., 2009. Pricing annuity guarantees under a regime-switching model. N. Am. Actuar. J. 13 (3), 316–332.

Lockwood, L.M., 2012. Bequest motives and the annuity puzzle. Rev. Econ. Dyn. 15 (2), 226–243.

MacKay, A., Augustyniak, M., Bernard, C., Hardy, M.R., 2017. Risk management of policyholder behavior in equity-linked life insurance. J. Risk Insurance 84 (2), 661–690.

Mahayni, A., Schneider, J.C., 2012. Variable annuities and the option to seek risk: Why should you diversify? J. Bank. Financ. 36 (9), 2417–2428.

Marshall, C., Hardy, M., Saunders, D., 2010. Valuation of a guaranteed minimum income benefit. N. Am. Actuar. J. 14 (1), 38–58.

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

Milevsky, M.A., Kyrychenko, V., 2008. Portfolio choice with puts: Evidence from variable annuities. Financ. Anal. J. 64 (3), 80–95.

Milevsky, M.A., Posner, S.E., 2001. The titanic option: valuation of the guaranteed minimum death benefit in variable annuities and mutual funds. J. Risk Insurance 68 (1), 93–128.

Milevsky, M.A., Salisbury, T.S., 2006. Financial valuation of guaranteed minimum withdrawal benefits. Insurance Math. Econom. 38 (1), 21–38.

Moenig, T., Bauer, D., 2016. Revisiting the risk-neutral approach to optimal policyholder behavior: A study of withdrawal guarantees in variable annuities. Rev. Financ. 20 (2), 759–794.

Moenig, T., Zhu, N., 2018. Lapse-and-reentry in variable annuities. J. Risk Insurance 85 (4), 911–938.

Peng, J., Leung, K.S., Kwok, Y.K., 2012. Pricing guaranteed minimum withdrawal benefits under stochastic interest rates. Quant. Finance 12 (6), 933–941.

Pfau, W.D., 2013. A broader framework for determining an efficient frontier for retirement income. J. Financ. Plan. 26 (2), 44–51.

Piscopo, G., Rüede, P., 2018. What if variable annuity policyholders with guaranteed lifelong withdrawal benefit were rational? J. Risk Insurance 85 (1), 203–217.

Sell, S., 2006. Guaranteed living benefits experience study. Soc. Actuaries Prod. Dev. News 64, 9–11.

Shevchenko, P.V., Luo, X., 2017. Valuation of variable annuities with guaranteed minimum withdrawal benefit under stochastic interest rate. Insurance Math. Econom. 76, 104–117.

Steinorth, P., Mitchell, O.S., 2015. Valuing variable annuities with guaranteed minimum lifetime withdrawal benefits. Insurance Math. Econom. 64, 246–258.

Sun, P., Kelkar, R., Dai, J., Huang, V., 2016. How Effective Is Variable Annuity Guarantee Hedging? Technical Report, Milliman Research Report, Available online: http://us.milliman.com/uploadedFiles/insight/2016/variable-annuity-guarantee-hedging.pdf.

Sun, J., Shevchenko, P.V., Fung, M.C., 2018. The impact of management fees on the pricing of variable annuity guarantees. Risks 6 (3), 103.

Trottier, D.-A., Godin, F., Hamel, E., 2018. Local hedging of variable annuities in the presence of basis risk. ASTIN Bull.:J. IAA 48 (2), 611–646.

Ulm, E.R., 2006. The effect of the real option to transfer on the value of guaranteed minimum death benefits. J. Risk Insurance 73 (1), 43–69.

Ulm, E.R., 2020. The effect of retirement taxation rules on the value of guaranteed lifetime withdrawal benefits. Ann. Actuar. Sci. 14 (1), 83–92.

Xiong, J.X., Idzorek, T., Chen, P., 2010. Allocation to deferred variable annuities with GMWB for life. J. Financ. Plan. 23 (2), 42–50.