



Financial valuation of guaranteed minimum withdrawal benefits

Moshe A. Milevsky^{a,*}, Thomas S. Salisbury^b

^a *Schulich School of Business, York University, Toronto, Ontario, Canada*

^b *Fields Institute for Research in Mathematical Science and York University Department of Mathematics and Statistics, Toronto, Ontario, Canada*

Received January 2005; received in revised form June 2005; accepted 30 June 2005

Abstract

Financial valuation OF GMWBs: We develop a variety of methods for assessing the cost and value of a very popular ‘rider’ available to North American investors on variable annuity (VA) policies called a Guaranteed Minimum Withdrawal Benefit (GMWB). The GMWB promises to return the entire initial investment, albeit spread over an extended period of time, regardless of subsequent market performance. First, we take a *static* approach that assumes individuals behave passively and holds the product to maturity. We show how the product can be decomposed into a Quanto Asian Put plus a generic term-certain annuity. At the other extreme of consumer behavior, the *dynamic* approach leads to an optimal stopping problem akin to pricing an American put option, albeit complicated by the non-traditional payment structure. Our main result is that the No Arbitrage hedging cost of a GMWB ranges from 73 to 160 basis points of assets. In contrast, most products in the market only charge 30–45 basis points. Although we suggest a number of behavioral reasons for the apparent under-pricing of this feature in a typically overpriced VA market, we conclude by arguing that current pricing is not sustainable and that GMWB fees will eventually have to increase or product design will have to change in order to avoid blatant arbitrage opportunities.

© 2005 Elsevier B.V. All rights reserved.

JEL classification: G22

Insurance Branch Code: IM10; IE43; IB13

Keywords: Portfolio insurance; Pensions; Annuities; Put options; Asian options

“They have stumbled onto a ‘killer app’ for the financial needs of today’s boomers, It’s called a GMWB. The deal is that for a half-percentage point a year, you can invest with a guarantee that your entire principal will be returned to you, provided that the principal is

not withdrawn at a rate greater than 7% annually. . .”
Washington Post, May 23, 2004

“The risks of variable annuities have come home to roost for insurers. . . To make matters worse, rating agencies, accountants and regulators never adequately publicized the risk of Variable Annuities for insurers balance sheets . . . their losses were compounded even further because the did such a poor job hedging their

* Corresponding author. Tel.: +1 416 736 2100x66014; fax: +1 416 736 5487.

E-mail address: milevsky@yorku.ca (M.A. Milevsky).

risk in many cases...” Financial Times, August 24, 2004

1. Introduction and motivation

Within the North American financial industry, Variable Annuities (VAs) are close cousins of mutual funds—which bundle individual securities, such as stocks and bonds, into diversified units or trusts—but they are formally classified as an insurance policy in addition to being registered as a security. They provide tax sheltered growth and embed a number of put-like derivatives that provide guarantees on the account value at the time of death.¹ As of early 2005 there is more² than \$ 1 trillion USD invested in Variable Annuities which, for all intents and purposes, can be viewed as tax efficient mutual funds with European-style puts that mature at death.

A recent innovation in this market is the above-referenced GMWB ‘rider’. In contrast to all the other bells and whistles³ it contains absolutely no life insurance component and is thus well within the domain of analysis of financial economics. The GMWB promises a minimal payout level from an initial investment capital—akin to a systematic withdrawal plan (SWiP)—regardless of the performance of the underlying policy. Typically, the policyholder might be guaranteed the ability to withdraw \$ 7 per annum per \$ 100 of initial investment until the original \$ 100 has been fully exhausted. According to recent estimates,⁴ close to 75% of variable annuity buyers selected this rider during the year 2004.

Thus, if the market performed poorly—and especially in the early years when the VA is purchased—the investor would be guaranteed a minimal *time-weighted average* return. Like all insurance riders, and in contrast to standard exchange traded options, most insur-

ance companies charge for this downside protection by deducting an ongoing *fraction of assets* as opposed to an up-front fee. These unique features differentiate the pricing of this derivative security from the standard Black and Scholes (1973) approach where the option premiums are paid up-front and in advance. This fact introduces subtle hedging issues which we will address in the body of the paper.

The GMWB is not a trivial wrinkle in a small market. It is being offered by a growing number of insurance companies as a substitute to payout annuities. Table 1 displays the names of companies offering this feature in the U.S. together with the GMWB rate and the insurance charge.

In light of the growing importance of this market, the aim and contribution of this paper is to (a) use No Arbitrage techniques to analyze insurance features in Variable Annuities, an area that has not received nearly as much academic attention as the mutual fund market, despite its \$ 1 trillion size; and (b) provide two extreme approaches to analyzing, valuing and managing the risks of a GMWB that are *predicated on the degree of investor rationality*. We differentiate our approach from a traditional insurance-oriented methodology that uses the law of large numbers—under the physical/statistical measure—to compute an expected loss. Our work is in the financial valuation tradition of Brennan and Schwartz (1976) and Boyle and Schwartz (1977) where the Black–Scholes methodology is extended to insurance contracts, except that in our analysis the payout is not necessarily tied to death.

And, at the risk of placing the cart before the horse, our main conclusion is that the GMWB that are most popular in the market seem to be *underpriced* for consumers who know how to utilize these options in a rational manner. This is in contrast to the wide-spread perception—see Clements (2004) for example—that the guarantees embedded within variable annuities are all overpriced.

Of course, it is now well established that individual retail investors do not adhere to the basic tenets of economic optimality. For example, Benartzi and Thaler (2001) document that investors in 401(k) pension plans use simple $1/n$ heuristics to select mutual funds as opposed to using a mean-variance approach to diversify their portfolio. Other papers in the behavioral finance literature provide evidence that consumers cannot be relied upon to optimally exercise financial

¹ See Milevsky and Posner (2001) or Brown and Poterba (2004) for a discussion of the many features available on VAs and the possible reasons underlying the demand for variable annuities. There is a clear relation to income taxes, and low-income households would find little value in the tax deferral which converts lightly-taxed capital gains into a (deferred) ordinary interest classification.

² Source: National Association of Variable Annuities.

³ So coined by Moody’s Investor Services in a November 2001 research report entitled: *The Growing Bells and Whistles on Variable Annuities*.

⁴ Source: National Underwriter Magazine, 2005.

Table 1
Who offers guaranteed minimum withdrawal benefits (GMWBs)?

Company name	No. of policies	Guaranteed rate (%)	Fee payment	Insurance fee
Lincoln National	10	7	% of G.B.	45 b.p.
Jackson National	2	7	% of A.V.	35 b.p.
Jefferson National	2	7	% of A.V.	35 b.p.
Pacific Life	5	7	% of A.V.	40 b.p.
Transamerica	2	7 or 5	% of G.B.	75 b.p.
ING Golden Life	4	7	% of G.B.	65 b.p.
Manulife Financial	4	5	% of G.B.	30 b.p.
Sun Life (U.S.)	9	7	% of A.V.	40 b.p.
Hartford Life	14	7	% of A.V.	50 b.p.
American Skandia	5	7	% of A.V.	35 b.p.
Travelers	10	5 or 10	% of A.V.	40 b.p.
U.S. Allianz	4	5 or 10	% of A.V.	N.A.

Notes: U.S. Market as of mid-2004. The insurance fee—to cover the GMWB—is listed in basis points (b.p.) and charged as a percent of a Guaranteed Base (G.B.) or the Account Value (A.V.). The G.B. is the total premiums paid, minus guaranteed withdrawals plus any bonuses.

options, such as executive and incentive options. This, of course, should impact the pricing of any (illiquid, non-tradeable) derivative security offered to retail investors where a portion of the value is based on the counter-party optimally exercising the option. Yet, the liability created by GMWBs should have a direct and measurable impact on the amount of capital (and reserves) insurance companies should be required to hold against these guarantees. Traditionally insurance companies have relied on the law of large numbers to set reserves which cover the risks $(1 - \varepsilon)\%$ of the time. But as insurance companies venture into offering products which merge life insurance and financial (downside) protection, there is a need to value the *financial economic* risks they are undertaking, especially given the recent movement towards fair value accounting and risk-based capital in the insurance industry.

Therefore, in this paper we present two extreme valuation algorithms—both within the framework of No Arbitrage pricing—for pricing the GMWB. First, we take a *static* approach that assumes individual investors behave passively in utilizing their guarantee. In this case we show how the rider can be decomposed into a Quanto Asian Put (QAP) plus a generic term-certain annuity. We believe this bifurcation has not been previously noted in the literature, and obviously allows the insurance company to use QAPs to hedge the product. There is a direct benefit to viewing the rider this way, in that a QAP is a modest variation on products that firms have extensive experience in hedging.

We rely on numerical techniques to price the embedded Asian options and do not give much atten-

tion to the algorithm per se, since this is a well studied problem in the literature. See Turnbull and Wakeman (1991), Milevsky and Posner (1998), Bakshi and Madan (2002), Nielsen and Sandmann (2003) or Dhaene et al. (2001), for a review of the various approaches. In this paper, we use PDE based numerical techniques to obtain values for the hedging cost of the GMWB under this *static* approach.

The opposite assumption is that all investors buying these GMWB features are dynamically rational and seek to maximize the embedded option value by lapsing (a.k.a. surrendering or terminating) the product at an optimal time, i.e. once the expected present value of fees exceeds the present value of benefits. We label this the *dynamic* approach, and its analysis leads to an optimal stopping problem akin to pricing an American put option, albeit complicated by the non-traditional payment structure. We formulate the optimal boundary as a linear complementarity problem—with an embedded ‘fixed point problem’—and then use numerical PDE techniques to obtain pricing results.

As alluded to earlier, we find that under a stylized product specification which guarantees a 7% withdrawal, and assuming a forward-looking investment volatility of $\sigma = 20\%$, the cost of providing a GMWB ranges from 73 to 160 basis points of assets per annum, with the variation depending on the degree of what we label, *lapsation rationality*. Of course, our pricing does not allow for any profits, commissions, fees and transaction costs, which is also the situation with the celebrated Black–Scholes formula. We are confident that more sophisticated pricing models that account

for stochastic volatility, jumps, term-structure effects, etc, and other recent innovations will only increase the price of the embedded option, although we do not pursue this approach in this paper. Yet, in contrast to our estimates, we find the recent GMWB products that have been introduced in the market are only charging 30–50 basis points, even though the underlying annuity sub-accounts contain high-volatility investment choices.

From a broader perspective, during the last 10 years there has been nothing short of an explosion of exotic options and financial guarantees being embedded within insurance policies. In fact, some have argued, for example Boyle and Boyle (2001), that the options embedded within insurance policies are even more complex than those in standard OTC and exchange traded contracts. And, while the rationale for this phenomena requires some justification, the embedded options are at times quite challenging to price, value and hedge. Historically, they have been analyzed by a variety of academics and practitioners under the label of equity-linked policies, starting with the extension of Black and Scholes (1973) by Brennan and Schwartz (1976) and Boyle and Schwartz (1977), or more recently Persson and Aase (1997) as well as Milevsky and Posner (2001) and Windcliff et al. (2001). In fact, we count more than 60 published papers—most of them in the insurance literature—written on the topic within the last 10 years alone. For a selected bibliography and recent books on the topic, we refer the interested reader to Hardy (2003). But as mentioned earlier, the contribution of this paper is to take a financial economic approach to the (new) GMWB features that differentiates between various forms of rationality and contrast these values with actual pricing in the market. Our work is also in the spirit of Pennacchi (1999), Sherris (1995) or Cox et al. (2004) where option pricing techniques are used to value guarantees embedded within certain pension funds or structured insurance products.

The remainder of the paper is organized as follows. Section 1.1 provides a numerical example to explain the mechanics of a GMWB. In Section 2 we provide a stochastic modeling framework for the GMWB and discuss the real-world probability the feature will end-up in-the-money. Of course, under optimistic enough assumptions for future market growth rates, it is easy to justify ‘ignoring’ the liability created by this option. The probability of usage will directly impact the actuarial insurance reserves held against the guarantee. Sec-

tion 3 uses the No Arbitrage (hedging) approach under the static analysis of the GMWB and decomposes the product into a term-certain annuity and a Quanto Asian Put. We also provide some numerical examples. Section 4 illustrates the dynamic perspective by solving the relevant optimal stopping problem, and Section 5 concludes the paper.

1.1. Numerical example of product specifics

Table 2 provides a numerical example of the payoff from a GMWB rider, assuming a particular sequence of quarterly investment returns for a typical variable annuity (VA) policy. The example assumes an initial investment of \$100 and a guaranteed withdrawal of \$7 per annum—the most prevalent structure—which is \$1.75 per quarter. At the end of each quarter, an investment return is recorded and applied to the previous end-of-quarter’s account balance. Thus, for example, after the first quarter return of negative 12.24%, the balance in the VA policy was \$87.76, and then \$1.75 was withdrawn. The next quarter resulted in a positive 10.06% return, and the account grew to \$94.66, etc.

We can decompose the GMWB payoff stream as account withdrawals plus an *insurance option* that steps in once the account is depleted (note: we will consider a different decomposition later, as a term-certain annuity plus a *Quanto Asian put*). At the end of the third quarter of the seventh year (i.e. after 27 quarters or 6.75 years), the balance in the account is 0.17 dollars under the above (randomly generated) sequence. This is not sufficient to finance the \$1.75 payment then due, and so the *insurance option* payoff starts. Under a standard systematic withdrawal plan there is no longer enough to withdraw the requisite \$1.75 per quarter and the policy is therefore ‘ruined’. In fact, under this particular state of nature, the total amount withdrawn prior to the end of 6.75 years is \$45.42, due to the mostly poor performance of the investments during the first few years. The insurance option kicks in and continues to provide an income of \$1.75 until the entire \$100 has been returned. Note that the entire \$100 will be returned in exactly $100/1.75 = 57.14$ quarters which is $57.14/4 = 14.285$ years. At the end of 14.28 years the entire sum is returned and the guarantee *matures*. The insurance company backing the VA policy and the guarantee would be ‘on the hook’ for the remaining $100 - 45.22 = 54.78$ dollars, albeit paid over the re-

Table 2
Numerical example

Time period	Investment return (%)	Balance E.O.Q. (\$)	SWiP (\$)	GMWB (\$)
0.25	−12.24	87.76	1.75	1.75
0.5	10.06	94.66	1.75	1.75
0.75	−6.43	86.94	1.75	1.75
1	−26.40	62.70	1.75	1.75
1.25	−17.30	50.41	1.75	1.75
1.5	−14.22	41.74	1.75	1.75
1.75	−7.05	37.17	1.75	1.75
2	−7.42	32.79	1.75	1.75
2.25	6.95	33.19	1.75	1.75
2.5	5.76	33.25	1.75	1.75
2.75	5.79	33.33	1.75	1.75
3	−1.68	31.05	1.75	1.75
3.25	3.13	28.38	1.75	1.75
3.5	16.57	22.22	1.75	1.75
3.75	−15.73	17.25	1.75	1.75
4	3.47	16.04	1.75	1.75
4.25	14.40	16.35	1.75	1.75
4.5	14.02	16.64	1.75	1.75
4.75	4.56	15.57	1.75	1.75
5	8.67	15.02	1.75	1.75
5.25	−9.40	12.02	1.75	1.75
5.5	0.70	10.34	1.75	1.75
5.75	−4.59	8.20	1.75	1.75
6	5.77	6.82	1.75	1.75
6.25	1.75	5.16	1.75	1.75
6.50	0.05	3.41	1.75	1.75
6.75	16.15	1.93	1.75	1.75
7	−6.83	0.17	0.17	1.75
7.25	14.98	−	−	1.75
7.50	0.33	−	−	1.75
7.75	−4.33	−	−	1.75
8	−1.22	−	−	1.75
8.25	−3.88	−	−	1.75
8.50	23.84	−	−	1.75
8.75	3.70	−	−	1.75
9	2.36	−	−	1.75
9.25	−18.28	−	−	1.75
9.50	3.68	−	−	1.75
9.75	8.31	−	−	1.75
10	−1.54	−	−	1.75
10.25	18.62	−	−	1.75
10.5	−15.57	−	−	1.75
10.75	−9.92	−	−	1.75
11	10.66	−	−	1.75
11.25	−1.58	−	−	1.75
11.5	−17.23	−	−	1.75
11.75	−1.05	−	−	1.75
12	13.94	−	−	1.75
12.25	−5.09	−	−	1.75
12.5	1.51	−	−	1.75
12.75	−4.33	−	−	1.75
13	12.06	−	−	1.75
13.25	19.85	−	−	1.75

Table 2 (Continued)

Time period	Investment return (%)	Balance E.O.Q. (\$)	SWiP (\$)	GMWB (\$)
13.50	4.76	–	–	1.75
13.75	–13.18	–	–	1.75
14	5.43	–	–	1.75
14.25	–9.83	–	–	1.75

Hypothetical return, cash flow and end-of-quarter account balance comparing a regular systematic withdrawal plan (SWiP) against the payoff from a Guaranteed Minimum Withdrawal Benefit (GMWB).

maining 7 years. Of course, Table 2 represents but one of many millions of possible scenarios for the 14.28 years of the guarantee’s life. Figs. 1–3 illustrate three other possible scenarios.

In the first case, the account is driven to a zero value at the beginning of the fifth year, and the insurance company pays the remaining \$ 1.75 per quarter until time 14.28. In the second case, the funds are exhausted in the middle of the 11th year and the insurance company makes 3 years of additional payments. In the final case the variable annuity survives a 7% withdrawal for 14.28 years and the company is relieved of its obligations. In the first two scenarios the policy holder receives no refund of capital at the end of the payment period. Whereas in the last scenario, there is a remaining account balance to which the policy holder is entitled at the end of 14.28 years. A more recent variation on this idea is a GMWB for life which—assuming the variable annuity is purchased and withdrawals are not made prior to age 65—implicitly guarantees an income stream akin to a life annuity.

The insurance option appears quite novel upon first inspection, since it starts paying-off at a random *ruin time* for the underlying investment-net-of-withdrawal process. A random maturity option was first analyzed by Carr (1998)—albeit in an attempt to price American put options—but our product specifics are quite different since the random maturity (ruin time) in our product is determined endogenously. We will discuss the precise nature of the embedded option in Section 3. The next section will set-up the model notation and examine the probability—under the real world measure—that the GMWB feature will pay off, which is relevant for traditional insurance pricing and reserving. We then contrast with a financial economic approach.

2. Static modeling framework

Let W_t denote the market value of the underlying VA at any future time $t \geq 0$, with an arbitrary (but innocuous) assumption that $w_0 = 100$ dollars. The most

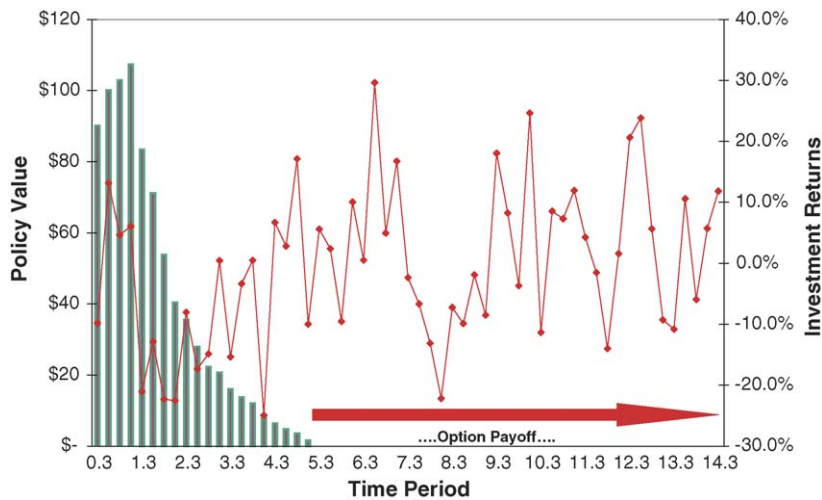


Fig. 1. Example of policy value under 7% withdrawal and random investment returns.

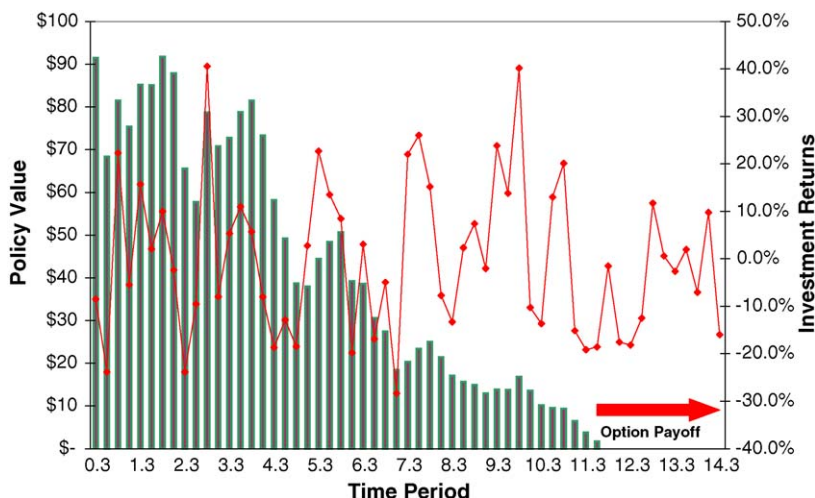


Fig. 2. Example of policy value under 7% withdrawal and random investment returns.

typical GMWB structure is that the policyholder is guaranteed to be able to withdraw *at least* $G = gw_0 = 7$ dollars per annum. The guarantee remains in effect until the entire \$ 100 has been disbursed which, at minimum, is a period of $100/7 = 14.28$ years. Thus, even in the extreme scenario where the initial $w_0 = 100$ collapses to a zero value 1 day after the policy is purchased, the investor will be made whole, albeit over an extended period of 14.28 years. Of course, in any given year the policyholder is entitled to withdraw an amount less than $G = 7$ dollars, which would extend

the life of the guarantee. Alternately, the policyholder could withdraw an amount greater than $G = 7$ dollars which would reduce both the value *and* life of the guarantee. These cases where the policyholder withdraws more or less than suggested by the guarantee—which falls under the category of dynamic strategies—will be carefully addressed in Section 4. In this section, we proceed by assuming the policyholder withdraws not more and not less than the $G = 7$ dollars per annum. This is called the *passive or static* approach. Most, if not all, insurance companies as-

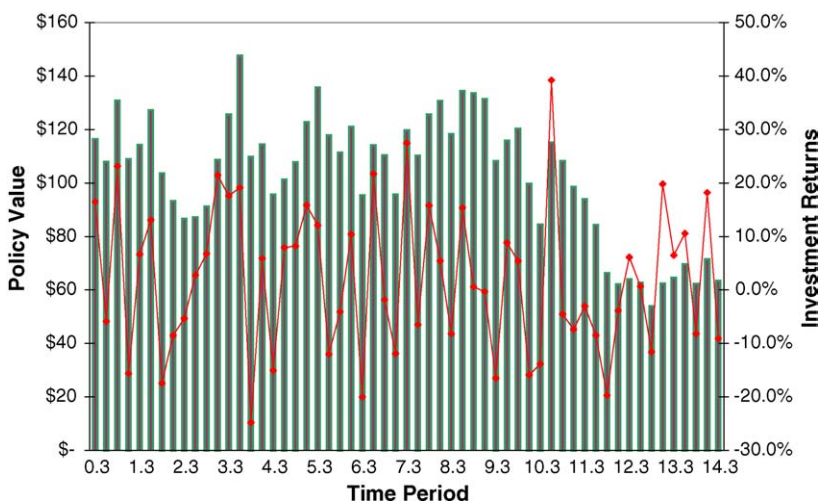


Fig. 3. Example of policy value under 7% withdrawal and random investment returns.

sume this type of behavior on the part of policyholders.

Following most of the option pricing literature, we assume the actual dynamics of the assets underlying the VA policy—i.e. before the deduction of any insurance fees—obey the following stochastic differential equation (SDE):

$$dS_t = \mu S_t dt + \sigma S_t dB_t. \tag{1}$$

The symbol B_t denotes a standard Brownian motion with mean zero and variance t . The parameter (μ) μ is the real-world expected growth rate of the asset class, and (σ) σ represents the volatility of the investment return. This assumption is critical to our valuation model and we are obviously sympathetic to criticism of the implicit lognormality assumption, which has been shown to be inconsistent with financial data. That said, our objective is not to provide a complete risk management system for these guarantees but rather to provide a conceptual analysis of the product and provide some initial estimates for the cost of hedging in capital markets.

The value W_t of the VA sub-account incorporates two additional effects: proportional insurance fees and withdrawals. In general, its dynamics will be

$$dW_t = (\mu - \alpha)W_t dt - \gamma_t dt + \sigma W_t dB_t, \tag{2}$$

while $W_t > 0$. That is, Eq. (2) holds while $t < \tau_0$, where τ_0 is the first time W_t hits 0. Here the parameter α captures the insurance fee that pays for the guarantee, and $0 \leq \gamma_t \leq W_t$ represents the discretionary withdrawals from the account, which can range from a low of zero, to as high as the actual account value W_t . The individual is assumed to invest an amount $W_0 = w_0$ in the variable annuity. The dynamic model we have selected for the underlying investment account is consistent with a large volume of research on “pricing insurance guarantees” such as Gerber and Shiu (2003) or Windcliff et al. (2001) for example.

In what follows next we assume that the withdrawal amount is exactly equal to the guaranteed amount $\gamma_t := G$, which is what we label the *passive or static* approach. Thus, in the passive case,

$$\begin{aligned} dW_t &= (\mu - \alpha)W_t dt - G dt + \sigma W_t dB_t, \\ W_0 &= w_0, \end{aligned} \tag{3}$$

at least while $W_t > 0$. If the account value W_t ever reaches 0, it remains 0. That is, equation (3) holds for $t < \tau_0$ and $W_t = 0$ for $t \geq \tau_0$.

Using standard techniques which can be verified by Ito’s lemma—see Karatzas and Shreve (1992) for details—the solution to the SDE in Eq. (3) can be written as:

$$\begin{aligned} W_T &= e^{(\mu - \alpha - (1/2)\sigma^2)T + \sigma B_T} \max \\ &\times \left[0, \left(w_0 - G \int_0^T e^{-(\mu - \alpha - (1/2)\sigma^2)t - \sigma B_t} dt \right) \right]. \end{aligned} \tag{4}$$

The first thing to note about the dynamics in Eqs. (3) and (4) is that since $G > 0$, which means that the process includes a forced dollar consumption, the value of W_t can in fact hit zero at some point $t > 0$. Although the exponential Brownian motion term is always positive, as soon as the integral term in Eq. (4) exceeds w_0/G , the quantity in the brackets will become negative. This is in contrast to a standard geometric Brownian motion, which is the term multiplying the brackets in Eq. (4) that can never hit zero in finite time. *Ex post*, the guaranteed ability to withdraw G per annum until time $T = w_0/G$ is of value *if and only if* the process W_t hits zero prior to T . Indeed, for those sample-paths for which the ruin time occurs after T , the insurance option has a zero payout since the minimum withdrawal would have been satisfied endogenously, *even* without an explicit guarantee provided by the insurance company.

Given the importance of the ruin time in the classification and understanding of this financial guarantee, we introduce the following notation for the probability of ruin of the process W_t , within the time period $[0, t]$. The function ξ_t is:

$$\begin{aligned} \xi_t &= P \left[\inf_{0 \leq s \leq t} W_s = 0 \right] \\ &= P \left[\int_0^t e^{-(\mu - \alpha - (1/2)\sigma^2)s - \sigma B_s} ds \geq \frac{w_0}{G} \right] \\ &:= P[\mathbf{X}_t \geq \frac{w_0}{G}], \end{aligned} \tag{5}$$

where the new term \mathbf{X}_t is defined equal to the integral in the middle of Eq. (5). The seemingly counter-

intuitive relationship between the infimum of a process and the integral of an exponential Brownian motion comes from the fact that Eq. (4) can only reach 0 once the integral \mathbf{X}_t exceeds w_0/G . Note also the fact that \mathbf{X}_t is monotonically increasing in t . Thus, once \mathbf{X}_t exceeds w_0/G , which means that $W_t = 0$, it can never recover and go back above zero.

It is quite easy to demonstrate that the probability of ruin ξ_t is increasing in the withdrawal rate G , and likewise, the greater the time t , the higher the probability of ruin. In fact, although it is beyond the scope of the current analysis, one can actually obtain a precise analytic expression for ξ_t when $t \rightarrow \infty$. In the current context, the traditional insurance company is most interested in the value of ξ_T where $T = w_0/G$. In other words, we would like to know what the odds are that the investor would actually run out of money by the end of the guaranteed period, assuming they withdrew the guaranteed amount as part of a systematic withdrawal plan.

2.1. So, what is the real-world probability of ruin?

Assume the arithmetic average return is expected (after management fees, but prior to insurance fees) to be $\mu = 9\%$ per annum jointly with a historical market volatility of $\sigma = 18\%$. According to Morningstar Principia Pro, the median sub-account volatility for the universe of variable annuity policies is 18%, with a 25th percentile of 16% and a 90th percentile of 25%. Also, we let the insurance fee for this particular GMWB rider be set to $\alpha = 0.40\%$ per annum, which is consistent with the current market pricing of these products. In this case, the parameterized dynamics of the investment become (while $W_t > 0$):

$$dW_t = ((0.086)W_t - 7)dt + 0.18W_t dB_t,$$

$$w_0 = 100. \quad (6)$$

Using numerical PDE methods—described at greater length in Appendix A—to obtain the ruin probability during the first $T = 14.28$ years, we find that $\xi_{14.28} = 11.7\%$. In other words, there is approximately an 88.3% chance that *even* if the policy holder withdraws the maximum allowable amount each year, the policy will *survive* to the end of the guaranteed horizon. But if we increase the investment return volatility to $\sigma = 25\%$ per annum, the ruin probability increases to $\xi_{14.28} = 26.2\%$. And, if we reduce the expected (arithmetic average) return to $\mu = 6\%$ and maintain a high $\sigma = 25\%$ volatility, the probability of ruin increases to $\xi_{14.28} = 39.9\%$, which are clearly non-trivial amounts. Table 3 displays the probabilities under various risk and return combinations.

Note that if the expected investment return is increased to $\mu = 12\%$ and the volatility of the return is set to $\sigma = 10\%$, the probability the withdrawals of $G = 7$ dollars per annum will actually exhaust (or ruin) the policy prior to time $T = 14.28$ is less than one half of a percent. Thus, an overly optimistic insurance actuary focused on the real-world probability of an ‘insurance payout’ would likely ignore this event all together. Note that some insurance companies impose restrictions on the asset allocation within the variable annuity policy. For example, they might force 20–40% of the sub-accounts into (less risky) bond and fixed income investments in order to limit the effective volatility σ and hence the probability that the option will mature in the money.

In any event, the probability of ex ante usage range from 0.5 to 20% depending on the subjective asset return assumption and characteristics, these will impact

Table 3
GMWB (SWiP) probability of ruin within 14.28 years (40 b.p. insurance fee with net investment return (μ) and volatility (σ))

	$\mu = 4\%$	$\mu = 6\%$	$\mu = 8\%$	$\mu = 10\%$	$\mu = 12\%$
$\sigma = 10\%$	19.0%	7.0%	1.7%	0.3%	0.04%
$\sigma = 15\%$	31.4%	18.5%	9.3%	4.1%	1.6%
$\sigma = 18\%$	37.8%	25.5%	15.5%	8.6%	4.4%
$\sigma = 25\%$	49.9%	39.6%	30.5%	22.2%	15.5%

In absence of an explicit GMWB, the process of withdrawing \$7 (per annum) for each \$100 of original principal—which is sometimes called a Systematic Withdrawal Plan—might drive the portfolio to ‘ruin’ within 14.28 year, the time over which the \$100 would be recouped. The above table computes the probability this event would occur under a variety of (real world) drift and volatility assumptions. Insurance reserves—in contrast to financial economic hedging—would be based on the above table. See Appendix A for algorithm used to compute ruin probabilities.

the setting of traditional insurance reserves. The relevant question to a financial economist interested in the fair value of liabilities, is: *How much does it cost the insurance company to hedge this guarantee in the capital market?*

3. Static analysis

In this section we illustrate how to bifurcate the product into a collection of strip-bonds (or a term-certain annuity) and a Quanto Asian Put (QAP) option. Note that $g = G/w_0$ and by definition $T = 1/g$ (since the product terminates or matures when all the funds have been returned) and so we have that:

$$W_T = w_0 e^{(\mu - \alpha - (1/2)\sigma^2)T + \sigma B_T} \max \left[0, \left(1 - \frac{1}{T} \int_0^T e^{-(\mu - \alpha - (1/2)\sigma^2)s - \sigma B_s} ds \right) \right]. \quad (7)$$

The payoff of the QAP option is:

$$\text{Option Payoff} := W_T, \quad (8)$$

since the holder of the variable annuity policy is guaranteed to receive any remaining funds in the account at time $T = 1/g$. Remember that the policyholder is also entitled to the periodic income flow in addition to the (possibly zero) maturity value of the account. Thus, focusing on the future value of all cash-flows and payments, the maturity value of the *periodic income* is:

$$w_0 g \int_0^T e^{rt} dt = \frac{w_0 g}{r} (e^{rT} - 1). \quad (9)$$

The (No Arbitrage) time-zero present value of the GMWB cash-flow package is therefore:

$$e^{-rT} E_Q[W_T] + \frac{w_0 g}{r} (1 - e^{-rT}), \quad (10)$$

where $E_Q[\cdot]$ denotes the expectation under the Q -measure, under which the real-world drift μ is replaced by the risk-free rate r . We refer the interested reader to any standard textbook on derivative pricing to justify this substitution of measures.

Finally, for the GMWB to be fairly priced we must have, at inception, that the amount invested in the prod-

uct w_0 , is equal to the value of the cash-flow package, where $T = 1/g$

$$w_0 = e^{-r/g} E_Q[W_{1/g}] + \frac{w_0 g}{r} (1 - e^{-r/g}). \quad (11)$$

Eq. (11) is a self financing condition for the GMWB and is one of our main results. It states that for the product to be fairly structured, the initial purchase price must equal the cost of the term-certain annuity plus the exotic option. For any given (r, σ) pair we can locate the (α, g) curve across which the product is fairly priced, which implies equality in Eq. (11).

The option component is effectively a Quanto Asian Put (QAP) defined on an underlying security that is the inverse of the account price process. To see this, define a new (reciprocal) process:

$$Y_t = S_t^{-1} = e^{-(r - \alpha - (1/2)\sigma^2)t - \sigma B_t}, \quad Y_0 = 1. \quad (12)$$

One can think of Y_t as the number of VA sub-account units that 1 dollar can buy, similar to the number of Euros or Yen than 1 dollar can purchase in the currency market. The inverse, $S_t = Y_t^{-1}$, is the value of one VA sub-account unit in dollars, similar to the price of one Euro or Yen in USD.

Now let:

$$\mathbf{A} := \frac{1}{T} \int_0^T Y_t dt, \quad \mathbf{Y} := Y_T, \quad (13)$$

which is an average of the reciprocal account value. The payoff from the QAP option at maturity is:

$$\text{Option Payoff} := w_0 \frac{\max[1 - \mathbf{A}, 0]}{\mathbf{Y}}. \quad (14)$$

This represents w_0 units of a Quanto (Fixed Strike) Asian Put option. In sum, scaling everything by the initial premium, a fairly priced product *at inception* implies the relationship:

$$e^{-r/g} E_Q \left[\frac{\max[1 - \mathbf{A}, 0]}{\mathbf{Y}} \right] + \frac{g}{r} (1 - e^{-r/g}) = 1. \quad (15)$$

Given values of the other parameters, the fair insurance fee α can be obtained by solving this equation.

Our main qualitative insight is that under a static perspective, this product can be decomposed into the following items:

1. a term-certain annuity paying G per annum for a period of $T = w_0/G$ years, plus,
2. a Quanto Asian Put (QAP) on the above-mentioned *reciprocal* variable annuity account.

For example, for an initial deposit of $w_0 = 100$, a guarantee withdrawal amount of $G = 7$ per annum, and an interest rate of $r = 0.06$, the time-zero cost of the term-certain annuity component is 67.15 dollars. The remaining 32.85 would go towards purchasing the option, and α is determined so that this represents the fair option value. One can think of a VA with a GMWB as consisting of 67% term-certain annuity and 32% Quanto Asian Put option. In contrast, at a (lower) interest rate of $r = 0.05$, the cost of the term-certain annuity would be (a higher) 71.46 dollars, and only 28.54 would go towards purchasing the required option.

Table 4 displays the required insurance fee that would lead to an equality in Eqs. (11) or (15) under a number of different volatility values. Note the fixed-point nature of the problem. Once the volatility σ , interest rate r and guarantee rate G have been selected, we must numerically search for a fee value α so that we get equality in Eq. (15). We actually price the (Quanto) Asian Put option using a numerical PDE technique which is described in Appendix A. For example, if the VA guarantees a 7% withdrawal and the pricing volatility is $\sigma = 20\%$, the fair insurance fee would be approximately $\alpha = 73$ basis points of assets per annum. Stated differently, a financial package which offers a stream of \$7-per-annum income (in continuous time) plus a Quanto Asian Put that matures in exactly $T = 14.29$ years is worth precisely $w_0 = 100$, when

the investment on which the option is struck is ‘leaking’ a dividend yield of 73 basis points per annum. If the guarantee is reduced to $g = 4\%$, which implies the product matures in $T = 25$ years, the fair insurance fee is only 23 basis points. Likewise, if the guarantee is increased to $g = 9\%$, which implies the product matures in $T = 11.11$ years, the fair insurance fee is 117 basis points. As we mentioned in Section 1, the most common GMWB guarantee (in mid-2004) being offered on variable annuities is $g = 7\%$, which (even) under a conservative $\sigma = 15\%$ volatility implies an insurance fee of 40 basis points.

The identification of the GMWB as a term-certain annuity plus a QAP is useful from several points of view. First, though we use PDE techniques to value it, this is not essential—there are a variety of other well studied approaches to the valuation of Asian options and our identification opens these approaches up for use in this case. For example, an alternative approach to (approximately) pricing Asian options is by Dhaene et al. (2001), which could also be used in this case. Second, there is an established market for Asian options, which raises the possibility of hedging using these products instead of via dynamic hedging. Finally, there is a body of practical experience with the hedging of Asian options, which turns the QAP into a much more familiar type of product than a GMWB first appears.

3.1. Impact of death and mortality

The inclusion of mortality—or death during the $T = 1/g$ life of the product—will reduce the value

Table 4
The impact of the GMWB rate and the volatility of the sub-account on the required fee α

Guarantee rate, g (%)	Maturity (years), $T = 1/g$	Investment volatility, σ	
		20%	30%
4	25.00	23b.p.	60 b.p.
5	20.00	37b.p.	90 b.p.
6	16.67	54b.p.	123 b.p.
7	14.29	73b.p.	158 b.p.
8	12.50	94b.p.	194 b.p.
9	11.11	117b.p.	232 b.p.
10	10.00	140b.p.	271 b.p.
15	6.67	272b.p.	475 b.p.

Assumes a 5% pricing interest rate. The table displays the required insurance fee α to hedge the GMWB assuming everyone holds the product to maturity (i.e. the static actuarial analysis). The maturity of the Quanto Asian Put (QAP) is the inverse of the guarantee withdrawal rate since that is the time at which the original principal has been recovered.

and hedging cost of the guarantee. Indeed, if we assume the guarantee will be terminated upon death of the policyholder who is currently aged x —and the beneficiary of the VA only receives the market value at the time of death—then the term-certain annuity should be replaced with a term-life annuity that terminates at $T = w_0/G$. If we let the actuarial symbol $({}_t p_x)$ denote the probability an individual currently aged- x survives for t more years, the time zero cost of the term-certain component becomes:

cost of term-certain annuity

$$= \int_0^T G({}_t p_x) e^{-rt} dt \leq \int_0^T G e^{-rt} dt, \quad (16)$$

and thus, mortality will reduce the cost of providing the guaranteed GMWB.

Thus, although Table 4 provides a value (or passive replicating cost) for the GMWB under the assumption that everyone ‘behaves’ exactly as predicted, in reality the insurance company can push the pricing to even lower levels by assuming a fairly high (real world) probability of death. This fact might explain the reason why observed GMWB fees in practice appear lower than dictated by Table 4. Of course, an alternative reason for apparent underpricing is that the base insurance fee on the VA product without the rider is enough to subsidize the extra cost of the relatively more expensive GMWB as well as the earlier-mentioned possibility that asset allocation constraints reduce the exposure. All of this is predicated on the *static* approach that policyholders do not deviate from the $\gamma_t = G$ dollars per annum withdrawal. But, it might in fact make sense for the policyholder to withdraw more or less than the minimum, even if it reduces the base of the guarantee, if the account has performed sufficiently well, thus making the original guarantee less valuable. It is not clear *a priori* the conditions under which this would make sense. This brings us to the next section which covers the pricing of these guarantees in perfect and complete financial markets where all counter-parties are fully rational and ‘option value’ maximizers.

4. Dynamic model and hedging

In this section we employ *American option* pricing techniques to obtain a dynamic model of the GMWB,

assuming policyholders are fully rational and lapse (i.e. withdraw more or less from) the product when it works to their economic advantage. As we argued in the introduction, the true ‘cost’ of the embedded guarantee lies somewhere between the *static* embedded option cost and the *dynamic* hedging cost.

It is important to note that once we include strategic lapsation as an option for the policyholder, the contingent deferred surrender charge (DSC) becomes an important factor in driving optimal behavior. Recall that most variable annuities impose a penalty if the product is lapsed or surrendered prior to maturity. This penalty is calculated either as a fraction of the account value at the time of surrender or the fraction of the original premium, and can range from 10 to 0% depending on the product, company, and the time that has elapsed since the policy was acquired. In what follows we will analyze products for which the DSC is imposed as a percent of account value. And, while current practice in the industry is such that the DSC goes exclusively towards paying commissions and brokerage fees—and is not used for risk management or hedging purposes—this penalty does induce the policyholders to continue holding the product and paying the ongoing asset-based management fees, even though the embedded option is far out-of-the-money. From a dynamic point of view, we must work with a DSC curve or schedule in any optimal stopping model that attempts to capture the salient features of the product.

4.1. Borrowing the methodology of the American put

We begin by presenting the American put option pricing problem in a slightly different way, which leads naturally to our methodology for the GMWB. In contrast to recent work by Ju (1998) or Carr (1998) for example, on extending the theory and practice of American option pricing in a variety of directions, we go back to basics and formulate the problem as a linear complementarity problem—a.k.a. free boundary value problem—leading to a partial differential inequality. This approach is not novel within itself, see for example Dempster and Hutton (1999). Rather, we do this short review to make clear the analogy between the American Put and a GMWB.

Consider a contingent claim which depends on the traded underlying security price S_t . When exercised at

time t , the claim pays off an amount $f(S_t)$. Under the real-world measure P , the security price process obeys:

$$dS_t = \mu S_t dt + \sigma S_t dB_t. \tag{17}$$

We can rewrite this as

$$dS_t = rS_t dt + \sigma S_t d\tilde{B}_t, \tag{18}$$

where $\tilde{B}_t = B_t + t(\mu - r)/\sigma$. Define a new probability measure Q by re-weighting the probabilities so that \tilde{B}_t is a Brownian motion under Q , which is the risk neutral probability measure. It follows by Itô's formula that $e^{-rt}S_t$ is a martingale under Q . This implies that the discounted value $e^{-rt}V_t^0$ of any self-financing portfolio is a martingale as well. Equivalently

$$dV_t^0 = rV_t^0 dt + dM_t, \tag{19}$$

where M_t is a Q -martingale.

In a complete market there is a perfect hedge for our contingent claim, which covers the option payoff at any exercise time, even the optimal time (the worst-case scenario). Let V_t denote the value of this hedge at time t . We will incorporate the cash flow resulting from exercising the option into the hedge, so V_t is no longer self-financing. But a formula analogous to (19) will still hold. Write η for the exercise time (which is not necessarily assumed to agree with the optimal exercise time τ). The cash flow from the hedge, if exercised at the stopping time η is $f(S_t) dR_t$, where

$$R_t = \begin{cases} 1, & t < \eta, \\ 0, & t \geq \eta. \end{cases} \tag{20}$$

The hedge now satisfies

$$dV_t = rV_t dt + dM_t + f(S_t) dR_t. \tag{21}$$

With optimal exercise (i.e. $\eta = \tau$), M_t will be a Q -martingale. In general—even with suboptimal exercise—it will be a Q -supermartingale. On the other hand, the value of hedge should be a function of the stock price, dropping to zero after exercise:

$$V_t = v(t, S_t)R_t. \tag{22}$$

Substituting into Itô's formula gives:

$$dV_t = \left[v_t + rS_t v_x + \frac{\sigma^2 S_t^2}{2} v_{xx} \right] R_t dt + v dR_t + \sigma S_t v_x R_t d\tilde{B}_t. \tag{23}$$

Equating the two expressions,

$$\begin{aligned} & \left[v_t + rS_t v_x + \frac{\sigma^2 S_t^2}{2} v_{xx} - rv \right] R_t dt + [v - f(S_t)] dR_t \\ & = dM_t - \sigma S_t v_x R_t d\tilde{B}_t. \end{aligned} \tag{24}$$

The RHS is a supermartingale in general, and a martingale under optimal exercise. Since the LHS is of bounded variation, it must be ≤ 0 in general, and $= 0$ under optimal exercise. That is,

$$v_t + rxv_x + \frac{\sigma^2 x^2}{2} v_{xx} - rv \leq 0, \quad f - v \leq 0 \tag{25}$$

with at least one expression holding with equality. This is a linear complementarity problem—a.k.a. free boundary value problem—whose solution can be found numerically to give the option price. With this background, we now return to the pricing of the GMWB.

4.2. Hedging the GMWB

Let W_t be the value of the variable annuity account with a guaranteed minimum withdrawal benefit (GMWB) rider. Associated with this extra guarantee is the insurance fee $\alpha\%$ and the contingent deferred surrender charge (CDSC) of $\kappa\%$. As in the *static* case, these fees are imposed solely for hedging purposes. The GMWB allows up to $G dt$ to be withdrawn in time dt , so the CDSC applies only to withdrawals in excess of $G dt$. The nominal withdrawal rate G is set as being a fixed percentage g , of the initial account value w_0 . The terms of the GMWB contract specify that as long as the rate of withdrawals stays below G , the account holder may eventually accumulate withdrawals of an amount A_t from the account, even if doing so would ordinarily drive the account to zero. Initially the guarantee level equals the account value $A_0 = x_0 (=W_0)$, but W_t then fluctuates. Withdrawals decrease both W_t and A_t . If withdrawals ever occur at a rate higher than G , then not only is the CDSC imposed, but after the guarantee level and account value are debited by the withdrawal, the guarantee level is reset to the smaller of its value and the account value. This reset provision acts as a disincentive to large withdrawals. As argued in the introduction, these provisions are idealized versions of those from several existing variable annuities. Other

terms could be analyzed the same way, for example with a CDSC that varies with time or with provisions for resetting the withdrawal rate G .

Mathematically, we use the same GBM model (as in the static section) for W_t :

$$dW_t = (r - \alpha)W_t dt + \sigma W_t d\tilde{B}_t - \gamma_t dt \quad (26)$$

under the Q -measure, while $W_t > 0$. Here γ_t models continuous withdrawals from the account, which may or may not equal the allowed amount G . Similarly

$$dA_t = -\gamma_t dt \quad \text{provided } \gamma_t \leq G = gx_0, \quad (27)$$

but if $\gamma_t > G$ dollars, then A_t jumps to $\min(A_t, W_t)$. There are similar expressions in the case of lump-sum withdrawals $\Delta W_t < 0$, but for simplicity sake we will carry out the analysis in the continuous case only. In fact, the same optimum is obtained regardless of whether lump sum withdrawals are permitted.

We wish to hedge this account. Write $V_t = v(A_t, W_t)$ for the value of the hedge. Insurance and DSC fees are retained in the hedge, so

$$dV_t = rV_t dt + dM_t - f(\gamma_t) dt, \quad \text{where} \quad (28)$$

$$f(\gamma_t) = \begin{cases} \gamma_t & \text{if } 0 \leq \gamma_t \leq G, \\ G + (1 - \kappa)(\gamma_t - G) & \text{if } \gamma_t > G. \end{cases} \quad (29)$$

Here M_t is a supermartingale, and a martingale under the optimal choice for γ_t .

An analysis similar to that of the American option lets us solve numerically for the hedging cost $v(a, w)$ by solving a free boundary value problem numerically. In contrast to the classical American put, there is no longer an initial fee on which to base the hedge—the initial value of the hedge is constrained to equal the initial account value, so the hedge must be financed through the insurance fee and DSC. In fact, our real problem is to determine, for a given value κ of the DSC, what the α is that allows the guarantee to be hedged. In mathematical terms, we carry out an iterative solution and locate a *fixed point* in terms of α . This is the same approach taken by Boyle (2005) as well. That is, for a given α we first solve the free boundary value problem to give an initial hedging cost $v_\alpha(w_0, w_0)$. We then adjust α , resolve the PDE, readjust α , etc., to converge on the value of α that makes

$$v_\alpha(w_0, w_0) = w_0. \quad (30)$$

In principle this might give an α that depends on the initial investment, but recalling that G is a linear function of w_0 , there is, in fact, a scale invariance in the problem from which it can be shown that the same α works for all levels of w_0 .

It turns out that the optimal withdrawal strategy γ_t amounts to withdrawing at an arbitrary large rate when W_t lies above some value $L(A_t)$, and to withdraw at the contracted rate G when $W_t \leq L(A_t)$. It should be emphasized that this *optimal* withdrawal strategy is not necessarily optimal from the point of maximizing the investor's expected utility—rather it should be viewed as the worst case scenario for the issuer (and hedger) of the policy. Again, the capital-markets cost is the price of eliminating all possibility of shortfall. If one is willing to accept some positive shortfall probability or to make modeling assumptions about sub-optimal withdrawal behaviors, the hedging cost can be reduced.

We now derive the equations for v . As was the case for the American option, we have two expressions for V_t , one from (28) and one from Itô's lemma:

$$dV_t = rV_t dt + dM_t - f(\gamma_t) dt, \quad (31)$$

$$\begin{aligned} dv(A_t, W_t) &= v_a dA_t + v_w dW_t + \frac{1}{2} v_{ww} d(W)_t \\ &= -v_a \gamma_t dt + (r - \alpha)W_t v_w dt \\ &\quad + v_w \sigma W_t d\tilde{B}_t - v_w \gamma_t dt + \frac{\sigma^2 W_t^2}{2} v_{ww} dt, \end{aligned} \quad (32)$$

while $W_t > 0$. Equating gives

$$\begin{aligned} &\left[(r - \alpha)W_t v_w + \frac{\sigma^2 W_t^2}{2} v_{ww} - rv \right] dt \\ &+ [f(\gamma_t) - \gamma_t v_w - \gamma_t v_a] dt = dM_t - v_w \sigma X_t d\tilde{B}_t, \end{aligned} \quad (33)$$

where the RHS is a supermartingale, and a martingale under the optimal choice of γ . Thus as before,

$$\begin{aligned} &\left[(r - \alpha)w v_w + \frac{1}{2} \sigma^2 w^2 v_{ww} - rv \right] \\ &+ [f(\gamma) - \gamma v_w - \gamma v_a] \leq 0 \quad \text{for every } \gamma, \end{aligned} \quad (34)$$

with equality for some γ . Because f is piecewise-linear, this reduces to three critical cases, namely $\gamma = 0$, $\gamma = G$, and $\gamma = \infty$. We arrive at the free boundary value problem

$$(r - \alpha)wv_w + \frac{\sigma^2 w^2}{2} v_{ww} - rv \leq 0, \tag{35}$$

$$(r - \alpha)wv_w + \frac{\sigma^2 w^2}{2} v_{ww} - rv + G[1 - v_w - v_a] \leq 0, \tag{36}$$

$$(1 - \kappa) - v_w - v_a \leq 0, \tag{37}$$

with equality in at least one case. Note in particular that the guarantee level α plays the same role in the second equation as time t did in the American put option problem. The numerical techniques used to solve the two problems are virtually identical.

The dynamic hedging strategy (i.e. delta-hedge) can be read off from (32), namely that one holds v_w units of W_t , with the balance $v - v_w W_t$ placed in a money-market account or bond. Of course, what one really wants is a hedge in terms of the traded security S_t rather than W_t . This can also be obtained, and can be shown to consist of $v_w W_t / S_t$ units of the security S_t , with the balance in the bond. In other words, one's S_t holdings have the same value $v_w W_t$ as the W_t holdings had in the delta-hedge. The values of v_w may be computed numerically.

4.3. Numerical comparison of static versus dynamic

Table 5 provides some comparisons between the static and dynamic valuation assuming the contingent deferred surrender charge of $k = 1\%$.

For example, under a $g = 7\%$ withdrawal rate and a pricing volatility of $\sigma = 20\%$, the numerical solution to the system of PDEs in Eqs. (35)–(37) leads to an insurance fee of 160 basis points of assets per annum. This can be compared to the 73 basis points required under the static actuarial case. When volatility is increased to $\sigma = 30\%$, the required insurance fee jumps to 565 basis points of assets. We remind the reader that these numbers are derived under the assumption that $\kappa = 1\%$ and that the insurance company can recover a portion of the hedging cost when the product

Table 5
Static vs. dynamic

Volatility, σ (%)	Static, α^*	Dynamic, α^* (DSC = 1%)
15	40 b.p.	97 b.p.
18	59 b.p.	136 b.p.
20	73 b.p.	160 b.p.
25	113 b.p.	320 b.p.
30	158 b.p.	565 b.p.

Assumes a 7% guaranteed withdrawal rate and 5% (pricing) interest rate. The expected (Q -measure) value of the account at maturity is 58.29% of the initial investment. The discounted expected (i.e. option) value is 28.53% of the initial investment. Under the static actuarial case, 28.53% of premium is used to purchase the Quanto Asian Put, while the remaining 71.47% should go towards buying a term-certain annuity paying \$ 7-per-\$ 100 for a period of 14.28 years. The purpose of this comparison is to note the higher required fees if we assume the individual will lapse the product if the option is no longer worth paying for (dynamic) vs. if the investor holds the product to maturity (static).

is lapsed or surrendered by imposing the 1% penalty. In practice, if the risk management division within the insurance company cannot ‘use’ or gain access to the 1% fee, the required α insurance fee would need to be even higher.

5. Conclusion

In this paper we develop two approaches for analyzing a novel type of derivative security that has become quite popular in North America—called a Guaranteed Minimum Withdrawal Benefit (GMWB)—which is an insurance rider offered on Variable Annuity (VA) policies. VA policies are a close cousin to mutual funds, but offer additional performance-based guarantees. First, we take a *static* approach that assumes individual investors behave passively in utilizing the embedded guarantee. In this case, we show how the product can be bifurcated into a type of Quanto Asian Put (QAP) plus a generic term-certain annuity. At the other extreme we assume that investors are fully rational and seek to maximize the embedded option value by lapsing (a.k.a. surrendering or terminating) the product at the optimal time, i.e. once the expected present value of fees exceed the present value of benefits. We label this the *dynamic* approach, which leads to an optimal stopping problem akin to pricing an American put option, albeit complicated by the non-traditional payment structure.

Our main contribution lies in (i) bifurcating the product into its respective (simpler) derivative components, (ii) managing the non-traditional payment scheme, which is basis-points-of-assets versus up-front payments, (iii) introducing the distinction between two extreme valuation approaches, and (iv) discussing the optimal policy in this case. The less ‘rational’ the insurance company believes its retail market is, the less they have to charge for the guarantee. This is quite different from pricing an American option freely traded in the open market, where the seller of protection cannot afford the luxury of assuming less than full rationality on the part of the buyer. In fact, the VA market in the U.S.—which is a \$1 trillion dollar industry—has not had anywhere near the intense level of financial economic analysis and scrutiny compared to the mutual fund industry. Yet, this market provides a robust laboratory for testing theories of incomplete markets and frictions.

On a practical side, our numerical results indicate that the current practice (as of mid-2004) of charging between 30 and 50 basis points of assets on an ongoing basis for a typical 7% GMWB is not sufficient to cover the capital market hedging (No Arbitrage) cost of the guarantee, assuming a 20% volatility. This is regardless of whether we take a static actuarial or dynamic financial approach to the problem. In fact, given the long-dated nature of the embedded options, it is quite likely that the *pricing* (implied) volatility would be even higher, if the company chose to use the capital markets to offset its risk. It is therefore not surprising to see insurance companies impose restrictions on asset allocations to reduce the effective portfolio volatility and hence the cost of the embedded option.

And, although we used the ‘old technology’ of GBM assumptions, we believe that more sophisticated models of volatility and interest will only serve to increase the value of the embedded put option (and of course, complicate the problem of solving PDEs as well).

This under-pricing result stands in contrast to work by Milevsky and Posner (2001)—and widely cited in the *Wall Street Journal* in Clements (2004)—in which they show that the standard return-of-premium rider available on variable annuity policies is worth less than 5 basis points of assets. Clearly, then, these rider are quite heterogenous in value and so-called death benefits must be distinguished from living benefits.

We conclude by offering a number of justifications as to why insurance companies (in mid-2004) might be pricing the GMWB rider at lower than what we perceive to be the fair *capital markets* cost:

1. The insurance company assumes a high level of irrational lapsation and possible mortality that would somehow further reduce the required fee, as we described in Section 3.1. In theory of course, astute (young and healthy) buyers could purchase the VA with a GMWB and guarantee a \$7 payout, thus arbitraging the insurance company.
2. The GMWB rider/feature is being subsidized by the basic insurance fee. This is consistent with the over-pricing of standard features in VA policies. Thus, a 50 basis point fee for the rider together with a baseline 120 basis point fee might add-up to 170 basis points, which is enough to cover the risk.
3. The company uses a reserving methodology to manage the risk that differs from the capital market perspective. In other words, they do not hedge the risk using financial derivatives, but simply compute a premium based on a real-world probability of loss, as per our discussion in Section 2.1. In the insurance (actuary) company mind, the financial risk is much lower since they are not concerned with hedging payouts in every state of nature.
4. Finally, as we mentioned in a number of places during the development of our model, portfolio and asset allocation restrictions that some companies have begun to impose might serve to reduce the hedging cost of the embedded guarantee.

Of course, neither of the first three explanations will protect the company in the event a secondary market develops for these products and consumer rationality increases to the point where they exercise their options at the optimal time.⁵ In sum, future research will con-

⁵ In fact, a company by the name of *Coventry First* has recently started advertising a service by which they “rescue” VAs whose owners want to lapse or surrender the policy—which only entitles them to market value—by paying them a fraction of the embedded option value. CF then holds these products to maturity, gaining any intermediate cash flows, and thus arbitrage the insurance company.

tinue to examine an appropriate and realistic hedging strategy for GMWBs—and other more recent innovations on the economic border of life insurance and financial markets—in the presence of the usual collection of market imperfections.

Acknowledgments

The authors would like to thank seminar participants at the University of Michigan and University of Illinois and an anonymous IME reviewer for comments as well as Anna Abaimova for helpful research assistance.

Appendix A. Technical appendix

A.1. Pricing the quanto and computing the ruin probability

Our model requires a quick and robust method of computing two important quantities for which there are no analytic expressions. They are:

$$A(w, t) := E[W_t | W_t > 0, W_0 = w],$$

$$P(w, t) := \Pr[W_t \leq 0 | W_0 = w], \tag{A.1}$$

when W_t obeys the following stochastic differential equation (SDE):

$$dW_t = (\mu W_t - 1) dt + \sigma W_t dB_t, \quad W_0 = w, \tag{A.2}$$

where μ, σ are the drift and diffusion coefficients and B_t is the Brownian motion driving the process. We now provide a high-level sketch of the algorithm. More details are available in Huang et al. (2004). To start, $P(w, t)$ in Eq. (A.1) satisfies the Kolmogorov backward equation

$$P_t + (\mu w - 1)P_w + \frac{1}{2}\sigma^2 w^2 P_{ww} = 0 \tag{A.3}$$

with a terminal condition

$$P(w_T, T) = 1 - H(w_T - y), \tag{A.4}$$

where $H(w)$ is the Heaviside function and w_T is the wealth at T . This is a second order linear partial differ-

ential equation which we can numerically solve by a θ -method that can be written as follows:

$$\begin{aligned} & \frac{P_j^{(n+1)} - P_j^{(n)}}{\delta t} + (\mu w_j - 1) \\ & \times \left(\theta \frac{P_{j^*}^{(n+1)} - P_{j^*-1}^{(n+1)}}{\delta w} + (1 - \theta) \frac{P_{j^*}^{(n)} - P_{j^*-1}^{(n)}}{\delta w} \right) \\ & + \frac{\sigma^2 w_j^2}{2} \left(\theta \frac{P_{j+1}^{(n+1)} + P_{j-1}^{(n+1)} - 2P_j^{(n+1)}}{\delta w^2} \right. \\ & \left. + (1 - \theta) \frac{P_{j+1}^{(n)} + P_{j-1}^{(n)} - 2P_j^{(n)}}{\delta w^2} \right) = 0, \tag{A.5} \end{aligned}$$

where $P_j^{(n)}$ is a grid function which approximates $P(w, t)$ on the grid points (w_j, t_n) . A uniform grid with equal spacing δt and δx is used in our algorithm. The parameter θ can be arbitrarily selected, but when $\theta = 1/2$, it corresponds to the well-known second order Crank–Nicolson scheme. An upwind scheme is used for the first order derivative P_w , where the variable j^* is either j or $j + 1$, depending on the sign of the coefficient. For any implicit method where $0 < \theta \leq 1$, numerical boundary conditions must be provided on the computational boundaries $j = 0$ and J . This can be derived as:

$$P_0^n = 1, \quad j = 0 \quad \text{and} \quad P_J^n = 0, \quad j = J. \tag{A.6}$$

The cases $j = 0$ and J correspond to the $w_0 = 0$ and $w_J = W$ which are the boundaries of the truncated computation domain for calculating the probability W_t is less than some value at a *fixed* time. Likewise, for calculating the probability W_t is less than some value at *any* time, we use $j = 0$ and J with respect to the $w_0 = y$ and $w_J = W$. These are the boundaries of the truncated computation domain. The terminal condition is:

$$P_j^N = 1 - H(w_j - y). \tag{A.7}$$

With these boundary conditions and the terminal conditions the discrete equations can be solved by matching from time t_n to t_{n+1} , starting from $n = 0$. At t_{n+1} , the equations for $P_j^{(n+1)}$ can be arranged from Eq. (A.5). In this space, we can solve for all the probabilities by

iteration. For the expected value, which is $A(w, t)$ in Eq. (A.1), we can apply the same method.

References

- Bakshi, G., Madan, D., 2002. Average rate claims with emphasis on catastrophe loss options. *Journal of Financial and Quantitative Analysis* 37 (1), 93–115.
- Black, F., Scholes, M.S., 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81 (2), 736–1653.
- Benartzi, S., Thaler, R.H., 2001. Naïve diversification strategies in defined contribution saving plans. *American Economic Review* 91 (1), 79–98.
- Brennan, M.J., Schwartz, E., 1976. The pricing of equity-linked life insurance policies with an asset value guarantee. *Journal of Financial Economics* 3 (1), 195–213.
- Brown, J., Poterba, J., 2004. The household demand for variable annuities. In: Presented at American Economics Association Annual Meeting, San Diego.
- Boyle, P.P., Schwartz, E., 1977. Equilibrium prices of guarantees under equity-linked contracts. *Journal of Risk and Insurance* 44 (2), 639–680.
- Boyle, P.P., Boyle, F., 2001. *Derivatives: The Tools that Changed Finance*, Risk Books, Approx. 200 pp.
- Boyle, P.P., 2005. Variable annuity guarantees: the GMWB. In: Presentation at the Fourth Annual IFID Conference, Available for download at www.ifid.ca.
- Carr, P., 1998. Randomization and the American put. *Review of Financial Studies* 11 (3), 597–626.
- Clements, J., 2004. For a conservative investment, variable annuities are too costly. *Wall Street Journal* C1, January 21.
- Cox, S.H., Fairchild, J.R., Pedersen, H.W., 2004. Valuation of structured risk management products. *Insurance: Mathematics and Economics* 34 (2), 259–272.
- Dempster, M.A.H., Hutton, J.P., 1999. Pricing American stock options by linear programming. *Mathematical Finance* 9, 229–254.
- Dhaene, J.M., Denuit, M., Goovaerts, R., Kaas, D., Vyncke, 2001. The concept of comonotonicity in actuarial science and finance: applications. *Insurance: Mathematics and Economics* 31 (2), 133–161.
- Gerber, H.U., Shiu, E., 2003. Pricing dynamic investment fund protection. *North American Actuarial Journal* 4 (2), 60–77.
- Hardy, M., 2003. *Investment Guarantees: Modeling and Risk Management for Equity-linked Life Insurance*. John Wiley & Sons, New Jersey.
- Huang, H., Milevsky, M.A., Wang, J., 2004. Ruined moments in your life: how good are the approximations? *Insurance: Mathematics and Economics* 34 (3), 421–448.
- Ju, N., 1998. Pricing an American option by approximating its early exercise boundary as a multipiece exponential function. *Review of Financial Studies* 11 (3), 627–646.
- Karatzas, I., Shreve, S.E., 1992. *Brownian Motion and Stochastic Calculus*, 2nd ed. Springer-Verlag, New York.
- Milevsky, M.A., Posner, S., 2001. The titanic option: valuation of guaranteed minimum death benefits in variable annuities and mutual funds. *Journal of Risk and Insurance* 68 (1), 55–79.
- Milevsky, M.A., Posner, S., 1998. Asian options, the sum of lognormals and the reciprocal gamma distribution. *Journal of Financial and Quantitative Analysis* 33 (3), 409–422.
- Nielsen, J.A., Sandmann, K., 2003. Pricing bounds on Asian options. *Journal of Financial and Quantitative Analysis* 38 (2), 449–473.
- Pennacchi, G., 1999. The value of guarantees on pension fund returns. *The Journal of Risk and Insurance* 66 (2), 219–237.
- Persson, S.A., Aase, K., 1997. Valuation of the minimum guaranteed return embedded in life insurance products. *Journal of Risk and Insurance* 64, 599–617.
- Sheris, M., 1995. The valuation of option features in retirement benefits. *The Journal of Risk and Insurance* 62 (3), 509–534.
- Turnbull, S., Wakeman, L.M., 1991. A quick algorithm for pricing European average options. *Journal of Financial and Quantitative Analysis* 26 (1), 377–389.
- Windcliff, H., Forsyth, P.A., Vetzal, K.R., 2001. Valuation of segregated funds: shout options with maturity extensions. *Insurance: Mathematics and Economics* 29 (2), 1–21.