Optimal Management of an Insurer's Exposure in a Competitive General Insurance Market

Paul Emms* and Steven Haberman†

Abstract

The qualitative behavior of the optimal premium strategy is determined for an insurer in a finite and an infinite market using a deterministic general insurance model. The optimization problem leads to a system of forward-backward differential equations obtained from Pontryagin's Maximum Principle. The focus of the modelling is on how this optimization problem can be simplified by the choice of demand function and the insurer's objective. Phase diagrams are used to characterize the optimal control. When the demand is linear in the relative premium, the structure of the phase diagram can be determined analytically. Two types of premium strategy are identified for an insurer in an infinite market, and which is optimal depends on the existence of equilibrium points in the phase diagram. In a finite market there are four more types of premium strategy, and optimality depends on the initial exposure of the insurer and the position of a saddle point in the phase diagram. The effect of a nonlinear demand function is examined by perturbing the linear price function. An analytical optimal premium strategy is also found using inverse methods when the price function is nonlinear.

1. Introduction

The actuarial price of a general insurance policy is calculated using a premium principle (Rolski et al. 1999), which relates the premium to the potential claims on the policy. However, many lines of general insurance are highly competitive, and this affects the price that insurers set for a policy. Indeed, a cycle is often observed featuring periods in which insurers price policies below and then above the actuarial price (Daykin et al. 1994). To understand and predict this cycle, one needs to model the competitive nature of insurance pricing.

Taylor (1986) formulates a competitive demand model to price general insurance policies using a single state equation governing the evolution of the insurer's exposure:

$$q_j = q_{j-1} F(p_j, \bar{p}_j), \tag{1.1}$$

where q_j is the exposure in year j, p_j is the insurer's premium, $\bar{p}_j$ is the market average premium, and F is a demand function. All premiums are measured per unit of exposure. Taylor also assumes that the objective of the insurer is to maximize its wealth after J years defined by

$$E = \sum_{j=1}^{J} v^{j-\frac{1}{2}} q_j (p_j - \pi_j), \tag{1.2}$$

*

†

where π_j is the breakeven premium in year j and v is a discount factor. Here the market average premium $\bar{p}_j$ and the break-even premium π_j ($j = 1, 2, \ldots J$) are assumed to be given now. This is a deterministic, discrete, optimal control problem (Sethi and Thompson 2000) where the controls are the p_j's and the objective is to maximize the terminal wealth of the insurer.

From (1.1), the exposure in year j is proportional to the exposure in the previous year, and from (1.2), the objective is a linear function of the exposures in years $j = 1, 2, \ldots J$. It is these two assumptions that allow Taylor to deduce the optimal premium strategy using a recurrence relation stepping backwards in time from the end of the planning horizon $j = J$. If we substitute (1.1) into (1.2), then the first-order conditions are independent of the current exposure q_0. In addition, the Jth first-order condition yields the terminal excess premium

$$p_J - \pi_J = -\frac{1}{\partial \log F/\partial p_J},$$

which can be determined implicitly *now* independently of the other premium values. The terminal premium $p_J > \pi_J$ because we expect that the demand for policies decreases as the premium increases. If the second-order condition holds, then this gives the (locally) optimal terminal excess premium as proportional to the inverse of the elasticity of demand. The insurer knows the optimal premium it should charge at termination, but must calculate the intervening premium values recursively, and this leads to a premium strategy independent of the current exposure. This means the optimal premium strategy is identical for all insurers in the market irrespective of their current size.

Emms and Haberman (2005) formulate a continuous-time version of Taylor's model, in which premium rates are held fixed over the course of the policy. The exposure equation in Emms and Haberman takes the form

$$\frac{dq}{dt} = q(G(q, p, \bar{p}) - \kappa), \tag{1.3}$$

where G is the fractional rate of generation of exposure through policy sales and renewals, κ^{-1} is the mean length of policies, and $\bar{p}$ is a process modeling the market average premium. Emms and Haberman suppose that G is independent of the exposure, so that the state equation (1.3) is linear in q. If no more insurance is sold at time t, then the exposure decreases exponentially at rate κ to account for policies currently in force. Notice that the change in exposure is split into two terms in order that there is an explicit expression for the generation of wealth by selling policies.

In practice, insurance companies have finite capacity for exposure because of the capital regulations necessary to cover potential claims (Cummins and Outreville 1987; Doherty and Garven 1995). In addition, new business is obtained at the expense of other insurers or from growth in the insurance market. This means that there is a limit to the amount of new business that an insurer can underwrite in a year. Consequently, the change in the exposure of the insurer cannot be proportional to its current exposure as the size of the insurer increases. The demand models in Taylor (1986) and Emms and Haberman (2005) assume that the market is not saturated with policies, and that there is sufficient new business available to take up the competitively priced policies offered by the insurer.

Many other volume/demand relationships are contained in the management science literature (Lilien and Kotler 1983). The Bass demand model (Bass 1969) is similar to Taylor's insurance model. Bass showed that the probability of an individual adopting a product can be represented as a linear function of the previous number of adopters. The Dolan-Jeuland model (Dolan and Jeuland 1981) assumes that the sales rate is proportional to the previous volume times the volume remaining in the market. This is an idea suggested by Robinson and Lakhani (1975), who calculate the optimal pricing strategy using dynamic programming.

Thus, in contrast to previous competitive insurance pricing models, we consider a more general parameterization for G in (1.3), which models a finite market for insurance policies. The pricing model is based on the paper of Emms (2007a), which uses premium values rather than rates but leads to an exposure equation that is identical to (1.3). However, the interpretation of the equation is changed. Each policy is of fixed length κ^{-1}, for which the insurer charges a premium p (per unit of exposure). Customers who renew their policies are treated as new policyholders. For tractability, we focus on a separable demand function, which leads to two parameterized functions: one for the exposure and one for the relative premium.

The resulting model is similar in form to the retail pricing model of Kalish (1983), who also considers a separable demand function for new retail products. Our work differs from his paper in that we consider demand as a function of the relative price rather than absolute price in order to model the competition in the insurance market. This formulation requires that the insurer remains relatively small, and so does not affect the market price of insurance. We also consider a more general objective than the maximization of terminal wealth because it is often found that loss-leading is optimal, and this is not a desirable objective for an insurer. Furthermore, we allow the market price to drift and lag the exposure of the insurer in order to model the loss of policies.

Section 2 introduces the deterministic insurance model. In Section 2.1 Pontryagin's Maximum Principle gives the necessary conditions for an optimal control and leads to a system of forward–backward differential equations. At this stage we pose the optimization problem in general terms and focus on how particular parameterizations simplify the problem. A new adjoint variable is defined in Section 2.2, which makes it clearer just how the optimization problem can be reduced. In general, the system of forward–backward differential equations must be solved numerically, and it is difficult to ascertain the properties of the optimal control without recourse to an exhaustive numerical study.

Instead, we study the phase diagram of the system of differential equations (Jordan and Smith 1977) in Section 3 using a linear price function. Phase diagrams are a powerful technique used to understand the qualitative features of a dynamical system. Each diagram consists of a number of phase paths, which correspond to a particular solution of the system of differential equations. The structure of phase space is determined by the equilibrium point(s) of the differential equations. By finding the type of each equilibrium point, we can classify the features of the optimal control (Léonard and van Long 1992).

A linear stability analysis in Section 3.1 gives the type of the equilibrium point(s) as a function of the model parameters for a terminal wealth objective. We find the equilibrium points for the total wealth objective in Section 3.2. The numerical computation of the phase diagrams is described for both objectives in Section 4 using a particular parameter set. Section 5 investigates how nonlinear parameterizations of demand affect the optimal strategy. Conclusions can be found in Section 6.

2. CONTINUOUS MODEL

First, we generalize the demand law in Emms and Haberman (2005) and adopt a separable parameterization following Kalish(1983):

$$G(q, k) = f(q)g(k), \tag{2.1}$$

where the premium relative to the market average premium is

$$k = \frac{p}{\bar{p}},$$

and both f and g are nonnegative decreasing functions of their respective arguments. We call f the exposure function and g the price function, and we adopt the convention that g is only defined to within a multiplicative constant. We use the notation of a prime ($'$) to denote the derivative of a function with respect to its argument so that no confusion can arise.

Notice that we use the current exposure of the insurer in the demand parameterization rather than the total volume of sales as adopted by Kalish (1983). Thus, we suppose the status of the insurer is measured by its present exposure rather than including policies that have expired. If we set $\kappa = 0$, then we obtain a model similar to Kalish (1983) for the pricing of new products because then policies are of infinite length, and so the current exposure does represent total sales.

Emms and Haberman (2005) consider the case

$$f \equiv a = \text{const.}, \qquad g = (b - k)^+.$$

Thus, the rate of increase in exposure is directly proportional to the current size of the insurer as measured by its exposure in the marketplace. If the insurer has a maximum possible exposure q_m, then a more plausible parameterization for the exposure function is

$$f(q) = a\left(1 - \frac{q}{q_m}\right)^+. \tag{2.2}$$

Once the insurer's exposure reaches the saturation value, q_m, then it is unable to achieve further market penetration through setting a competitive premium. Notice that q_m is not the total possible exposure in the market for a particular line of insurance in contrast to Robinson and Lakhani (1975), who considered the retail market. Our model differs from theirs because policies are contracts that hold for a fixed length of time. Policyholders must either break their contract or await renewal in order to take up a lower-priced policy at another insurer.

We assume that the market prices insurance according to the expected value principle, so that the market average premium is

$$\bar{p}(t) = (1 + \theta)\pi(t),$$

where θ is the constant market loading and $\pi(t)$ is the (given) break-even premium including expenses. The ratio of the breakeven premium to the market average premium is therefore constant:

$$\gamma := \frac{\pi}{\bar{p}} = \frac{1}{1 + \theta}. \tag{2.3}$$

It is not difficult to generalize the model by considering a time-dependent loading, but this leads to further unknown parameters. The drift in the market average premium (and the break-even premium) is defined by

$$\mu(\bar{p}, t) := \frac{1}{\bar{p}}\frac{d\bar{p}}{dt}. \tag{2.4}$$

Suppose the state of the insurer is described by its current exposure q and its wealth w. From (1.3), the sale of insurance policies over time δt generates exposure $\delta q = qG\delta t$, which in turn changes the wealth of the insurer by $\delta w = (p - \pi)\delta q = (p - \pi)qG\delta t$. Thus, following Emms (2007a), we suppose the state evolves according to

$$\frac{dq}{dt} = q(f(q)g(k) - \kappa), \tag{2.5}$$

$$\frac{dw}{dt} = -\alpha w + \pi(k\gamma^{-1} - 1)qf(q)g(k), \tag{2.6}$$

using (2.3), and where α represents the loss of wealth due to returns to shareholders. Let us define the objective function by

$$J(t) = \int_t^T U_1(w(t))\, dt + U_2(w(T)),$$

where the utility functions U_1 and U_2 are concave functions of wealth and T is the planning horizon.

The insurance pricing problem is now in the form of a deterministic optimal control problem (Sethi and Thompson 2000). The objective of the insurer is to maximize $J(0)$ over the control set U, which consists of all finite relative premium strategies $k(t)$. For simplicity, we do not impose any constraints on the control or state (Emms 2007b). This objective is based solely on the wealth of the insurer over the planning horizon or its wealth at termination: we classify the problem as a terminal wealth problem if $U_1 \equiv 0$ or a total wealth problem if $U_2 \equiv 0$. In the total wealth problem, we ignore the time value of money because other studies (Emms and Haberman 2005) have found that the optimal strategy is relatively insensitive to discounting.

2.1 Maximum Principle

Following the definition in Sethi and Thompson (2000), the Hamiltonian for this general insurance model is

$$H(q, w, \lambda_1, \lambda_2, k) = U_1(w) + \lambda_1 q(f(q)g(k) - \kappa) + \lambda_2(-\alpha w + q\pi(k\gamma^{-1} - 1)f(q)g(k)), \quad (2.7)$$

where λ_1 and λ_2 are the adjoint variables for q and w, respectively. If $g'(k) = 0$ at an extremum of H, then the first-order condition gives $g(k) = 0$, which determines k. If the extremum occurs at a point where $g'(k) < 0$, then the first-order condition yields

$$k + \frac{g(k)}{g'(k)} = \gamma\left(1 - \frac{\lambda_1}{\pi\lambda_2}\right). \quad (2.8)$$

Providing the right-hand side is defined and the equation has a real root, then this gives a well-defined value of k. The second-order condition for a (local) maximum of H is

$$H_{kk} = (qf)(\lambda_1 g'' + \lambda_2\pi(2\gamma^{-1}g' + (k\gamma^{-1} - 1)g'')) < 0, \quad (2.9)$$

which becomes on applying (2.8)

$$\lambda_2\left(2 - \frac{g''g}{g'^2}\right) > 0,$$

providing $g'(k) < 0$ and g is sufficiently differentiable.

The adjoint equations of the Maximum Principle are

$$\frac{d\lambda_1}{dt} = -\lambda_1((qf)'g - \kappa) - \lambda_2\pi(k\gamma^{-1} - 1)(qf)'g, \quad (2.10)$$

$$\frac{d\lambda_2}{dt} = -U_1'(w) + \alpha\lambda_2, \quad (2.11)$$

with transversality conditions

$$\lambda_1(T) = 0, \qquad \lambda_2(T) = U_2'(w(T)).$$

For the terminal wealth problem, $U_1 \equiv 0$ and so

$$\lambda_2(T) = U_2'(w(T))e^{\alpha(t-T)} > 0,$$

since utility functions increase with wealth (Gerber and Pafumi 1998). Similarly, for the total wealth problem, $U_2 \equiv 0$ and $U_1'(w) > 0$, and so from the adjoint equation (2.11)

$$\frac{d(e^{-\alpha t}\lambda_2)}{dt} < 0.$$

If λ_2 is smooth and continuous, then $e^{-\alpha t}\lambda_2 \geq 0$ and so $\lambda_2 \geq 0$. Providing $\lambda_2 \neq 0$, then the second-order condition for a (local) maximum is

$$\frac{g''g}{g'^2} < 2, \quad (2.12)$$

which is similar to the result given by Taylor (1986) and Kalish (1983). Taylor's condition for a maximum is for a discrete system and a slightly different insurance model, so the result is not directly comparable. His condition requires that the value function

$$E_i = \sum_{j=i}^{J} v^j(p_j - \pi_j) \exp \sum_{k=1}^{j} \log F(p_k) > 0.$$

The value function is dependent on the demand function F, and it is not clear what further restrictions this places on that function. For a sufficiently large discount factor v, E_i is positive, but, as Taylor states, this is preventing loss-leading today if tomorrow's profits are heavily discounted. Kalish (1983) considers the demand to be a function of price rather than relative price, so the derivatives in (2.12) are replaced with derivatives with respect to price.

Condition (2.12) is a necessary condition for (2.8) to yield a local maximum of H. The local maximum must also yield the global maximum for H over the control set U in order that the conditions of the Maximum Principle are satisfied. Thus, we can only say that these strategies are locally optimal. The same statement applies to the strategies found in Taylor's paper. Moreover, the Maximum Principle itself is a set of *necessary* conditions, which the optimal control and state trajectory must satisfy. For the current problem, if the control k and state (q, w) satisfy the Maximum Principle and the Hamiltonian is concave in the control and state variables, then this trajectory is the optimal premium strategy (see Sethi and Thompson 2000, p. 64).

2.2 Relative Marginal Cost of Underwriting

The form of (2.8) suggests that we introduce the variable

$$\Lambda = \frac{\lambda_1}{\bar{p}\lambda_2} = \frac{\gamma\lambda_1}{\pi\lambda_2},$$

so that the first-order condition can be written

$$k - \gamma = -\frac{g}{g'} - \Lambda. \tag{2.13}$$

From the adjoint equations (2.10), (2.11), and the definition of the drift (2.4) we obtain

$$\begin{aligned}\frac{d\Lambda}{dt} &= \frac{1}{\bar{p}\lambda_2}\frac{d\lambda_1}{dt} - \frac{\Lambda}{\lambda_2}\frac{d\lambda_2}{dt} - \Lambda\mu(\bar{p}, t)\\ &= \Lambda\left(\frac{U_1'(w)}{\lambda_2} - (qf)'g + \kappa - \mu(\bar{p}, t)\right) - g(qf)'(k - \gamma)\\ &= \Lambda\left(\frac{U_1'(w)}{\lambda_2} + \kappa - \alpha - \mu(\bar{p}, t)\right) + \frac{g^2(qf)'}{g'}.\end{aligned} \tag{2.14}$$

For the terminal wealth problem, the transversality condition for this new adjoint variable is $\Lambda(T) = 0$ from the boundary conditions of (2.10) and (2.11). For the total wealth problem, $\lambda_1(T) = \lambda_2(T) = 0$. However, Taylor series expansions of λ_1, λ_2 yield $\Lambda(t) \approx \frac{1}{2}\lambda_1''(T)(T - t)/\bar{p}(T)U_1'(w(T))$ for $t \approx T$ so that $\Lambda(T) = 0$. Thus, provided that the first-order condition (2.13) yields the control that maximizes the Hamiltonian, and the adjoint variables are sufficiently smooth, then the terminal optimal premium for the terminal wealth problem is identical to the terminal optimal premium for the total wealth problem.

The dimensionless adjoint variable Λ is the marginal cost of underwriting a unit of exposure relative to the marginal change in wealth generated by underwriting. Thus, we call Λ the relative marginal cost of underwriting. We can see this identification explicitly by relating each adjoint variable with the value function defined by

$$V(q, w, t) = \max_{k \in U} J(t).$$

From Yong and Zhou (1999, p. 229) we have

$$V_q = -\lambda_1, \qquad V_w = -\lambda_2,$$

so that

$$\Lambda = \frac{V_q}{\bar{p} V_w}.$$

Conventionally, the elasticity of demand is defined as the relative change in demand divided by the relative change in price (Lilien and Kotler 1983). It is a dimensionless quantity, which here we can identify with

$$\frac{1}{g}\frac{dg}{dk}.$$

This dimensionless expression is the relative change in demand with respect to the change in *relative premium*. The identification is appropriate because the demand law is a function of the relative premium rather than just the premium. Consequently, $-g/g'$ is the inverse elasticity of demand, and it is positive because $g < 0$. Therefore, the terminal premium is greater than ~~break even~~ from the transversality condition as in Taylor's model (Taylor 1986). Loss-leading occurs when $k < \gamma$ or

$$\Lambda > -\frac{g}{g'} > 0.$$

The economic interpretation of the first-order condition (2.13) is a nondimensional relationship that must be satisfied along an optimal state trajectory. It says that at each instant $k - \gamma$, which is

$$\text{excess relative premium} = \frac{1}{\text{elasticity of demand}} - \text{relative marginal cost of underwriting}.$$

When there is significant loss-leading, the relative marginal cost of underwriting is very large and substantial capital is required to finance borrowing. In view of the transversality condition, at the end of planning horizon the optimal excess relative is equal to 1/(elasticity of demand). So for an inelastic line of insurance, the optimal terminal relative premium is very large. This is because in a very inelastic market, it is difficult to generate sales from lower prices, so that it is optimal to leave the market by setting a very high premium.

Now, by differentiating the first-order condition (2.13) with respect to time, we find

$$\frac{d\Lambda}{dk} = \frac{gg''}{g'^2} - 2, \tag{2.15}$$

so that from (2.12), the adjoint variable Λ increases as the relative premium k decreases. Moreover, from (2.14) we can find an explicit expression for the evolution of the relative premium

$$\frac{dk}{dt}\left(2 - \frac{gg''}{g'^2}\right) = \left(k - \gamma + \frac{g}{g'}\right)\left(\frac{U_1'(w)}{\lambda_2} + \kappa - \alpha - \mu(\bar{p}, t)\right) - (qf)'\frac{g^2}{g'}.$$

At the end of the planning horizon, the terminal premium $k(T)$ is given implicitly by the root of $k - \gamma = -g/g'$ because $\Lambda(T) = 0$. In addition, we know how the optimal premium approaches the terminal value, because at termination

$$\frac{dk}{dt} = (qf)'\left(\frac{g^2(-g')}{2g'^2 - gg''}\right).$$

Thus, if $(qf)'(T) < 0$, then $dk/dt < 0$ at $t = T$, and the optimal relative premium decreases toward the terminal value $k(T)$. If the market is infinite, then $(qf)'(T) = a$, and so this expression yields an explicit Taylor series approximation for the optimal control near termination.

3. Linear Price Function

Henceforth we neglect the return to shareholders by setting $\alpha = 0$ because Emms and Haberman (2005) find that this parameter has little impact on the optimal strategy. Emms and Haberman take the demand as linearly dependent on the relative premium:

$$g(k) = (b - k)^{+}. \tag{3.1}$$

Let us define

$$k^{i} = \frac{1}{2}(b + \gamma - \Lambda), \tag{3.2}$$

using the first-order condition for a maximum of the Hamiltonian (2.13). If $k^i < b$, then k^i is the control that maximizes the Hamiltonian; we call this control interior. If $k^i \geq b$, then it is optimal not to sell insurance. If the control is interior, then the corresponding demand function is

$$G^{i} = \tfrac{1}{2}f(q)(b - \gamma + \Lambda). \tag{3.3}$$

If $k^i < \gamma$, then the interior premium is less than ~~break even~~ $p^i < \pi$, and so the strategy leads to a loss at that instant. In terms of the adjoint variable loss-leading strategies satisfy

$$\Lambda > b - \gamma. \tag{3.4}$$

If the control is interior, then the two adjoint equations can be written in terms of Λ and λ_2:

$$\frac{d\Lambda}{dt} = \Lambda\left(\frac{U_1'(w)}{\lambda_2} + \kappa - \mu(\bar{p}, t)\right) - \tfrac{1}{4}(b - \gamma + \Lambda)^2(qf)', \tag{3.5}$$

$$\frac{d\lambda_2}{dt} = -U_1'(w), \tag{3.6}$$

with transversality conditions

$$\Lambda(T) = 0, \qquad \lambda_2(T) = U_2'(w(T)).$$

For given U_1, U_2 we have a system of forward-backward differential equations: the state equations are integrated forwards from $t = 0$, whereas the adjoint equations are integrated backwards from $t = T$. In general, there are four dependent variables: q, w, λ_1, and λ_2, which means the phase diagram is four dimensional.

3.1 Terminal Wealth

For $U_1 \equiv 0$ we obtain the optimization problem for maximizing the terminal utility of wealth and Λ is independent of λ_2, the wealth w, and the choice of utility function U_2 (see [3.5]). Moreover, the equation for Λ is autonomous if we suppose the drift in $\bar{p}$ is

$$\mu(\bar{p}, t) = \mu = \text{constant}.$$

Consider the phase plane of the exposure q and the adjoint Λ given by

$$\frac{dq}{dt} = X(q, \Lambda) := q(\tfrac{1}{2}f(b - \gamma + \Lambda) - \kappa), \tag{3.7}$$

$$\frac{d\Lambda}{dt} = Y(q, \Lambda) := -\tfrac{1}{4}(qf)'\Lambda^2 - \Lambda(\tfrac{1}{2}(qf)'(b - \gamma) - \kappa + \mu) - \tfrac{1}{4}(qf)'(b - \gamma)^2. \tag{3.8}$$

If we identify a trajectory passing through the q-axis and move backwards in time T years along that trajectory, then we arrive at an initial exposure q_0 and an optimal premium strategy. Consequently the phase plane yields the optimal strategies for a given parameter set, and it is a good way to determine the qualitative form of these strategies.

The structure of the phase diagram is determined by its equilibrium points (Smith 1985), which are given by $X = Y = 0$ or[1]

$$q(\tfrac{1}{2}a^{-1}f(\phi + \Lambda) - \psi) = 0, \tag{3.9}$$

$$\tfrac{1}{4}\Lambda^2 + \Lambda\left(\tfrac{1}{2}\phi + \frac{\zeta - \psi}{a^{-1}(qf)'}\right) + \tfrac{1}{4}\phi^2 = 0, \tag{3.10}$$

where we have have introduced the dimensionless parameters

$$\phi = b - \gamma > 0, \qquad 0 < \psi = \frac{\kappa}{a} < 1, \qquad \zeta = \frac{\mu}{a}.$$

The parameter ϕ is a measure of the extent an insurer can raise its premium above ~~break even~~ and still obtain positive demand for its policies, while ψ indicates how much the exposure increases for a given relative premium. The nondimensional drift in the market average premium, ζ, can be positive or negative. In an infinite market the number of free parameters can be reduced further by introducing the parameter

$$\rho = \zeta - \psi.$$

The linear stability of an equilibrium point (q_e, Λ_e) can be determined by considering the local coordinate system

$$\xi = q - q_e,$$

$$\eta = \Lambda - \Lambda_e,$$

following Jordan and Smith (1977, p. 51). Using a Taylor series expansion we have

$$\frac{1}{a}\frac{d}{dt}\begin{pmatrix}\xi\\ \eta\end{pmatrix} = \begin{pmatrix} m_{11} & m_{12}\\ m_{21} & m_{22}\end{pmatrix}\begin{pmatrix}\xi\\ \eta\end{pmatrix}, \tag{3.11}$$

where the elements of the matrix $M = (m_{ij})$ are

$$m_{11} = a^{-1}X_q(q_e, \Lambda_e) = \tfrac{1}{2}a^{-1}(qf)'(q_e)(\phi + \Lambda_e) - \psi,$$

$$m_{12} = a^{-1}X_\Lambda(q_e, \Lambda_e) = \tfrac{1}{2}a^{-1}(qf)(q_e),$$

$$m_{21} = a^{-1}Y_q(q_e, \Lambda_e) = -\tfrac{1}{4}a^{-1}(qf)''(q_e)(\phi + \Lambda_e)^2,$$

$$m_{22} = a^{-1}Y_\Lambda(q_e, \Lambda_e) = -\tfrac{1}{2}a^{-1}(qf)'(q_e)(\phi + \Lambda_e) - \rho.$$

Solutions of the linear system (3.11) have the form

$$\begin{pmatrix}\xi\\ \eta\end{pmatrix} = \begin{pmatrix}r\\ s\end{pmatrix}e^{a\omega t},$$

which on substituting into (3.11) shows that the eigenvalues of the matrix M determine the stability of the equilibrium point. In terms of the matrix coefficients, the two eigenvalues for each equilibrium point are

[1] In Kalish (1983) equilibrium points occur where the sales rate is zero, for example, when the market is saturated with goods. Here the equilibrium points occur when the sale of exposure equals the loss of exposure due to nonrenewal.

$$\omega_{\pm}(q_e, \Lambda_e) = \tfrac{1}{2}(m_{11} + m_{22} \pm ((m_{11} - m_{22})^2 + 4m_{12}m_{21})^{1/2}). \tag{3.12}$$

It can be seen that the eigenvalues are a function of the exposure function f and the three nondimensional parameters ϕ, ψ, and ζ. It is this parameterization and these three parameters that determine the type of equilibrium point.

3.1.1 Infinite Market

If $f = a$ corresponding to an infinite market, then $(qf)' = a$. In this case the adjoint equation (3.8) uncouples from the state equation (3.7), and it has the analytical solution given in the Appendix. The optimal premium strategy can then be determined from (3.2) providing k remains interior.

From (3.9) and (3.10) there are at most two equilibrium points in the phase plane given by $q = 0$ and the real roots of

$$\tfrac{1}{4}\Lambda^2 + \Lambda(\tfrac{1}{2}\phi + \rho) + \tfrac{1}{4}\phi^2 = 0. \tag{3.13}$$

If the discriminant of this quadratic

$$\Delta^0 = \rho(\rho + \phi) > 0, \tag{3.14}$$

then there are two real roots

$$\Lambda^0_{\pm} = -2\rho - \phi \pm 2\sqrt{\Delta^0} \tag{3.15}$$

with $\Lambda^0_+ \geq \Lambda^0_-$. These roots are either both positive or both negative. We can use the equilibrium points to characterize the optimal premium strategy in a similar way to Emms (2007b).

Figure 1*a* shows the behavior of these roots as we vary ρ with ϕ held fixed. It is easy to show that there is the symmetry $\Lambda^0_{\pm}(\rho, \phi) = -\Lambda^0_{\mp}(-\rho - \phi, \phi)$ as shown in the figure, and that the roots lie below $\Lambda = 0$ if $\rho > 0$ and above $\Lambda = 0$ if $\rho < -\phi$. Figure 1*b* shows the evolution of Λ using graphical arguments similar to those in Emms (2007b). The direction of the arrows shows that, for the optimal strategy, Λ always decreases as one integrates forwards in time toward $\Lambda = 0$. This means that the optimal relative premium is always increasing.

If $\Delta^0 < 0$, then there are no real roots and $d\Lambda/dt < 0$. Consequently, on any trajectory, k increases over the planning horizon. If the planning horizon T is sufficiently large, then by the form of (A.3) in the Appendix, we can see that the optimal strategy may be loss-leading because negative premium values are possible. In addition, if the premium is constrained to positive values (Emms, 2007b), then the optimal strategy, should it exist, may not be smooth.

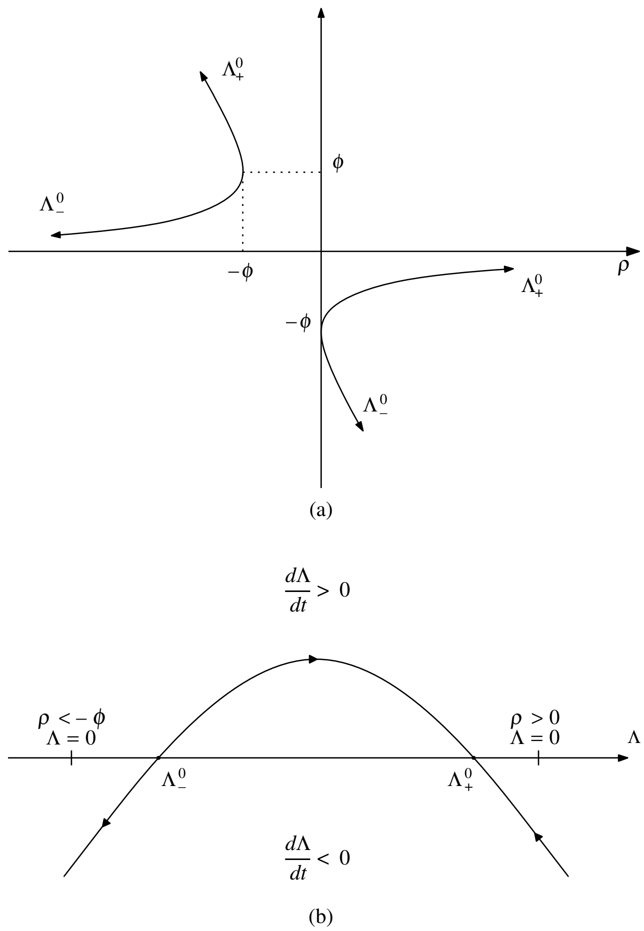
In the (q, Λ) plane the type of each equilibrium point can be determined from the matrix M, which now has simplified coefficients:

$$m_{11} = \tfrac{1}{2}(\phi + \Lambda^0_{\pm}) - \psi, \qquad m_{12} = m_{21} = 0, \qquad m_{22} = -m_{11} - \zeta. \tag{3.16}$$

The eigenvalues for each equilibrium point are just $\omega_+ = m_{11}$ and $\omega_- = m_{22}$ from (3.12). We expect that $\zeta \ll 1$ because the drift in the market average premium should be much less than the growth in the exposure. The remaining parameters are all $O(1)$, which means that we expect that m_{11}, m_{22} are $O(1)$ also. As a result the eigenvalues for both equilibrium points differ in sign, and therefore both are unstable saddle points. The matrix M is diagonal, so the normalized eigenvectors for both equilibrium points are $(0, 1)^T$ and $(1, 0)^T$. Consequently the axes of the saddle points are in the direction of the q and Λ axes.

In summary, the strategy can be classified according to the sign of the discriminant Δ^0 and the position of the equilibrium points in relation to the axis $\Lambda = 0$. If $-\phi < \rho < 0$, then there are no equilibrium points, and blow-up can occur within the planning horizon. If $\rho < -\phi$, then all optimal trajectories tend to Λ^0_- as $t \to -\infty$ providing $\Lambda^0_- < b + \gamma$ so that the control remains positive. If $\rho > 0$,

Figure 1
Qualitative Behavior of Adjoint Variable Λ for an Insurer Maximizing Terminal Wealth in an Infinite Market



Notes: (*a*) Position of the two roots $\Lambda_\pm^0$ as ρ is varied for fixed ϕ. (*b*) Behavior of adjoint variable Λ as time varies. The arrows in (*b*) show the direction of increasing time *t*.

then the optimal trajectories do not tend to a limit as $t \to -\infty$. For simplicity, we define type A strategies as those which do not have this limiting behavior, and type B strategies as those that do.

3.1.2 Finite Market

For a finite market we have from (2.2)

$$(qf)' = a\left(1 - \frac{2q}{q_m}\right) \quad \text{for } q < q_m, \tag{3.17}$$

and we find that the phase diagram has additional equilibrium points. Henceforth we shall omit the caveat that $q < q_m$.

If $q = 0$, then $(qf)' = a$, and the equilibrium equations (3.9) and (3.10) correspond to the infinite market case. Consequently, providing that $\Delta^0 > 0$, then there are two equilibrium points at $(0, \Lambda_\pm^0)$ in the finite market case as well, and these are both saddle points with their axes in the direction of the q and Λ axes.

If $q \neq 0$, then we can eliminate q from (3.10) and obtain a cubic for Λ:[2]

$$(\Lambda + \phi)(\Lambda^2 + \Lambda(2\phi - 4\zeta) + \phi(\phi - 4\psi)) = 0.$$

The first root $\Lambda = -\phi$ is not relevant because it corresponds to infinite demand for policies. The remaining two roots are

$$\Lambda_\pm = 2\zeta - \phi \pm 2(\zeta^2 - \phi\zeta + \phi\psi)^{1/2}, \tag{3.18}$$

$$q_\pm = q_m\left(1 - \frac{2\psi}{\phi + \Lambda_\pm}\right), \tag{3.19}$$

and these yield distinct real-valued equilibrium points providing the discriminant

$$\Delta = \zeta^2 - \phi\zeta + \phi\psi > 0.$$

These new equilibrium points are relevant in the phase plane if $\Lambda_\pm$ yield a relative premium in the range $0 < k < b$ so that the control is positive and interior. This restriction requires that

$$-\phi < \Lambda_\pm < b + \gamma, \tag{3.20}$$

from (3.2). In addition, the equilibrium points are of interest only if $q_\pm > 0$ because the exposure of the insurer must be non-negative. Thus, from (3.18) and (3.19) we require

$$-\rho < \pm(\zeta^2 - \phi\zeta + \phi\psi)^{1/2}.$$

If $\rho > 0$, then q_+ trivially satisfies this condition, whereas for $q_- > 0$ we require

$$\frac{\psi}{\phi}\left(1 + \frac{\zeta}{\rho}\right) < 1. \tag{3.21}$$

If $\rho < 0$, then q_- fails to satisfy this condition, and for $q_+ > 0$ we obtain (3.21) again. In the forthcoming numerical work we shall focus on the case that $\rho < 0$, so that only (q_+, Λ_+) is present in the phase diagram provided (3.20) and (3.21) are satisfied and $\Delta > 0$. Thus, we suppose the drift in the market average premium $\mu < \kappa$, which rules out large increases in $\bar{p}$ over the planning horizon.

The linear stability of the equilibrium points is determined by the matrix M, which now has coefficients

$$m_{11} = \tfrac{1}{2}\left(1 - \frac{2q_e}{q_m}\right)(\phi + \Lambda_e) - \psi, \qquad m_{12} = \tfrac{1}{2}q_e\left(1 - \frac{q_e}{q_m}\right),$$

$$m_{21} = \frac{1}{2q_m}(\Lambda_e + \phi)^2, \qquad m_{22} = -\tfrac{1}{2}\left(1 - \frac{2q_e}{q_m}\right)(\Lambda_e + \phi) - \rho.$$

Using the equilibrium value for $q_e = q_\pm$ in (3.19), we can rewrite these expressions in terms of the adjoint equilibrium variables $\Lambda_\pm$ and obtain

$$m_{11} = \tfrac{1}{2}(2\psi - \phi - \Lambda_\pm), \qquad m_{12} = \frac{q_m\psi}{\phi + \Lambda_\pm}\left(1 - \frac{2\psi}{\phi + \Lambda_\pm}\right),$$

$$m_{21} = \frac{1}{2q_m}(\Lambda_\pm + \phi)^2, \qquad m_{22} = \tfrac{1}{2}(\phi + \Lambda_\pm) - \zeta - \psi.$$

The two eigenvalues for each equilibrium point are denoted by $\omega_\pm(q_\pm, \Lambda_\pm)$ and are determined from (3.12) using (3.18).

[2] Notice that we can rearrange (3.10) as

$$(\Lambda + \phi)^2 = \frac{4\Lambda(\psi - \zeta)}{a^{-1}(qf)'} = \frac{4\Lambda(\psi - \zeta)(\Lambda + \phi)}{(4\psi - \phi - \Lambda)},$$

using (3.17) with $q/q_m = 1 - 2\psi/(\phi + \Lambda)$ from (3.9).

The expressions for the eigenvalues $\omega_{\pm}(q_+, \Lambda_+)$ are rather complicated and difficult to analyze. Consequently we plot $\omega_{\pm}(q_+, \Lambda_+)$ as a function of the three nondimensional parameters ϕ, ζ, and ψ in Figure 2 to understand how the equilibrium point moves in the phase diagram as the parameters of the model are changed. As each parameter is varied, we keep the others fixed using the base parameter set in Table 1, except that we have raised the drift to $\mu = 0.25$ to highlight some features of the graphs. Notice that the eigenvalues are independent of q_m because this parameter cancels in the term $m_{12}m_{21}$. As well as these two eigenvalues, each panel *a–c* shows the position of the equilibrium point (q_+, Λ_+) and the discriminant Δ^0, which if positive indicates that there are two further equilibrium points on the $q = 0$ axis. Over the majority of parameter space the two eigenvalues are real and of differing sign so that (q_+, Λ_+) is a saddle point. This feature of the phase diagram alters the qualitative structure of the optimal strategies for a finite market.

From Figure 2*a*, we can see that for $q_+ > 0$ the equilibrium point is a saddle point because the eigenvalues $\omega_{\pm}(q_+, \Lambda_+)$ are real and differ in sign, and that as q_+ increases, the discriminant Δ^0 becomes negative. This means the equilibrium points on $q = 0$ axis disappear, and the phase diagram is dominated by the saddle point (q_+, Λ_+). For the parameter range on the graph $\Delta_+ > 0$, so that the saddle remains in the $\Delta > 0$ half-plane. A similar picture emerges as we vary ζ, the nondimensional drift in the market average premium, in Figure 2*b*. Over the parameter scale in the graph, the equilibrium point is always a saddle point in the quadrant $q, \Lambda > 0$. As the drift becomes large, the two equilibrium points on the $q = 0$ disappear. There is a qualitative difference in the phase diagram as we vary ψ in Figure 2*c*. For ψ sufficiently small, corresponding to policies of a long duration, the saddle point enters the quadrant $q > 0$, $\Lambda < 0$, and we expect the form of the optimal control to change.

Clearly the behavior of the optimal control is quite complex, so we defer further discussion to the numerical results in Section 4. Next, we focus on the analysis of the total wealth problem and its analytical reduction.

3.2 Total Wealth

There are utility functions U_1 that lead to a three-dimensional phase diagram. If we consider the problem of maximizing the total utility of wealth ($U_2(w) \equiv 0$), then we introduce the relative marginal wealth

$$\Lambda_2 = \frac{\lambda_2}{\cancel{U_1(w)}} = -\frac{V_w}{dU_1/dw} = \frac{\delta V}{\delta U_1},$$

where Λ_2 has units of time. The corresponding adjoint equation is

$$\begin{aligned} \frac{d\Lambda_2}{dt} &= -\Lambda_2 \frac{U_1''(w)}{U_1'(w)} \pi \left(\frac{k}{\gamma} - 1\right) Gq - 1 \\ &= -\frac{U_1''(w)}{4\gamma U_1'(w)} \pi(qf)\Lambda_2((b - \gamma)^2 - \Lambda^2) - 1, \end{aligned} \tag{3.22}$$

providing the control is interior. Therefore, if the absolute risk aversion of the insurer $-U_1''(w)/U_1'(w) = \mathscr{A}$ is constant, then Λ_2 is independent of w, and the dimension of the phase diagram is reduced by one. If U_1 is linear in w or the utility function is exponential, then the risk aversion is constant and such a reduction occurs. From the transversality conditions, the optimal trajectories must pass through the line $\Lambda = \Lambda_2 = 0$ (the q-axis), which fixes the planning horizon T.

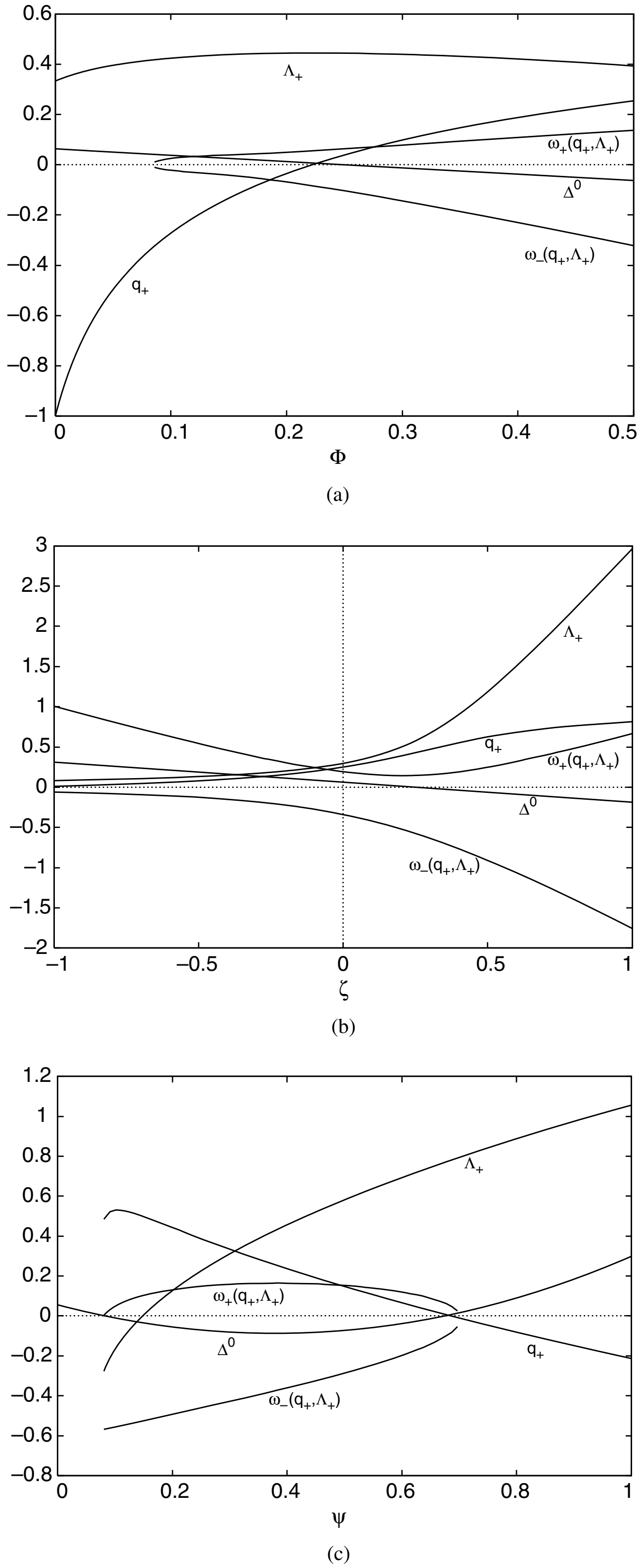
We restrict the analysis to the case that the utility function is exponential

$$U_1(w) = \frac{1 - e^{-\mathscr{A}w}}{\mathscr{A}},$$

where $\mathscr{A} \geq 0$ is the constant risk aversion. The case that the utility function is linear in wealth can be obtained by setting $\mathscr{A} = 0$. The adjoint equations and state equation now become

Figure 2

Linear Stability of Equilibrium Point (q_+, Λ_+) as a Function of Nondimensional Parameters $\phi = b - \gamma$, $\zeta = \mu/a$, and $\psi = \kappa/a$



Note: The eigenvalues of the equilibrium point are denoted by $\omega_{\pm}(q_+, \Lambda_+)$.

Table 1
Sample Data Set

Time horizon T	2.0 yr
Demand parameterization a	3 p.a.
Demand parameterization b	1.5
Market loading θ	0.1
Length of policy $\tau = \kappa^{-1}$	1 yr
Market average premium growth μ	0.1 p.a.

$$\frac{dq}{dt} = q(\tfrac{1}{2}f(b - \gamma + \Lambda) - \kappa), \tag{3.23}$$

$$\frac{d\Lambda}{dt} = -\tfrac{1}{4}(qf)'\Lambda^2 - \Lambda(\tfrac{1}{2}(qf)'(b - \gamma) - \kappa + \mu) - \tfrac{1}{4}(qf)'(b - \gamma)^2 + \frac{\Lambda}{\Lambda_2}, \tag{3.24}$$

$$\frac{d\Lambda_2}{dt} = \frac{\mathscr{A}\pi}{4\gamma}(qf)((b - \gamma)^2 - \Lambda^2)\Lambda_2 - 1, \tag{3.25}$$

with boundary conditions

$$q(0) = q_0, \qquad \Lambda(T) = 0, \qquad \Lambda_2(T) = 0.$$

If the market is infinite and $\mathscr{A} = 0$, then the adjoint equations uncouple from the state equations, but there are no equilibrium points in the phase diagram (Λ, Λ_2) in contrast to the terminal wealth case. We can integrate the second adjoint variable immediately to obtain $\Lambda_2 = T - t$, while the first adjoint equation becomes

$$\frac{d\Lambda}{dt} = -\tfrac{1}{4}a\Lambda^2 - \Lambda(\tfrac{1}{2}a(b - \gamma) - \kappa + \mu) - \tfrac{1}{4}a(b - \gamma)^2 + \frac{\Lambda}{T - t}.$$

Thus, the optimal strategy is independent of the initial size of the insurer, just as it was for the terminal wealth objective, but the adjoint equation is nonautonomous and not integrable analytically.

If the market is infinite and $\mathscr{A} \neq 0$, then there is a single equilibrium point in the phase diagram given by

$$q = \frac{\gamma(\psi^2 - (2\psi - \phi)(\psi - \zeta))}{\mathscr{A}\pi\psi(\phi - \psi)(2\psi \cancel{+} \phi)}, \qquad \Lambda = 2\psi - \phi, \qquad \Lambda_2 = \frac{\gamma}{\mathscr{A}\pi a q(\phi - \psi)\psi}. \tag{3.26}$$

This leads to a qualitative difference in the optimal strategies from the terminal wealth objective. Now, optimal premium strategies depend on the initial exposure of the insurer because this affects the evolution of the wealth of the insurer over the course of the planning horizon.

For an insurer in a finite market, the equilibrium points are determined by the cubic:

$$\gamma(\Lambda + \phi)(\Lambda - \Lambda_+)(\Lambda - \Lambda_-) = 2\mathscr{A}\pi q_m\psi\Lambda(\Lambda - \phi)(\Lambda + \phi - 2\psi), \tag{3.27}$$

where $\Lambda_\pm$ are given by (3.18). From (3.19) and (3.20), we require $\Lambda_+ > 2\psi - \phi$ in order that k is interior and $q_+ > 0$. Thus, by plotting the cubic on the left-hand side of (3.27) superimposed on the cubic of the right-hand side, we can find the intervals in which real roots exist. For the parameter set in Table 1, (3.27) has a real root in $(2\psi - \phi, \Lambda_+)$. It is easy to calculate the roots numerically to find the equilibrium point of the phase diagram.

4. NUMERICAL RESULTS

In this section we describe the numerical techniques used to solve the boundary value problems that determine the optimal control. We start with the simplest problem first and progressively increase the sophistication of the model.

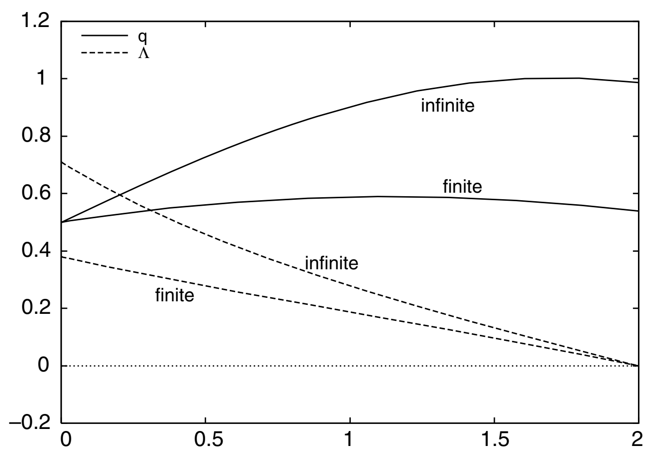
Figure 3 shows the optimal exposure q and adjoint variable Λ for an infinite and a finite market for a terminal wealth problem with constant drift in the market average premium μ. The boundary conditions for this problem are $q(0) = 0.5$, $\Lambda(T) = 0$, and the numerical results are calculated using a so-called shooting method. Thus, we fix the value $\Lambda(0)$, integrate the exposure and adjoint equations (3.7), (3.8) forward in time to give $\Lambda(T)$, and then vary $\Lambda(0)$ until $\Lambda(T) = 0$. We can see in Figure 3 that for both finite and infinite markets the adjoint variable Λ is decreasing, so that the premium is increasing over the planning horizon. However, in a finite market the exposure of the insurer decreases more rapidly at the end of the planning horizon. This is because in a finite market the insurer cannot guarantee to increase its exposure through premium reduction if there are not sufficient customers in the marketplace. This inability to gain exposure means that for the given parameter set, it is optimal for the insurer in a finite market to set a higher premium.

Without an exhaustive numerical study, it is not easy to determine the qualitative properties of the optimal control, and to determine which are the important parameters in the model. To tackle these problems, we consider the phase diagram of the system in Figure 4 for a terminal wealth objective in an infinite market. Each line in panels a and b represents an integration of equations (3.7), (3.8), and because the system is autonomous, we can see how the system evolves *for all* initial conditions (q, Λ). The arrows on each trajectory show increasing time and are separated at half-yearly intervals. It is apparent in panel b that the arrows lie on lines of constant Λ if the trajectories start at the same Λ. This reflects the uncoupling of the adjoint equation from the state equation, so that the relative premium is independent of the current exposure. All the phase diagrams are shown for $q > 0$ and $-\phi < \Lambda < b + \gamma$ because that leads to an interior relative premium. The dashed lines show the extent of this region in the diagrams. In addition, we have plotted the line $\Lambda = \phi$. For $\Lambda > \phi$ the trajectory is loss-leading from (3.4).

Optimal controls must pass through the line $\Lambda = 0$ to satisfy the transversality condition. Thus, in Figure 4a corresponding to $\kappa = 2$, only those trajectories in the lower half of the diagram are optimal, and they all represent a market withdrawal because q decreases along the trajectory. The phase diagram has the merit of showing both the optimal control and the optimal state trajectory. Superimposed on Figure 4a are the two equilibrium points $(0, \Lambda_{\pm})$ given by (3.15). It is clear by comparing panels a and b that the existence of the equilibrium points significantly alters the optimal premium strategy.

Figure 3

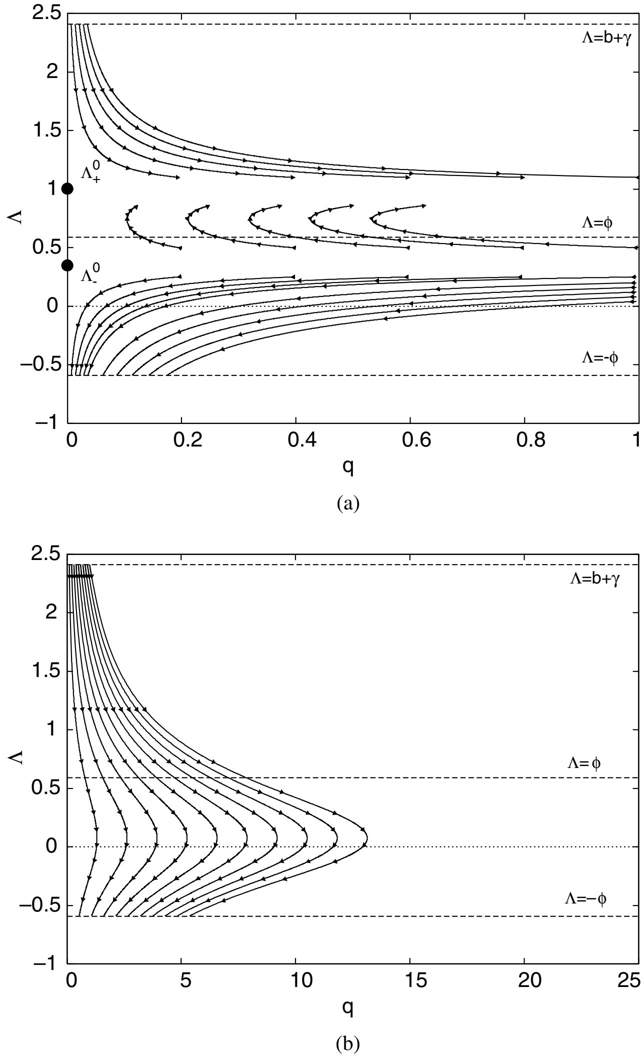
Calculation by Shooting of Optimal State Trajectory q and Adjoint Variable Λ for an Insurer with Finite and Infinite Capacity Using a Terminal Wealth Objective Function



Note: The parameter set is given in Table 1 with $q(0) = 0.5$, and $q_m = 5$ for a finite market.

Figure 4

Phase Diagrams for a Terminal Wealth Objective in an Infinite Market Corresponding to *f* = *a*



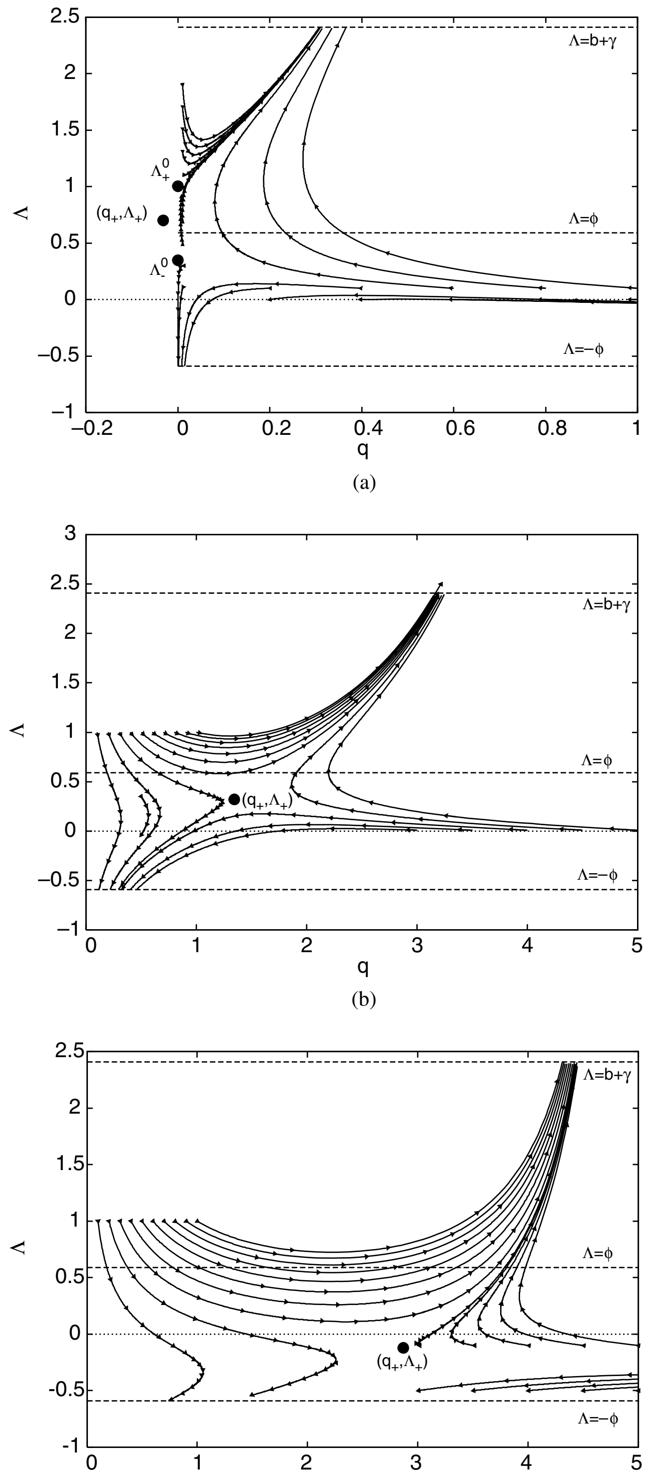
Notes: Parameters are taken from Table 1 except that in (*a*) $\kappa = 2.0$, while in (*b*) $\kappa = 1.0$. Arrows show increasing time *t* with arrow heads at half-yearly intervals.

In panel b, the optimal strategy is always increasing and generates significant market exposure, as one can see from the much larger q scale. Consider an insurer with initial exposure $q_0 = 0.5$. An infinite number of paths intersect the lines $q = q_0$ and $\Lambda = 0$, and these are all optimal trajectories. Arrows are plotted at half-yearly intervals, so that by counting the number of arrows between $q = q_0$ and $\Lambda = 0$, we find the solution of the optimization problem with planning horizon T. For example, we can see that for a time horizon $T \gtrsim 3$ years, there is no smooth *positive* optimal control, because all trajectories extend beyond the line $\Lambda = b + \gamma$ corresponding to $k = 0$. In addition, for $T \gtrsim 1.75$ years the optimal control is loss-leading because $\Lambda > \phi$ on all paths.

When the market is finite there can be additional equilibrium points in the phase diagram, and the optimal control depends on the current exposure of the insurer. Phase diagrams for a finite market using the exposure function (2.2) and a terminal wealth objective are shown in Figure 5a, 5b, and 5c for $\kappa = 2.0$, 1.0, and 0.3, respectively. Figures 4a, 4b, and Figure 5 are phase diagrams for comparable parameter sets and reveal how the structure of the optimal control changes in a finite market.

Figure 5

Phase Diagrams for Optimal Premium Strategy in a Finite Market and Terminal Wealth Objective with $q_m = 5$



(c)

Note: Parameters are taken from Table 1 except that in (*a*) $\kappa = 2.0$, (*b*) $\kappa = 1.0$, and (*c*) $\kappa = 0.3$.

In Figure 5*a* the optimal strategy is similar to the infinite market case: that is, market withdrawal is optimal, because the equilibrium point (q_+, Λ_+) has $q_+ < 0$. The optimal premium is larger than in the infinite market case because optimal trajectories are lowered in Λ. Notice that the two equilibrium points on the $q = 0$ axis still affect the optimal premium strategy so that either strategy type A or B is possible.

In Figure 5*b* policies are longer, and the saddle point is now in the quadrant $q, \Lambda > 0$. Notice that in the lower left of the diagram, we have superimposed the initial condition for the problem solved by shooting in Figure 3. The type of optimal strategy varies according to the initial exposure of the insurer, and the position of the equilibrium point on the phase diagram. The optimal strategy originates from one of the quadrants of the saddle point, and this forms the basis of the classification in Figure 6. Thus, four types of optimal premium strategy emerge:

C: (Loss-leading) These strategies are optimal for a small insurer and feature an increasing premium over the course of the planning horizon. As T increases, loss-leading becomes the dominant optimal premium strategy. The exposure of the insurer can increase or decrease depending on the value on Λ_+. The optimal strategy calculated by shooting in Figure 3 is an example of a type C strategy.

D: (Expansion) These optimal strategies occur when $\Lambda_+ < 0$ and the insurer is relatively small. This strategy represents an expansion into the market because q always increases. The premium is lowered toward the end of the planning horizon as the saturation exposure q_m is approached.

E: (Close to saturation) These strategies are optimal for an insurer that is currently close to its saturation exposure. They can be categorized by a decreasing relative premium over the planning horizon. The exposure can increase as the premium is decreased because demand for policies is generated.

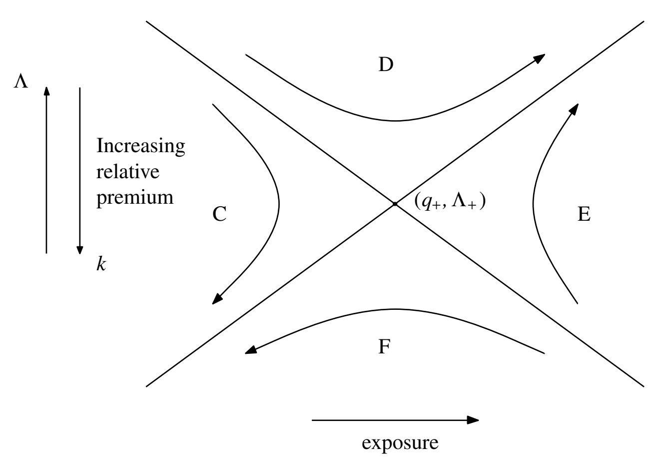
F: (Withdrawal) These strategies occur when $\Lambda_+ > 0$ and the current exposure of the insurer is relatively large. They represent a withdrawal from the insurance market because the exposure always decreases.

Figure 5*c* illustrates the case that $\Lambda_+ < 0$ so that strategies of type C, D, or E can be optimal depending on the initial exposure of the insurer.

Next we consider the numerical solution of the total wealth problem. There are numerical difficulties with the equation (3.24) because the last term on the right-hand side is undefined as $t \to T$ on an optimal trajectory. This is because at $t = T$ the optimal premium is undefined for a total wealth

Figure 6

Classification of Optimal Premium Strategy for an Insurer Maximizing Terminal Wealth in a Finite Market



Note: This classification is appropriate when the mean policy length is large, leading to small κ and an equilibrium point $q_+ > 0$.

problem. We can determine the local behavior of a smooth optimal trajectory by considering the limit as $t \to T$ for the adjoint equations. From (3.25),

$$\Lambda_2 \to T - t, \tag{4.1}$$

because $\Lambda(T) = \Lambda_2(T) = 0$. The leading order term in a power series expansion for Λ substituted into (3.24) gives

$$\Lambda \to \tfrac{1}{8}(qf)'(T)(b - \gamma)^2(T - t) + o(T - t). \tag{4.2}$$

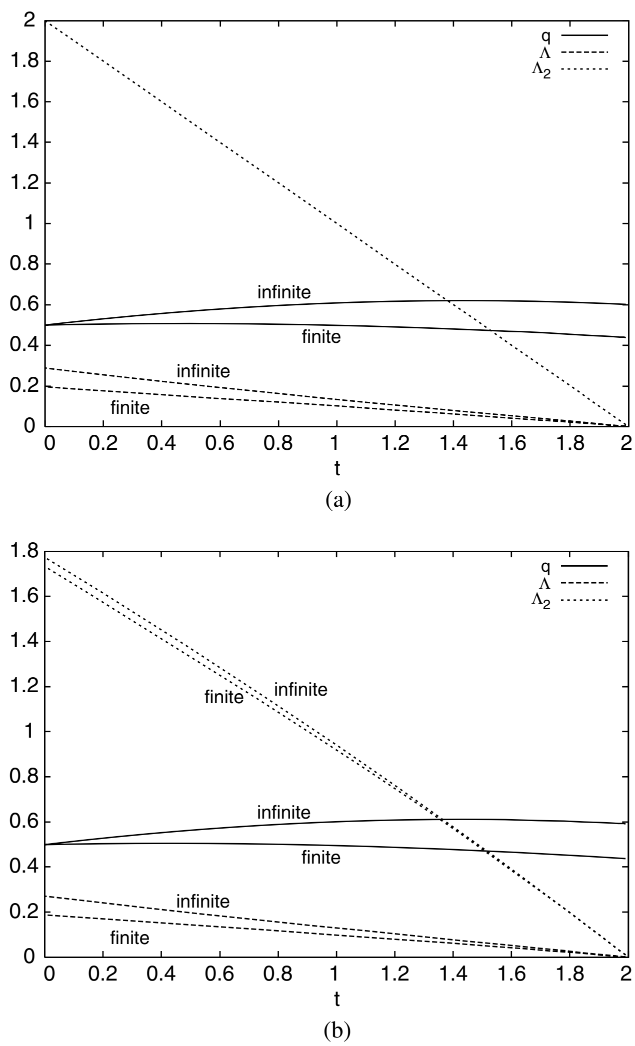
Consequently the optimal phase trajectories approach the origin on the plane

$$\frac{\Lambda}{\Lambda_2} = \tfrac{1}{8}(qf)'(T)(b - \gamma)^2.$$

If the market is finite, then $(qf)'$ is not independent of q, and we cannot determine this behavior analytically without knowledge of the terminal exposure.

Figure 7

Calculation by Shooting of Optimal State Trajectory *q* and Adjoint Variables Λ, Λ_2 for an Insurer in an Infinite or Finite Market Using a Total Utility of Wealth Objective



Notes: (*a*) A linear utility function ($\mathscr{A} = 0$); (*b*) The results for an exponential utility function with risk aversion $\mathscr{A} = 1$. The parameter set is given in Table 1 with $q(0) = 0.5$ and $q_m = 5$ for an insurer with finite capacity.

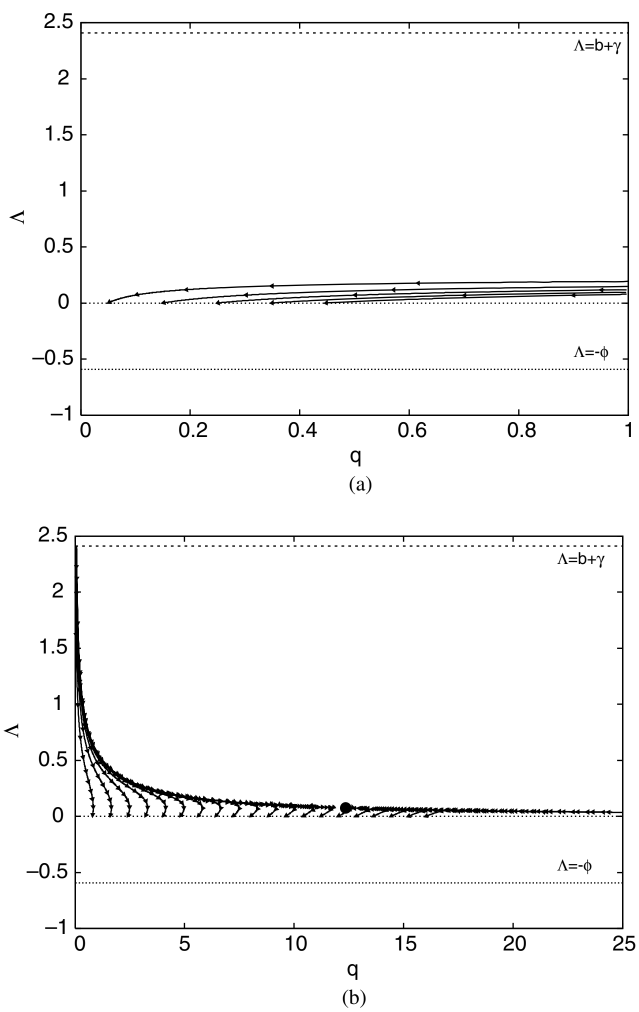
Figure 7 shows the calculation of the optimal premium strategy by shooting for the total wealth problem and is comparable with the terminal wealth problem in Figure 3. Figure ~~3a~~ uses a linear utility function so $\mathscr{A} = 0$, whereas in Figure ~~3b~~ we have used an exponential utility function with risk aversion $\mathscr{A} = 1$. In both plots we compute the optimal trajectory by setting Λ and Λ_2 from (4.2) and (4.1) at $t = T - \varepsilon$ with $\varepsilon = 0.01$ and choosing $q(T - \varepsilon)$. Then we integrate backwards to $t = 0$, which yields $q(0)$, and vary $q(T - \varepsilon)$ until $q(0) = q_0$. This procedure avoids the singularity at $t = T$.

Both panels in Figure 7 show that the optimal premium is increased over the terminal wealth objective, and this leads to an overall lower exposure. In a finite market the premium is still higher and the exposure smaller than in the infinite case. As the risk aversion $\mathscr{A}$ is increased, the optimal premium is raised because this generates lower exposure and decreases the tendency to loss-lead.

To compute the phase diagrams in Figures 8 and 9 we adopt a similar technique and integrate the phase trajectories backwards in time from $t = T - \varepsilon$. Therefore in both phase diagrams we show only trajectories that cross the $\Lambda = \Lambda_2 = 0$ plane, and so these are optimal trajectories. Notice in Figure 8*b* there is an equilibrium point in the phase diagram even though the market is infinite. This behavior is different from the terminal wealth problem where equilibrium points occur only on $q = 0$. With

Figure 8

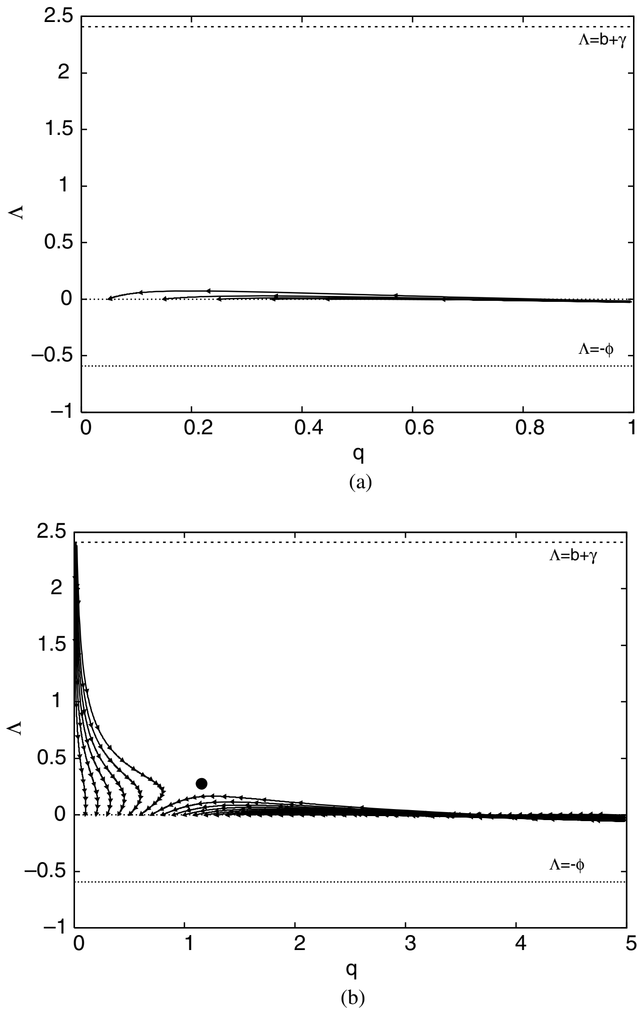
Phase Diagram for a Total Wealth with Exponential Utility Function ~~with~~ Risk Aversion $\mathscr{A}$ = 1.0 and an Infinite Market



Note: (*a*) $\kappa = 2.0$; (*b*) $\kappa = 1.0$.

Figure 9

Phase Diagram for a Total Wealth with Exponential Utility Function ~~with~~ Risk Aversion $\mathscr{A}$ = 1.0 and a Finite Market



Notes: (*a*) κ = 2.0; (*b*) κ = 1.0. The solid circle shows the position of the equilibrium point calculated numerically from (3.27).

strictly positive risk aversion it is optimal for a relatively large insurer to withdraw from the insurance market to avoid the risk of loss-leading. For a finite market shown in Figure 9, the equilibrium value of q is much reduced, and it is optimal to leave the market for a much larger range of initial exposures.

5. NONLINEAR DEMAND FUNCTIONS

When the price function $g(k)$ is a nonlinear function it becomes more difficult to simplify the optimization problem. The linear parameterization (3.1) leads to an explicit linear relationship between the relative premium, k, and the adjoint variable, Λ, given by (3.2). We can see from the form of the first-order condition (2.13), that the form of the inverse elasticity $-g/g'$ determines whether any other parameterization leads to an explicit maximiser of the Hamiltonian. In addition, the ratio g^2/g' determines whether the adjoint equation (2.14) can be integrated analytically, and so yield an explicit expression for the relative premium k^* at time t.

5.1 Inverse Problem

If we specify a simple parameterization for either $(-g/g')(k)$ or $(g^2/g')(\Lambda)$, then we can see if that yields a reasonable price function g and an analytical optimal premium strategy k^*. We call this an inverse problem.

In this section we focus on the pricing problem for an infinite market $f = a$, constant drift in the market average premium μ, and a terminal wealth objective $U_1 = 0$. In this case the adjoint equation becomes

$$\frac{d\Lambda}{dt} = \Lambda(\kappa - \mu) + \frac{ag^2}{g'} \tag{5.1}$$

with transversality condition $\Lambda(T) = 0$. If one relaxes these assumptions, then the state variables enter the adjoint equation, and analytical progress seems unlikely.

It is tempting to specify g^2/g' as a function of the adjoint variable Λ so that one can integrate (5.1) analytically. For example, suppose that $g^2/g' = 1/A$, a constant. Then $g = -1/(Ak + B)$ and B is the constant of integration. However, it is easy to show $g''g/g'^2 = 2$, which violates the second-order condition (2.12) so that the parameterization does not yield a maximizer of the Hamiltonian. Instead, if we try $g^2/g' = A\Lambda + B$, then

$$\frac{d}{dk}\left(\frac{g^2}{g'}\right) = A\,\frac{d\Lambda}{dk} = A\left(\frac{gg''}{g'^2} - 2\right),$$

using (2.15). Applying the differential operator yields $g = -A$, so this cannot provide a suitable parameterization for the demand function. If one poses a quadratic function of Λ for g^2/g', then one obtains an equation similar to (A.1) in the Appendix, which leads to a linear price function. Consequently we shift focus to the ratio g/g'.

Suppose that $g/g' = A$, a constant, then $g = e^{k/A}$ and $A < 0$ for the demand to be a decreasing function of k. The adjoint equation (5.1) becomes

$$\frac{d\Lambda}{dt} = \Lambda(\kappa - \mu) + aAe^{k/A},$$

which is not integrable analytically. Consequently there is not an explicit analytical optimal premium strategy for this price function.

Next, suppose the inverse elasticity is a linear function of the relative premium

$$g/g' = Ak + B.$$

We can easily integrate this ODE to find

$$g = (Ak + B)^{1/A}, \qquad \frac{g''g}{g'^2} = 1 - A, \qquad \frac{g^2}{g'} = (Ak + B)^{1/A+1},$$

where we have set the constant of integration as unity without loss of generality. Consequently, to satisfy the necessary condition for a maximum (2.12), we require $A > -1$. The third expression determines whether we can integrate the adjoint equation (2.14) analytically. For example, if we take $A = -1/2$ and $B = -1/\sqrt{b}$ (so that the linear and nonlinear demand functions agree at $k = 0$), then

$$g(k) = \frac{1}{\left(\frac{1}{2}k + \frac{1}{\sqrt{b}}\right)^2}. \tag{5.2}$$

The first-order condition is now

$$k = 2\left(\gamma + \frac{1}{\sqrt{b}} - \Lambda\right).$$

With this price function the adjoint equation (5.1) becomes

$$\frac{d\Lambda}{dt} = \Lambda(\kappa - \mu) + \frac{a}{\Lambda - \gamma - \dfrac{2}{\sqrt{b}}}, \tag{5.3}$$

which can be integrated analytically. The resulting implicit optimal premium strategy is given in the Appendix.

5.2 Quadratic Price Function

It is not easy to see how the curvature of the price function in the previous section changes the optimal premium strategy without an exhaustive numerical study. Therefore, in this section we perturb the linear price function (3.1) by a quadratic function and see how that modification changes the qualitative features of the phase diagram.

Consider the quadratic price function

$$g(k) = ((b - k)(1 + ck))^{+}. \tag{5.4}$$

When $c = 0$, g reduces to the linear price function, whereas if $c \gtrless 0$, then the demand function is convex/concave, respectively. If $0 \le k \le b$, then the derivative of g must be such that g is a positive decreasing function of k. Consideration of the end points of the domain yields the condition that

$$|c| < \frac{1}{b}. \tag{5.5}$$

Substituting g into the first-order condition (2.13) yields a quadratic for k:

$$3ck^2 + 2k(1 + c(\Lambda - b - \gamma)) + (1 - bc)(\Lambda - \gamma) - b = 0,$$

and we must determine which, if any, of the roots of this equation yields the maximizer of the Hamiltonian defined by (2.7). For the terminal wealth, finite market problem the Hamiltonian becomes

$$H = -\kappa\lambda_1 q + \mathcal{H},$$

where

$$\mathcal{H} = qf((b - k)(1 + ck))^{+}(\lambda_1 + \pi\lambda_2(k\gamma^{-1} - 1)).$$

Now $\mathcal{H}$ has three roots: $k = -1/c$, $k = b$, and $k = \gamma - \Lambda$, and the first of these roots must come before the second from (5.5). At termination $\Lambda = 0$ and $\gamma < b$ for there to be premium values above ~~break even~~ that yield positive demand. Thus, at termination $-1/c < b < \gamma$, and from the panels in Figure 10, we see that the interior control is given by

$$k^i(\Lambda) = \frac{-(1 + c(\Lambda - b - \gamma)) + ((1 + c(\Lambda - b - \gamma))^2 - 3c((1 - bc)(\Lambda - \gamma) - b))^{1/2}}{3c},$$

irrespective of the sign of c. Moreover, this is the maximizer of the Hamiltonian if $k^i < b$, for then the demand is nonzero.

With this demand function the adjoint equation is

$$\frac{d\Lambda}{dt} = \Lambda(\kappa - \mu) + (qf)'\,\frac{(b - k^i)^2(1 + ck^i)^2}{c(b - 2k^i) - 1},$$

Figure 10
Determination of the Interior Control for Quadratic Demand Function (5.4)

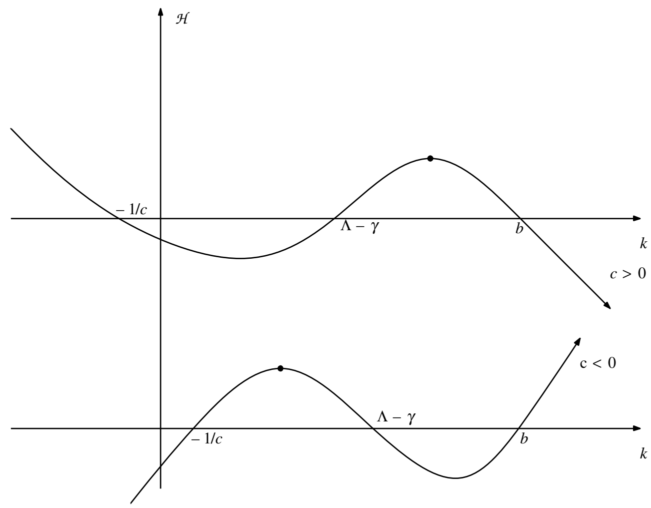


Figure 11
Phase Diagrams for the Quadratic Demand Law: (*a*) *c* = 0.25, (*b*) *c* = 0.5

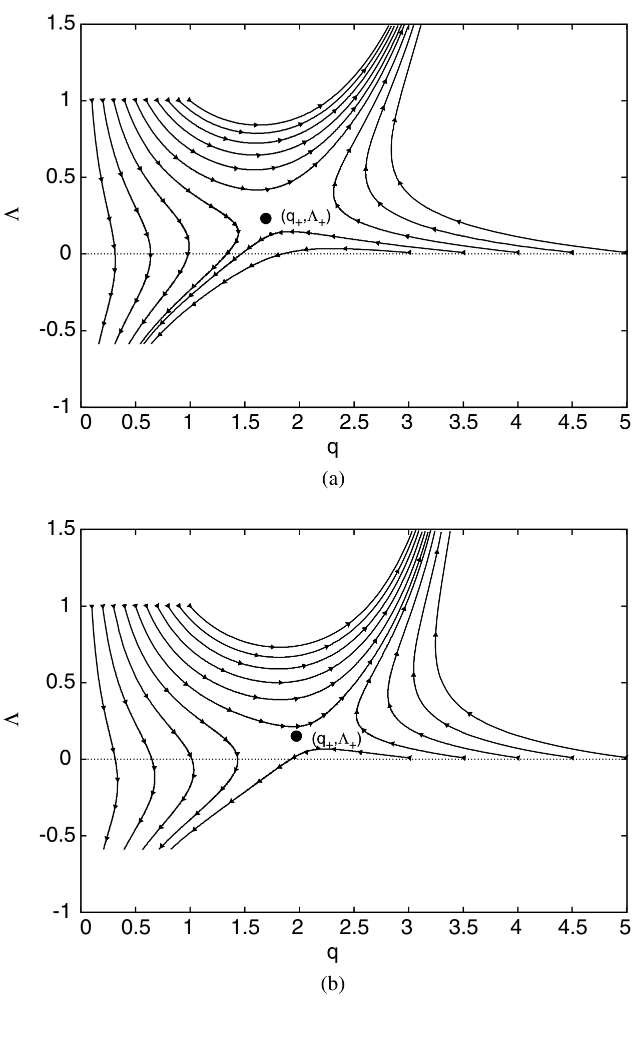
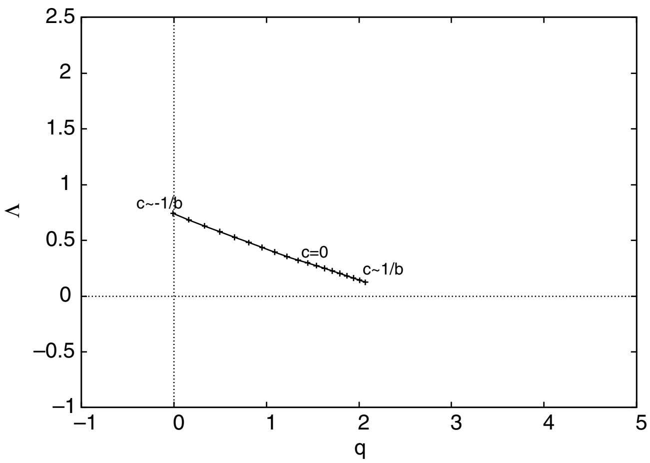


Figure 12
Position of Equilibrium Point in the Phase Diagram for a Quadratic Demand Law as Parameter *c* Is Varied



where $(qf)'$ is given by (3.17).

The second-order condition (2.12) is

$$\frac{gg''}{g'^2} = \frac{2c(k - b)(1 + ck)}{(c(b - 2k) - 1)^2} < 2$$

providing $k \leq b$. Expanding out the inequality gives

$$3c^2k^2 + 3c(1 - bc)k + 1 - bc + b^2c^2 > 0,$$

and the quadratic on the left-hand side has discriminant

$$\Delta^c = -3c^2(1 + bc)^2 < 0.$$

for all $c \neq 0$. Consequently, there are no real roots of this quadratic so that the second-order condition is always satisfied, providing $k \leq b$.

Figure 11 shows two phase diagrams corresponding to two fixed positive values of c for the quadratic price function. It is clear that the equilibrium remains a saddle point, but that the nonlinearity of demand moves the position of the equilibrium point in phase space. Figure 12 shows the movement of this saddle point as the parameter c is varied. As c is increased, which corresponds to increased demand for a fixed relative premium, the equilibrium point is pushed to the right in the diagram. This favors loss-leading optimal strategies in the leftmost quadrant of the saddle point as one might expect. Conversely, as c is decreased, the equilibrium point is pushed to the left in the diagram, and the rightmost quadrant of the saddle becomes dominant. These optimal strategies represent market withdrawal.

6. CONCLUSIONS

We have described a general deterministic model for pricing general insurance using optimal control theory. The theory encompasses different parameterizations of the demand for policies and different objectives for the insurer. Any model tackled via control theory becomes more difficult to analyze as one increases the number of state variables to accurately model the underlying processes. We have focused on how the optimization problem is simplified as the assumptions of the model are changed.

The simplest problem, that of an insurer in an infinite market with a terminal wealth objective, requires only the backwards integration of the adjoint variable of the exposure. This has an explicit analytical solution if the price function is linear. We have also found an implicit analytical solution for a nonlinear price function, although for this parameterization there is no cutoff in relative premium beyond which there is no demand for insurance. Thus, it becomes difficult to classify the optimal strategy because it is always optimal to sell insurance policies.

When the market is finite, the simplest optimization problem becomes a boundary value problem, where the exposure is integrated forwards in time, and simultaneously the adjoint of the exposure is integrated backwards. No analytical solutions have been found in this case. However, by analyzing the phase diagram of the state/adjoint system, we have explored the optimal strategies for the insurer. It is found that premium strategies vary according to the equilibrium point(s) in the phase diagram, and that these points are always unstable saddle points over the parameter set of interest. The type of optimal strategy can be classified according to in which quadrant of the saddle the insurer lies as given by its initial exposure and the position of the equilibrium point. For example, one quadrant corresponds to a loss-leading strategy where it is optimal to set an increasing premium and build up exposure if the insurer is particularly small. For the terminal wealth problem, there is an explicit expression for the position of the equilibrium point.

The demand function is the parameterization that most affects the optimal premium strategy. There are certain restrictions on the form of the demand function: most notably we require $gg'' < 2g'^2$, where g is the price function, in order that the first-order condition of the Hamiltonian gives a maximum. This is an analogous result to that given by Taylor (1986) and Kalish (1983).

In an infinite market, the optimal premium strategy for the total wealth objective depends on the current size of the insurer if the utility function is nonlinear. The nonlinearity of the concave utility function means that low wealth is relatively more favorable over high wealth, and this affects the premium strategy of a relatively large insurer where the insurer is close to its saturation exposure. If the demand function is concave indicating lower demand for a given relative premium ratio, then that favors market withdrawal over a loss-leading strategy. Similarly, convex demand functions push the equilibrium point in the phase diagram toward the region of withdrawal so that loss-leading is favored.

In further research we shall consider a stochastic generalization of (1.1) to understand how the uncertainty in gaining exposure from given premium values changes the optimal premium strategy. The present paper forms the foundation for this stochastic problem because we have determined the optimal premium strategy in the limit as the volatility of exposure tends to zero.

APPENDIX

ANALYTICAL OPTIMAL PREMIUM STRATEGIES

Here we give the two analytical optimal premium strategies described in the main text. The strategies are given in nondimensional form to highlight which parameter values are important. Both strategies are for a terminal wealth problem ($U_1 \equiv 0$) with an infinite market ($f = a$) and constant market drift in the market average premium μ.

If the price function is linear (3.1) and the control interior, then there is an analytical optimal premium strategy. This is identical to that described in Emms (2007a) and is given here for completeness. From (3.7) the adjoint equation is

$$\frac{d\Lambda}{dt} = -\tfrac{1}{4}a\Lambda^2 - \Lambda(\tfrac{1}{2}a(b - \gamma) - \kappa + \mu) - \tfrac{1}{4}a(b - \gamma)^2. \tag{A.1}$$

We can nondimensionalize (A.1), remembering that Λ is nondimensional, by defining the nondimensional time to termination $s = a(T - t)$. The nondimensional adjoint equation is then

$$\frac{d\Lambda}{ds} = \tfrac{1}{4}(\Lambda + \phi)^2 + \rho\Lambda. \tag{A.2}$$

On applying the boundary condition $\Lambda = 0$ at $s = 0$, this can be integrated to obtain

$$\Lambda = \begin{cases} \dfrac{(e^{sD_+} - 1)(2(\rho - D_+) + \phi)}{1 - \left(\dfrac{2(\rho - D_+) + \phi}{2(\rho + D_+) + \phi}\right) e^{sD_+}} & \text{if } \Delta^0 > 0, \\ 2\rho + \phi - 2D_- \tan\left(\tfrac{1}{2}sD_- + \tan^{-1}\left(\dfrac{2\rho + \phi}{2D_-}\right)\right) & \text{if } \Delta^0 < 0, \\ \dfrac{s(2\rho + \phi)^2}{4 - s(2\rho + \phi)} & \text{if } \Delta^0 = 0, \end{cases} \tag{A.3}$$

where $D_\pm = \sqrt{\pm\Delta^0}$. The analysis of this form of control when there are constraints and when $k^i > b$ is contained in Emms (2007b).

For the infinite time horizon problem $T \to \infty$, which corresponds to the limit $s \to \infty$. Therefore, we can see that if $\Delta^0 < 0$, then the limit is not finite and no optimal strategy exists. If $\Delta^0 \geq 0$, then the limit is finite, and in the case $\Delta^0 = 0$

$$k^i = b + \rho.$$

There is another analytical optimal premium strategy when the demand law is given by (5.2). In nondimensional form the adjoint equation (5.3) is

$$\frac{d\Lambda}{ds} = \rho\Lambda + \frac{1}{z - \Lambda} = \frac{1 + \rho z\Lambda - \rho\Lambda^2}{z - \Lambda}, \tag{A.4}$$

where we write $z = \gamma + 2/\sqrt{b}$. To simplify the notation we define the discriminant for this equation as

$$\Delta^i = \rho(\rho z^2 + 4)$$

and integrate to obtain

$$s = \begin{cases} \log(1 + \rho z\Lambda - \rho\Lambda^2) \quad \dfrac{z}{2D_+} \log\left(\left(\dfrac{2\rho\Lambda - D_+ - \rho z}{2\rho\Lambda + D_+ - \rho z}\right)\left(\dfrac{\rho z - D_+}{\rho z + D_+}\right)\right) & \text{if } \Delta^i > 0, \\ \dfrac{\log(1 + \rho z\Lambda - \rho\Lambda^2)}{2\rho} - \dfrac{z}{D_-}\left(\tan^{-1}\left(\dfrac{2\rho\Lambda - \rho z}{D_-}\right) + \tan^{-1}\left(\dfrac{\rho z}{D_-}\right)\right) & \text{if } \Delta^i < 0, \end{cases} \tag{A.5}$$

where we have redefined $D_\pm = \sqrt{\pm\Delta^i}$. If $\Delta^i = 0$, then either $\rho = 0$ in which case

$$s = \tfrac{1}{2}\Lambda(2z - \Lambda) \tag{A.6}$$

or $\rho = -4/z^2$, and we integrate to obtain

$$s = \tfrac{1}{4}z^2\left(\frac{2\Lambda}{z - 2\Lambda} - \log\left(1 - \frac{2\Lambda}{z}\right)\right).$$

Notice these are implicit relationships between s and Λ, and so therefore we obtain only an implicit form for the optimal premium strategy. This inversion is not always possible, indicating that there is no optimal strategy. For example, from (A.6), if $\Delta^i = \rho = 0$ and $s > \frac{1}{2}z^2$, then there is no optimal premium strategy.

REFERENCES

BASS, F. M. 1969. A New Product Growth Model for Consumer Durables. *Management Science* 15(5): 215–27.

CUMMINS, J. D., AND J. F. OUTREVILLE. 1987. An International Analysis of Underwriting Cycles in Property-Liability Insurance. *Journal of Risk and Insurance* 54: 246–62.

DOHERTY, N. A., AND J. R. GARVEN. 1995. Insurance Cycles: Interest Rates and the Capacity Constraint Model. *Journal of Business* 68(3): 383–404.

DOLAN, R. J., AND A. P. JEULAND. 1981. Experience Curves and Dynamic Demand Models: Implications for Optimal Pricing Strategies. *Journal of Marketing* 45: 52–62.

EMMS, P. 2007a. Dynamic Pricing of General Insurance in a Competitive Market. *ASTIN Bulletin* 37(1): 1–34.

———. 2007b. Pricing General Insurance with Constraints. *Insurance: Mathematics and Economics* 40: 335–55.

EMMS, P., AND S. HABERMAN. 2005. Pricing General Insurance Using Optimal Control Theory. *ASTIN Bulletin* 35(2): 427–53.

GERBER, H. U., AND G. PAFUMI. 1998. Utility Functions: From Risk Theory to Finance. *North American Actuarial Journal* 2(3): 74–100.

JORDAN, D. W., AND P. SMITH. 1977. *Nonlinear Ordinary Differential Equations*. Oxford: Oxford University Press.

KALISH, S. 1983. Monopolistic Pricing with Dynamic Demand and Production Cost. *Marketing Science* 2: 135–60.

LÉONARD, D. AND N. VAN LONG. 1992. *Optimal Control Theory and Static Optimization in Economics*. Cambridge: Cambridge University Press.

LILIEN, G. L., AND P. KOTLER. 1983. *Marketing Decision Making*. New York: Harper & Row.

ROBINSON, B., AND C. LAKHANI. 1975. Dynamic Price Models for New-Product Planning. *Management Science* 21: 1113–22.

SETHI, S. P., AND G. L. THOMPSON. 2000. *Optimal Control Theory*. 2nd ed. Dordrecht: Kluwer.

SMITH, G. D. 1985. *Numerical Solution of Partial Differential Equations: Finite Difference Methods*. Oxford: Oxford University Press.

TAYLOR, G. C. 1986. Underwriting Strategy in a Competitive Insurance Environment. *Insurance: Mathematics and Economics* 5(1): 59–77.

YONG, J., AND X. Y. ZHOU. 1999. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. New York: Springer.

Discussions on this paper can be submitted until July 1, 2009. The authors reserve the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.