Optimal Control of Biological Invasions in Lake Networks

Alexei. B. Potapov
Dept of Mathematical and Statistical Sciences and Centre for Mathematical Biology, University
of Alberta, Edmonton, AB, T6G 2G1 Canada
E-mail address: apotapov@math.ualberta.ca
Mark. A. Lewis
Dept of Mathematical and Statistical Sciences and Centre for Mathematical Biology, and
Dept of Biological Sciences
University of Alberta, Edmonton, AB, T6G 2G1 Canada
E-mail address: mlewis@math.ualberta.ca
David C. Finnoff
Department of Economics and Finance,
University of Wyoming, Laramie, WY, 82071,
E-mail address: finnoff@uwyo.edu

ABSTRACT. A metapopulation model for alien species invasion of a lake network is coupled with an economic model of prevention. The model restates a stochastic problem in deterministic terms. It provides a macroscopic description of the lake network with prevention methods controlling both the outflow of invaders at infected lakes and the inflow of invaders at uninfected lakes. Results indicate that optimal control implements no more than one of these methods at any moment in time. Typical optimal control measures change over time as the lake ecosystem becomes successively more invaded. Early control of outflow from infected lakes is replaced by later control of inflow to remaining uninfected lakes. Closed-loop control trajectories are analytically characterized in the phase-plane for a limiting case, while in general a simple and stable numerical algorithm is developed for solving the optimal control problem.

KEY WORDS: optimal control, natural resources, biological invasions, invasive species

1. INTRODUCTION

Successful invaders can change the balance of natural ecosystems, lead to the extinction of native species, and cause damage to natural resources and economic activity. This problem of nonindigenous species is not purely biological; its description requires both biological and economic factors, and their feedbacks (Crocker and Tschirhart, 1992; Settle et al., 2001). In lake ecosystems the economic changes induced by invaders are diverse and include pipeline fouling, fishery degradation, water quality impacts, and loss of recreation potential (e.g. O'Neill, 1997).

When an invader is introduced into a lake or a river, its establishment depends on the state of the existing ecosystem. If it has successfully established and starts to grow considerably, it may change the ecosystem itself, which creates both short and long-term biological and economic consequences. After the species has established, it may start to spread to other lakes, via human-related and natural dispersal mechanisms.

While it may be impractical to consider all of the factors mentioned in a single model — such a model would be very complex, a number of different models can be used, each one capturing components of the complex phenomenon. Each model can focus on a particular scale: macroscopic (description of average characteristics of a large region, Shigesada and Kawasaki, 1997; Clark, 1990), middle-scale (e.g. transport of invaders between lakes, Buchan and Padilla, 1999), or small-scale (processes in a single lake or a part of it, analysis of a single population etc., Turchin, 2003). Lower-level models may be necessary to estimate parameters of the larger-scale models and, in turn, the latter can be used to estimate basic trends, which are to be needed on smaller scales.

In this paper the optimal management of a lake system subject to nonindigenous species invasion is analyzed at the macroscopic level. We simplify a complex stochastic process into a model that takes into account important effects on average, or proportionally. This allows the

1

restatement of a stochastic problem in deterministic terms. The method permits the incorporation of intersecting biology and economics at a large scale, and investigates their joint influence on decision-making and optimal invasion management.

The analysis provides several contributions. First we demonstrate, both analytically and numerically, how optimal prevention policy depends upon several key factors, including the mean economic damage per lake, the efficiency of prevention, the planning horizon, initial magnitude of the invasion, and the discount rate. Given our choice of functional form, we find that at each moment it is optimal to implement at most only one type of prevention: control only at infected lakes, only at uninfected lakes, or at neither. There may be a moment of control switching, which depends on the efficiency and per unit cost of each control.

Second, we perform a comprehensive phase-plane analysis on the dynamical system. We are able to analytically characterize control trajectories in the phase-plane and closed-loop optimal control polices for a limiting case where the discount rate is zero. For the general case, we present a simple and stable numerical algorithm for the control problem. Unlike usual methods in optimal control, where the equation governing the stock dynamics has to be integrated forward, and that for the shadow price backward, on the phase plane there is only a single equation, simplifying the integration considerably. While we are not able to analytically characterize the complete general case, we find several comparison propositions that provide considerable insight into the dynamics of the system.

There is a large literature on the successful application of macroscopic models of fishery management as laid out in Clark (1990). In spite of the simplicity of such macroscopic models, they provide many explanations and insights into the organization of fisheries. Similar attempts have been made in resource extraction, agriculture, national parks management and similar fields (van Kooten & Bulte, 2000). Olsen and Roy (2002) use a spaceless, stochastic dynamic

optimization model to characterize situations where the eradication of an invasive species is optimal and not optimal, given natural growth and spread are subject to exogenous environmental risk. Their results highlight the importance of the speed of growth/spread of the invasion, dependent in turn on the exogenous risk. Herein the speed of the spread is not taken as exogenous, rather it is endogenous to the policy maker. Brown, Lych and Zilberman (2002) abstract from all dynamic issues and employ a spatial, static approach to investigate transmission and source control of insect-transmitted plant diseases. Allowing for removal of diseased source plants and/or barriers restricting insect transmission of the disease, the results demonstrate that while (complete) source control and transmission barriers improve welfare over only barriers, when non-market values over source areas are included the less effective transmission barriers solely may be optimal.

This paper extends the literature by viewing noninvaded space as a nonrenewable resource, and investigates its optimal management. This allows us to find analogies between our approach and other models, e.g. optimal control of epidemics or optimal use of an antibiotic (Laxminarayan, 2001). Nonetheless, the models are different, as are the main conclusions. The most important difference arises from the highly nonlinear structure of the control terms in the state dynamics.

The paper is organized as follows. In Section 2 the main assumptions underlying a macroscopic model of optimal invasion prevention are described. In Section 3 the model is developed, followed by a preliminary analysis, which enables the characterization of a number of important conclusions. Section 4 presents a phase-plane analysis. Numerical simulations are employed to illustrate several propositions in Section 5, followed by a brief conclusion.

2. MACROSCOPIC MODEL OF INVASION AND PREVENTION

2.1 Biological background

Suppose there is an alien species introduced into one or several lakes. It has established and starts to spread to other lakes in a region. A primary vector of human influenced spread is the transportation of boating and fishing equipment from lake to lake (see e.g. Johnson et al. (2001) for an analysis of zebra mussel dispersal). Consider a resource manager who allocates scarce resources to maximize social welfare subject to the damages of invasion, costs of preventing further spread, and the spread of the invader. The primary means of preventing further spread is a disinfecting process at some point in the transportation mechanism, for example washing equipment on its way from one lake to another (Buchan and Padilla, 1999). In this setting, we call invaded lakes "donor" since they are the source of traveling invader propagules, and the noninvaded lakes we call "recipient".

The total number of lakes N is assumed to be large enough, such that (i) we can characterize the invasion process by a single variable, the proportion of invaded lakes p (the number of invaded lakes N_I divided by the total number of lakes, $p=N_I/N$), and (ii) change of p with time may be reasonably approximated by a continuous and differentiable function. In this case it is possible to derive a quadratic growth model for $p(t)$ that is both biologically reasonable and mathematically convenient

$$\frac{dp}{dt} = Ap(t)(1 - p(t)). \tag{1}$$

The derivation follows from the assumption that the transport of boating and fishing equipment is independent of the invasion process and this transportation mechanism is also assumed to connect all lakes under consideration. The mean potential number of invader propagules that can be transported from a donor lake to a recipient one, characterizes traffic in the transportation

mechanism. In the absence of any prevention effort, we assume the average number of propagules that can be transported from any donor lake to any given recipient lake per unit time (defined as the intensity of propagule transport) be constant and denoted by A_1. For a small period of time Δt the mean total number of propagules transported from each of the N_I invaded lakes to any other given lake is then $K=N_I A_1 \Delta t$. Likewise the total number introduced into N_R lakes is $N_I N_R A_1 \Delta t$. If the probability a propagule survives following introduction is given by A_2, the number of surviving propagules is $N_I N_R A_1 A_2 \Delta t$. As $N_R=N-N_I$, the increase in the number of invaded lakes during Δt is,

$$\Delta N_I = A_1 A_2 \Delta t N_I (N-N_I).$$

Dividing through by N and taking into account that $\Delta N_I/N=\Delta p$, $N_I=Np$, and $\Delta N_I=N\Delta p$ we come to

$$\Delta p = A_1 A_2 N \Delta t p (1-p).$$

Assuming $p(t)$ differentiable and Δt small, we can replace this by differential equation (1) with $A=A_1 A_2 N$. This model of invasion *without prevention* can be interpreted as Levins metapopulation model (Levins, 1969) without extinction or simply as logistic growth equation (Clark, 1990) in the proportion of invaded lakes over time.

2.2 Economic and control assumptions

Let the resource manager believe the benefits of preventing the spread of the invasion be such that costly investments in prevention are justified. Assume it is possible to estimate the cost of damage to ecosystems due to the invasion. Let the cost of damage per unit time be proportional to the total number of invaded lakes, and let g be the (constant) monetized damage cost per single lake per unit time. Total damage costs per unit time are then $C_1(t)=gN_I(t)$.

The resource manager employs prevention, which adds to the costs of invasion per unit time. For example, to wash boats it is necessary to create checkpoints and supply materials, both

5

of which are costly. While not necessary, for brevity let the disinfection process (prevention) be general enough such that it may be performed at every lake, donor or recipient. (In the context of epidemics, this is similar to the treatments of those infected with a disease and as well as those not infected). Let prevention efforts at donor lakes be $x(t)$ with (constant) unit cost w_x, and prevention efforts at recipient lakes be $s(t)$ with a unit cost w_s. The total cost of prevention per unit time is then

$$C_2(t)=w_x x(t)N_I(t)+w_s s(t)(N-N_I(t)). \tag{2}$$

Combining damage and prevention costs delivers the total cost per unit time,

$$TC(t)=C_1(t)+C_2(t)=gN_I(t)+w_x x(t)N_I(t)+ w_s s(t)(N-N_I(t)).$$

To mesh the manager's behavior with the macroscopic biological model and maintain a focus on important effects in terms of proportion of lakes invaded, we compute the average cost of invasion per lake per unit time,

$$C(t)=gp(t)+w_x x(t)p(t)+w_s s(t)(1-p(t)). \tag{3}$$

Average costs are comprised of average cost of damages due to the invasion (first term), average cost of prevention at donor lakes (second term), and the average cost of prevention at recipient lakes (third term).

Prevention on a donor lake $x(t)$ reduces the average number of propagules transported to any recipient lake from A_1 to $A_1 a(x(t))$, $0<a\leq1$. Here $a(x(t))$ is the probability of a propagule escaping treatment, and $1-a(x(t))$ can be interpreted as the proportion treated at donor lakes. Similarly $s(t)$ and $b(s(t))$, $0<b\leq1$ are prevention effort and probability of escaping treatment at any recipient lake.

To complete the specification of the problem it is necessary to define an explicit relationship between the proportions treated ($1-a(x(t))$ and $1-b(s(t))$) and the treatment efforts ($x(t)$ and $s(t)$).

In deriving this relationship we assume that the effects of two successive treatments are independent, and that treatments with efforts x_1 and x_2 are equivalent to a single treatment with the efforts x_1+x_2. Then for the probabilities of two independent events we have $a(x_1+x_2)= a(x_1) a(x_2)$. We further assume that a small effort $x_2=\Delta x<<1$ treats a proportionate fraction of propagules $1-a(\Delta x)\approx k_1\Delta x$, where k_1 is the control efficacy. Upon substitution into the previous rule we have $a(x+\Delta x)=a(x)a(\Delta x)\approx a(x)-k_1a(x)\Delta x$, or $\Delta a=-k_1a(x)\Delta x$ as Δx approaches zero. Similarly $\Delta b=-k_2b(s)\Delta s$ as Δs approaches zero. This leads to simple exponential forms that incorporate diminishing effectiveness of prevention,

$$a\big(x(t)\big)= e^{-k_1x(t)}, \quad b\big(s(t)\big)= e^{-k_2s(t)} \tag{4}$$

where k_1 and k_2 characterize heterogeneous efficiencies in boat processing. Under these assumptions, the spread of the infection is described by the equation (1) with $A=A_1A_2N$ replaced by $A=A_1aA_2bN$ and a and b as given in equation (4),

$$\frac{dp}{dt} = Ae^{-k_1x(t)-k_2s(t)}p(t)\big(1-p(t)\big) \tag{5}$$

with the initial proportion of infected lakes $p(0)=p_0$.

3. OPTIMAL CONTROL OF INVASION

To characterize optimal management, the problem facing the resource manager is to choose prevention efforts $x(t)\geq0$ and $s(t)\geq0$ so as to minimize the discounted stream of costs J during a given time horizon T,

$$J[x(t),s(t)]= \int_0^T e^{-rt}C(t)dt ,$$

where r is the discount rate (e.g. Clark, 1990) and $C(t)$ the average cost of invasion. If x and s are chosen such that the total cost of invasion J is minimal, an optimal control task is obtained. Optimal trajectories for x and s are denoted $x^*(t)$, $s^*(t)$.

Pontryagin's maximum principle is used to solve the task facing the manager (Pontryagin *et al*., 1964). Optimal rules for $x(t)$ and $s(t)$, given an initial proportion of lakes invaded, can be found from the maximum principle as applied to the current value Hamiltonian,

$$H = -\left(gp(t) + w_x x(t)p(t) + w_s s(t)(1 - p(t))\right) + \mu(t)Ae^{-k_1 x(t) - k_2 s(t)} p(t)(1 - p(t)), \tag{6}$$

where $\mu(t)$ is the current value shadow price of the proportion of lakes invaded. p and μ satisfy differential equations

$$\frac{dp}{dt} = Ae^{-k_1 x - k_2 s} p(1 - p) = F(p, x, s), \quad 0 \le t \le T, \tag{7}$$

$$\frac{d\mu}{dt} = r\mu - \frac{\partial H}{\partial p} = \mu\left[r - Ae^{-k_1 x - k_2 s}(1 - 2p)\right] + (g + w_x x - w_s s) = G(p, \mu, x, s). \tag{8}$$

$$p(0) = p_0, \quad \mu(T) = 0.$$

Given the underlying assumption of only piecewise continuous, bounded controls, this implies that $p(t)$ and $\mu(t)$ are continuous and piecewise continuously differentiable as solutions of ODEs (7) and (8) with piecewise continuous right hand sides.

The necessary condition for optimality of control is maximum value of H at each moment of time, that is for optimal $x = x^*$ and $s = s^*$

$$\frac{\partial H}{\partial x} = -w_x p - k_1 \mu Ae^{-k_1 x - k_2 s} p(1 - p) = 0, \quad \text{or} \quad x = 0, \frac{\partial H}{\partial x} < 0, \tag{9}$$

$$\frac{\partial H}{\partial s} = -w_x(1 - p) - k_2 \mu Ae^{-k_1 x - k_2 s} p(1 - p) = 0, \quad \text{or} \quad s = 0, \frac{\partial H}{\partial s} < 0, \tag{10}$$

where time notation is suppressed for brevity. The first condition in each of (9) and (10)

corresponds to an "internal maximum" for x and s respectively, whereas the second conditions corresponds to "boundary maximum". Note that generically the conditions $\partial H/\partial x=0$ and $\partial H/\partial s=0$ cannot be satisfied together because then simultaneously

$$w_x = -k_1\mu Ae^{-k_1 x-k_2 s}(1-p), \quad w_s = -k_2\mu Ae^{-k_1 x-k_2 s}p, \tag{11}$$

or $k_2 pw_x = k_1(1-p)w_s$. This can hold only for single p value

$$p = p_S = \frac{w_s k_1}{w_s k_1 + w_x k_2}. \tag{12}$$

Therefore not more than one type of control is nonzero. In what follows it is necessary to know the sign of μ.

Proposition 1. *The shadow price $\mu(t)$ for $t<T$ is always negative provided the invasion is costly (i.e. $g>0$).*

A proof is given in the Appendix.

To determine what p ranges correspond to what type of control consider (12) and the necessary conditions. Let $p<p_S$, then from the expressions for $\partial H/\partial x$ and $\partial H/\partial s$ it follows that assumption $s\geq 0$ and $\partial H/\partial s=0$ implies $\partial H/\partial x>0$ and $x\geq 0$, which gives a contradiction. On the other hand, assumption $x\geq 0$, $\partial H/\partial x=0$, implies $\partial H/\partial s<0$ and $s=0$, which is consistent. Similarly we can check the case $p>p_S$. This allows us to conclude that

a) for $p<p_S$ only control at invaded lakes x may be nonzero;

b) for $p>p_S$ only control at uninvaded lakes s may be nonzero;

c) the value $p=p_S$ corresponds to switching between the two types of control. The conditions (9), (10) give only the value of aggregate control $\varphi=k_1 x+k_2 s$ at this point, but not x and s separately. Therefore one of x or s can be chosen at our convenience. For example we can choose $s=0$, then one of the controls is zero for $p=p_S$ as well.

Therefore, at any moment only one control variable can be nonzero. This allows the sufficient conditions for a maximum of H to be checked as if it were a function of only one variable, that is just check the signs of the second derivatives $\partial^2 H/\partial x^2$ and $\partial^2 H/\partial s^2$. This is because the maximum of H is always reached at the boundary of the admissible control set. Thus the standard procedure for internal maximum of a 2D function, involving the Hessian matrix, is not applicable.

For the sufficient conditions it is straightforward to show that, since $\mu < 0$,

$$\partial^2 H/\partial x^2 = k_1^2 \mu A e^{(-k_1 x - k_2 s)} p(1-p) < 0, \quad \partial^2 H/\partial s^2 = k_2^2 \mu A e^{(-k_1 x - k_2 s)} p(1-p) < 0,$$

and therefore the conditions $\partial H/\partial x = 0$ or $\partial H/\partial s = 0$ imply that there is a point of maximum of H with respect to x or s.

In sum, we have the following types optimal solution:

I. Donor control, $x^* > 0$, $s^* = 0$, the optimality conditions $H_x = 0$, $H_s < 0$, mean that $p < p_S$ and

$$x^* = \frac{1}{k_1} \ln\left(\frac{-k_1 A \mu (1-p)}{w_x} \right); \tag{13}$$

II. Recipient control, $x^* = 0$, $s^* > 0$, the optimality conditions $H_x < 0$, $H_s = 0$ mean that $p > p_S$ and

$$s^* = \frac{1}{k_2} \ln\left(\frac{-k_2 A \mu p}{w_s} \right); \tag{14}$$

III. No control, $x^* = 0$, $s^* = 0$, the optimality conditions $H_x < 0$, $H_s < 0$ give $k_1 A(1-p)\mu > -w_x$ and $k_2 A p \mu > -w_s$.

Combing these solutions gives

$$\mu > \mu_{SW}(p) = \max\left\{ -\frac{w_x}{k_1 A(1-p)}, -\frac{w_s}{k_2 A p} \right\}. \tag{15}$$

The curve $\mu = \mu_{SW}(p)$ corresponds to another switching condition: from positive control to no

control.

As we have expressed x^*, s^* through p and μ, we can rewrite (7), (8) as an autonomous system of ordinary differentiable equation (ODE)

$$\frac{dp}{dt} = F\big(p, x^*(p,\mu), s^*(p,\mu)\big) = \tilde{F}(p,\mu),$$

$$\frac{d\mu}{dt} = G\big(p, \mu, x^*(p,\mu), s^*(p,\mu)\big) = \tilde{G}(p,\mu),$$

(16)

and hence we can apply the technique of phase plane analysis to study it. The curves $\mu=\mu_{SW}(p)$, $0<p<1$, and $p=p_S$, $\mu\leq\mu_{SW}(p)$ split the phase plane into three domains, according to the three types of control. When the trajectories of (16) cross these lines, switching of the control takes place (Fig. 1).

This approach allows us to deduce the behavior of the optimal "aggregate control" $\varphi^*(t)=k_1x^*+k_2s^*$ (remembering that only a single control or none will be nonzero, apart from the switching moments) which is important for construction of solutions.

Lemma 1. *The optimal "aggregate control", $\varphi^*(t)=k_1x^*+k_2s^{*,}$ is a continuous function of time on [0,T], and it is piecewise differentiable on this interval.*

A proof is given in the Appendix.

Corollary 1. *The flow of (16) is C^1 within each of the regions I, II, III, and is C^0 for $p>0$, $\mu<0$.*

Corollary 2. *At the point $t=t_1$ of control switching from $x>0$ to $s>0$ ($p=p_S$) $\lim_{t\to t_1^-} k_1x(t) = \lim_{t\to t_1^+} k_2s(t)$. At switching points where the aggregate control turns to zero, the value of $x^*=0$ or $s^*=0$ must be optimal, that is one of the relations (11) must hold with $\varphi^*=0$. This allows us to relate values of p and μ at the switching points as $\mu(1-p)=-w_x/(k_1A)$ or $\mu p=-w_s/(k_2A)$.*

Further simplifications are related with replacing the boundary value problem for p and μ

11

by an initial problem for μ only. The idea is to use p as an independent variable. Since $0<p<1$, $dp/dt>0$ in equation (7), so $p(t)$ is a strictly increasing one-to-one function on $[p_0,p_e]$, where the final proportion of lakes invaded is

$$p_e=p(T).$$

Therefore $p(t)$ has an inverse $t(p)$. Substituting it into $x(t)$, $s(t)$, and $\mu(t)$ we obtain the following:

Lemma 2. *The controls x, s, and the shadow price μ can be expressed as functions of p: x(p), s(p), and $\mu(p)$, defined on the interval $[p_0,p_e]$, $p_e=p(T)$.*

This lemma allows us to solve only one equation instead of two in (16),

$$\frac{d\mu}{dp} = \frac{\tilde{G}(p,\mu)}{\tilde{F}(p,\mu)}. \tag{17}$$

We can solve this numerically, or sometimes analytically, for any p_e such that $0<p_0<p_e<1$. The procedure is quite simple — find the equation's solution within each of the phase plane domains. Since in domains I and II initial conditions are not given, the solution contains an arbitrary constant. For domain III the initial condition is $\mu(p_e)=0$, which then provides a solution $\mu_{III}(p)$. The process is continued until the switching line $\mu=\mu_{SW}(p)$ is intersected at some $p=p_Z$. Suppose this happens at the boundary between domains III and II, that is $p_Z>p_S$. Then, since μ is continuous, we have the initial condition for the domain II, $\mu_{II}(p_Z)=\mu_{SW}(p_Z)$, and due to continuity of the aggregate control, $s(p_Z)=0$. Hence we have the solution inside domain II with $s(p)$ and $\mu_{II}(p)$. Again continue the process till the next switching line $p=p_S$, and obtain the initial condition for the domain I, $k_1x(p_S)=k_2s(p_S)$ and $\mu_I(p_S)=\mu_{II}(p_S)$. Then we can construct the solution in the domain I and continue it down to $p=p_0$. If p_e is small enough, the solution does not enter the domain II ($p_0<p_Z\leq p_S$), then the recipient lake control is never used, and solution remains within domains I and III. Finally, it may appear that p_e is so small, that $p_Z\leq p_0$, and then the solution

always remains within domain III, and hence it is optimal for there to be a complete absence of any control. These situations are illustrated in Fig. 1, panels b and c.

However, it is not known which p_e corresponds to the given T. In the Appendix we show that there is one-to-one dependence between T and p_e, and the functions $p_e(T)$ and $T(p_e)$ are monotonous and increasing, as in Fig. 1d. Some examples calculated numerically are shown in Fig. 4. This allows us to set up an iterative numerical procedure for solving a problem with the given time horizon $T=T_0$:

a) Set some p_e value,

b) For the given p_e solve (17), and find $\mu(p)$;

c) Solve $\dfrac{dt}{dp} = -\dfrac{1}{\tilde{F}(p, \mu(p))}$, then $t(p_0)=T(p_e)$.

d) If within given accuracy $T(p_e)\approx T_0$, then stop, otherwise update p_e and return to step b.

This numerical algorithm proved quite simple and in most cases more efficient than the method of gradient projection (Rosen, 1960), often used to solve optimal control tasks.

4. PROPERTIES OF THE OPTIMAL SOLUTION

First, note that in domains I and II it is more convenient to express μ through x or s and p, and to obtain the equation for dx/dp or ds/dp respectively, which is simpler than (17). The resulting equations for all three domains are (see Appendix for more details),

$$\text{Domain I: } \frac{d}{dp}\left[k_1 px + \left(1 + \frac{k_1 g}{w_x}\right)p\right] = \frac{r}{A}\frac{e^{k_1 x}}{(1-p)}, \qquad \mu = -\frac{w_x e^{k_1 x}}{k_1 A(1-p)}, \tag{18}$$

$$\text{Domain II: } \frac{d}{dp}\left[k_2(1-p)s + \left(\frac{k_2 g}{w_s} - 1\right)p\right] = \frac{r}{A}\frac{e^{k_2 s}}{p}, \qquad \mu = -\frac{w_s e^{k_2 s}}{k_2 A p}, \tag{19}$$

13

Domain III:
$$\frac{d}{dp}\left[p(1-p)\mu - \frac{g}{A}p\right] = \frac{r}{A}\mu. \tag{20}$$

In the case $r=0$ the equations (18)–(20) can be solved analytically. The solution satisfying $\mu(p_e)=0$ for the domain III is

$$\mu(p) = -\frac{g}{A}\frac{p_e - p}{p(1-p)}, \qquad \max\{p_0, p_Z\} \le p \le p_e, \quad r = 0,$$

which allows us to obtain the expression for p_Z as the intersection point with $\mu_{SW}(p)$ (15),

$$p_Z = \max\left\{\frac{k_2 g p_e - w_s}{k_2 g - w_s}, \frac{k_1 g p_e}{k_1 g + w_x}\right\}. \tag{21}$$

For domains I and II the analytical solution is obtained as described above, and has the following form ($r=0$)

Domain II:
$$k_2 s(p) = \left(\frac{k_2 g}{w_s} - 1\right)\frac{p_Z - p}{1-p}, \qquad p_S \le p < p_Z \tag{22}$$
$$\text{(only if } p_S < p_Z, \quad p_0 < p_Z\text{)},$$

Domain I:
$$k_1 x(p) = \frac{k_1 g}{w_x}\frac{p_e - p}{p} - 1, \qquad p_0 \le p < \min\{p_S, p_Z\} \tag{23}$$
$$\text{(only if } p_0 < p_Z, \quad p_0 < p_S\text{)}.$$

The trajectory remains within domain III. It is always optimal for zero control provided

$$p_0 \ge \max\left\{\frac{k_2 g p_e - w_s}{k_2 g - w_s}, \frac{k_1 g p_e}{k_1 g + w_x}\right\}.$$

Examples of analytical solutions are shown in Fig. 2.

For $r>0$ the equations (18)–(20) cannot be solved analytically and we can only make a number of comparison results. For berevity, only the main points of the comparison results are discussed, although complete statements and proofs are in the Appendix. It is natural to use analytical solution (21)–(23) as a reference point for the case $r>0$, with other parameters ideally

14

held constant. However, here we come to a problem. If the discount rate is changed to a nonzero value, this typically changes optimal values of controls, and hence changes the terminal p value. Therefore, if we need to compare two solutions with the same time horizon T, they have different p_e and vice versa. Given these issues, a summary of the comparison results finds:

- If we consider two solutions with $r=0$ and $r>0$ and the same p_0 and p_e, then for the solution with $r>0$ controls are less and time horizon is shorter (Proposition 2, Corollary 3, 4).

- If we consider two solutions with $r=0$ and $r>0$ and the same p_0 and T, then the solution for $r>0$ has greater p_e (Corollary 5).

- From the above two statements it is then possible to use the analytical solutions as upper bound estimates for the control values (Corollary 6).

- Consequently, if in the non-discounted case it is optimal for there to be no control, the same is true for any $r>0$ and the same p_e (Corollary 7).

- There is one-to-one correspondence between p_e, $p_0<p_e<1$, and time horizons T. That is, for any such p_e there exists T for which $p(T)=p_e$. There then exists a function $T(p_e)$, which is continuous and strictly increasing (Proposition 3).

The last result is important for the validity of both analytical and numerical techniques.

5. NUMERICAL RESULTS

5.1. Parameters dependence

Various parameter combinations and resulting trajectories are shown in Figure 3. Note that all our equations and solutions depend only on the following four combinations of parameters:

$\frac{k_1 A}{w_x}, \frac{k_2 A}{w_s}, \frac{g}{A}, \frac{r}{A}$, and only p_0 and p_e remain independent. If the parameters are changed in such

a way that these four combinations remain the same, the control pattern does not change. If the solution curve is known for one case, it is not necessary to solve all equations again, only to rescale the existing curve. This property allows an essential reduction in numerical work: setting e.g. $A=w_x=w_s=1$, we can vary only k_i, g, and r (although this means we may use parameter values which at first glance seem unreasonable). For example, $r=3$ never arise in practice (2000% annual), but for r/A this may be a reasonable value if, say, $r=0.03$ and $A=0.01$. The latter value means that the processes develop on a time scale of hundred years.

Unfortunately solutions $x(t)$, $s(t)$, $\mu(t)$ and the relation between p_e and T cannot be found analytically even for $r=0$, except in the case of zero control. Numerically these dependencies can be obtained easily, and it appears that some trajectories may correspond to a very long time horizon. In Fig. 4 the relationship between p_e and T is shown for several parameter combinations and scaled to $\log_{10}AT(p_e)$. This figure illustrates the statement of Corollary 6 in Appendix that implies the greater r, the smaller T for given p_e.

When control is very efficient (large k_i) or if there are large losses g, and the discount rate is 0 or very small, it becomes optimal to maintain high levels of control. This makes the invasion spread very slowly, and hence the values of p close to 1 can be reached only after a very long period of time (T large). Because of this drastic dependence of time horizon on the final invasion level, we consider influence of parameters on control policy for fixed p_e and for fixed T.

5.2 Control regimes for fixed p_e.

The dependence of analytical solution on k_i and g can be seen from the formulas: the greater are their values, the more intensive is the control. This remains true in case $r>0$ (Fig. 6).

Discounting brings three new qualitative features:

1. x remains a decreasing function of p, but its growth as $p \to 0$ becomes limited (Fig. 6), in contrast with undiscounted case where $x \sim 1/p$ (Fig. 3).

2. $s(p)$ may be growing for some p interval, in nondiscounted case it is always decreasing (Figs. 5 and 6).

3. If r/A is big enough (~3), then x- and s-control can be separated by an interval of no control (Fig. 5).

5.3 Control regimes for fixed time horizon T

Returning to the original problem formulation (from the phase plane analysis) introduces the parametric influence of T. In this case the influence on the optimal control pattern from perturbations in other parameters depends on the given T. Fig. 7 shows examples of such controls.

In contrast with the phase plane, it is clear to see that essential dependence of control on k_i is observed only in a relatively small range of k values (the same is true for g). Outside of this range the control saturates; either there is practically no control, or there is full control.

Fig. 8 illustrates this effect from the dependence of control time T_C (that is the total time when x or s is nonzero) on $k=k_1=k_2$ for various g values. For the nondiscounted problem this dependence is rather steep, and the width of the transitional interval is proportional to g. Although discounting makes this interval wider, it remains proportional to g. Overall, these numerical results demonstrate that in many cases the problem of optimal control of invasion can be approximately reduced to a constant level of intensity.

6. CONCLUSIONS

The simple macroscopic model of invasion control developed herein demonstrates a number of important implications for resource management. From analytical and numerical viewpoints optimal prevention policy was shown to be sensitive to several key factors, including the mean economic damage per lake, the efficiency of prevention, the planning horizon, initial magnitude of the invasion, and the discount rate.

While the analytic results are dependent on the choice of functional form, we believe they provide substantial insight into the invasion process at the macroscopic level. Given the choice of functional form, at each moment it is optimal to implement only one type of prevention (donor or recipient) or none at all. There may be a moment of control switching, which depends on the efficiency and per unit cost of each control.

Through a comprehensive phase-plane analysis on the dynamical system closed-loop optimal control polices were derived for the limiting case of a zero discount rate. For the general case a simple and stable numerical algorithm was presented for the control problem. In the phase-plane the complexity of usual methods was substantially reduced. With a single equation in the phase plane the usual method requiring the equation governing the dynamics of the stock to be integrated forward, and that for the shadow price backward, was considerably simplified. While a full analytical characterization of the system proved impossible, several comparison propositions were developed and provide considerable insight into the dynamics of the system.

There is one more important consequence arising from the models considered. If managers are confined to only considering the control of an invader spreading from lake to lake it is possible to only delay the total invasion, not completely prevent it. There are natural mechanisms of invasion spread for which there are no controls. Through the treatment of boats it may be possible to provide extra time for designing other control measures enabling the existing

lake ecosystems either to destroy invaders or to integrate them without allowing catastrophic abundance.

19

APPENDIX

This appendix contains proofs for results in the main body of the paper.

Proposition 1. *The shadow price $\mu(t)$ for t<T is always negative provided the invasion losses*

g>0.

Proof. According to the boundary conditions, $\mu(T)=0$, which implies $\partial H / \partial x = -w_x p < 0$ and

$\partial H / \partial s = -w_s (1-p) < 0$ at *t=T*. In this case, the maximum of *H* is reached at *x=s=0*, and at the

last moment zero control is optimal. Therefore $d\mu/dt(T)=g>0$, hence near *T* for *t<T* $\mu(t)<0$. To

attain nonnegative values the continuous function $\mu(t)$ must cross the *t* axis at least once, and at

the point of crossing $d\mu/dt \leq 0$. At this point $\mu(t_C)=0$. Hence $\partial H / \partial x < 0$ and $\partial H / \partial s < 0$ at *t=t_C*,

and therefore $x(t_C)=s(t_C)=0$. This implies that $d\mu/dt(t_C)=g>0$, which gives a contradiction proving

the proposition. ☐

Lemma 1. *The optimal "aggregate control", $\varphi^*(t)=k_1 x^* + k_2 s^*$ is a continuous function of time on*

[0,T], and it is piecewise differentiable on this interval.

Proof. If $\varphi(t)>0$, one of the optimality conditions (11) or (8a) must be satisfied. Let $p<p_S$, then

$s^*=0$, and hence (11) is satisfied for x^* and s^*, and $e^{\varphi^*} = k_1 |\mu| A(1-p)/w_c$. Since both $p(t)$ and

$\mu(t)$ are continuous, so is $\varphi^*(t)$. Similarly, in case $\varphi^*(t)>0$ and $p>p_S$ we obtain continuity from

(8a). For $p=p_S$ continuity follows from both conditions. In case $\varphi^*(t)=0$ the statement of the

theorem is trivial. It remains to show that the control cannot stop abruptly, that is $\varphi^*(t)$ cannot

jump from a positive value down to zero. Consider the continuous function $F(t)=k_1|\mu(t)|A(1-p(t))/w_2$. Our approach is to assume that for $t<t_{SW}$ $\varphi^*(t)>f_0>0$, and for $t>t_{SW}$ $\varphi^*(t)=0$, and then to

proceed by showing a contradiction. Without loss of generality, let (11) be satisfied for $t<t_{SW}$.

Then $F(t) = e^{\varphi^*} > e^{f_0} > 1$. For $t>t_{SW}$ optimal $s^*=x^*=0$, which implies $\partial H / \partial x < 0$ and therefore

$F(t)<1$. Hence $F(t)$ has discontinuity at $t=t_{SW}$ having jumped from $e^{(f_0)}$ to a value less than one, without taking any intermediate values, for example, it has not passed through $e^{\left(\frac{f_0}{2}\right)}$. This contradiction implies that $\varphi^*(t)$ must approach zero continuously in time. Hence it is a continuous function on the whole interval $[0,T]$. The continuity of $\varphi^*(t)$ means that equation (7) implies $p(t)$ is continuously differentiable everywhere except switching points, and equation (8) implies $\mu(t)$ has the same property. Therefore $\varphi(t)=\log(F(t))$ is also differentiable everywhere except the switching points. $\square$

Deriving (18), (19), (20).

Domain I, $x \geq 0$, $s=0$. The relation (11) allows us to express x through μ and p,

$$x = \frac{1}{k_1}\ln\left(-\frac{k_1 A \mu (1-p)}{w_x}\right), \qquad e^{-k_1 x} = -\frac{w_x}{k_1 \mu A (1-p)},$$

and to rewrite (7), (8) as a closed system of differential equations, which do not contain unknown controls:

$$\frac{dp}{dt} = -\frac{w_x p}{k_1 \mu},$$

$$\frac{d\mu}{dt} = \mu r + \frac{w_x (1-2p)}{k_1 (1-p)} + g + \frac{w_x}{k_1}\ln\left(-\frac{k_1 A \mu (1-p)}{w_x}\right)$$

and

$$\frac{d\mu}{dp} = -\frac{k_1 \mu}{w_x p}\left[\mu r + \frac{w_x (1-2p)}{k_1 (1-p)} + g + \frac{w_x}{k_1}\ln\left(-\frac{k_1 A \mu (1-p)}{w_x}\right)\right].$$

However it is complicated for the analysis. It is possible to derive a more convenient form for the subsequent analysis by using x through μ and p:

$$\mu = -\frac{w_x e^{k_1 x}}{k_1 A(1-p)} \quad .$$

Then after simplifications we come to the system

$$\frac{dp}{dt} = Ae^{-k_1 x} p(1-p),$$

$$k_1 \frac{dx}{dt} = r - Ae^{-k_1 x}(1-p)\left(1 + \frac{k_1 g}{w_x} + k_1 x\right)$$

and the equation for $x(p)$

$$k_1 \frac{dx}{dp} = \frac{re^{k_1 x}}{Ap(1-p)} - \frac{1 + \dfrac{k_1 g}{w_x} + k_1 x}{p} \quad .$$

After multiplying it by p and combining terms together we obtain (18).

Domain II, $x=0$, $s \geq 0$. In a similar fashion we can express μ through s and p,

$$\mu = -\frac{w_s e^{k_2 s}}{k_2 Ap} \quad .$$

Following the same steps as in domain I, we obtain

$$k_2 \frac{ds}{dp} = \frac{re^{k_2 s}}{Ap(1-p)} - \frac{\dfrac{k_2 g}{w_s} - 1 - k_2 s}{1-p},$$

After multiplying by $1-p$ and combining terms, it can be written in the form (19).

Domain III, $x=0$, $s=0$. In this case the control terms vanish, and we can use (7), (8) directly,

$$\frac{dp}{dt} = Ap(1-p),$$

$$\frac{d\mu}{dt} = \mu[r - A(1-2p)] + g \quad .$$

Then (17) takes the form

$$\frac{d\mu}{dp} = \mu\left[\frac{r}{Ap(1-p)} - \frac{1-2p}{p(1-p)}\right] + \frac{g}{Ap(1-p)}.$$

After multiplying by $p(1-p)$ and combining terms it can be written as (20).

Comparison results for $r>0$

Proposition 2 (comparison). *Let $x(p)$, be the solution to (18) for $r>0$ on $[p_1,p_2]$, and $\tilde{x}(p)$ the solutions for $r=0$ such that $\tilde{x}(p_2) \geq x(p_2)$. Then $\tilde{x}(p) \geq x(p)$ on $[p_1,p_2]$. Similar propositions can be proved for $s(p)$, and for $\mu(p)$: if $s(p)$, $\mu(p)$, $\tilde{s}(p)$, $\tilde{\mu}(p)$ the solutions to (19) and (20), and $\tilde{s}(p_2) \geq s(p_2)$, $\tilde{\mu}(p_2) \leq \mu(p_2)$, then $\tilde{s}(p) \geq s(p)$, $\tilde{\mu}(p) \leq \mu(p)$ on $[p_1,p_2]$ respectively.*

Proof. We can consider only $x(p)$; the other two cases can be proved similarly. We have

$$\frac{d}{dp}\left[k_1 px + \left(1 + \frac{k_1 g}{w_x}\right)p\right] = \frac{r}{A}\frac{e^{k_1 x}}{(1-p)}, \qquad \frac{d}{dp}\left[k_1 p\tilde{x} + \left(1 + \frac{k_1 g}{w_x}\right)p\right] = 0.$$

Subtracting the second from the first we have

$$\frac{d}{dp}[k_1 p(x - \tilde{x})] = \frac{r}{A}\frac{e^{k_1 x}}{(1-p)} > 0.$$

Since $p>0$, $k_1 p(x(p) - \tilde{x}(p)) < k_1 p_2(x(p_2) - \tilde{x}(p_2)) \leq 0$, or $\tilde{x}(p) \geq x(p)$. If $\tilde{x}(p_2) > x(p_2)$, then $\tilde{x}(p) > x(p)$. □

Corollary 3. *Let us denote by $x(p)$, $s(p)$, $\mu(p)$ and $\tilde{x}(p)$, $\tilde{s}(p)$, $\tilde{\mu}(p)$ to be the solution to the optimal control problem for the same initial and final infection levels p_0 and p_e, and $r>0$ and $r=0$ respectively. Then, $\tilde{x}(p) \geq x(p)$, $\tilde{s}(p) \geq s(p)$, $\tilde{\mu}(p) \leq \mu(p)$, and for values of p where, $\tilde{x}(p) > 0$ or $\tilde{s}(p) > 0$, or $\tilde{\mu}(p) < 0$ the corresponding inequality is strict, that is $\tilde{x}(p) > x(p)$, $\tilde{s}(p) > s(p)$, or $\tilde{\mu}(p) < \mu(p)$.*

Proof. It is convenient to analyze the situation backwards, from p_e to p_0. As $\tilde{\mu}(p_e) = \mu(p_e) = 0$,

23

according to Proposition 2, for $p<p_e$ $\tilde{\mu}(p)<\mu(p)$ where they satisfy (20) i.e. before they

intersect the switching curve $\mu_{SW}(p)$ (15). Since $\tilde{\mu}(p)<\mu(p)$, $\tilde{\mu}(p)$ intersects the switching

curve first. This implies for the switching points $\tilde{p}_Z > p_Z$. On the interval $[p_Z,\tilde{p}_Z]$ $\tilde{\mu}(p)<\mu(p)$

because they are separated by the switching curve. For simplicity assume that $p_Z>p_S$, then $s(p_Z)=0$

while $\tilde{s}(p_Z)>0$. If Proposition 2 is applied again on the interval $[p_S,p_Z]$ then $\tilde{s}(p)>s(p)$.

Manipulating the optimality condition so that $\mu = -\dfrac{w_s}{k_2 Ap}e^{k_2 s}$ then $\tilde{\mu}(p)<\mu(p)$. Finally for $p=p_S$

$\tilde{x}(p_S)>x(p_S)$, and hence $\tilde{x}(p)>x(p)$ $p<p_S$. Applying $\mu = -\dfrac{w_x}{k_1 A(1-p)}e^{k_1 x}$ allows $\tilde{\mu}(p)<\mu(p)$

to be obtained. Situations $\tilde{p}_Z > p_S > p_Z$ and $p_S > \tilde{p}_Z > p_Z$ can be analyzed similarly. If for the

non-discounted problem the optimal choice is no control, then solutions coincide.

Corollary 4. *Let the conditions of Corollary 3 be satisfied. Then time horizon for the discounted*

problem is shorter than that for the non-discounted problem, $T<\tilde{T}$.

Proof. We can write

$$T = T(p_e) = \int_{p_0}^{p_e} \frac{1}{(dp/dt)}dp = \int_{p_0}^{p_e} \frac{e^{k_1 x(p)+k_2 s(p)}}{Ap(1-p)}dp. \qquad (24)$$

Since $\tilde{x}(p)>x(p)$ for $p_0{\leq}p<\min(p_S,p_Z)$, and the integrand is always positive, we obtain the

statement of the Corollary. ☐

Corollary 5. *If discounted and non-discounted problems have the same time horizon T and the*

initial infection level p_0, the discounted problem has greater final level of infection, $p_e > \tilde{p}_e$.

Proof. For the non-discounted problem $x(p)$, $s(p)$, and p_Z are increasing functions of p_e, and

hence $\tilde{T}$ is an increasing functions of p_e as well. To make time horizons in both problems equal it

is necessary to decrease $\tilde{T}$, and consequently decrease the final level of infection. This proves

the statement.

For very long time horizons T the final level of infection is very close to 1, so it is reasonable to consider what happens in the limit $p_e \rightarrow 1$. Then for the case of $r=0$ analytical solutions give

$$p_Z \rightarrow 1, \quad k_2 s(p) \rightarrow \left(\frac{k_2 g}{w_s} - 1\right), \quad k_1 x(p) \rightarrow \frac{k_1 g}{w_x} \frac{1-p}{p} - 1.$$

Since both x and s increase with p_e, it is possible to obtain bounds for controls:

Corollary 6. *The values of the control variables under optimal control for any r satisfy*

$$0 \le k_1 x(p) \le \frac{k_1 g}{w_x} \frac{1-p}{p} - 1, \quad 0 \le k_2 s(p) \le \left(\frac{k_2 g}{w_s} - 1\right).$$

It follows that the recipient lakes control is used only in case $k_2 g \ge w_s$. From (19), when $p=p_Z$, s turns to zero and cannot be growing, so $ds/dp \le 0$. This means that the necessary condition for the recipient lakes control is

$$\frac{r}{Ap_Z} - \frac{k_2 g}{w_s} + 1 \le 0 \quad \text{or} \quad k_2 g \ge w_s \left(1 + \frac{r}{Ap_Z}\right) > w_s \left(1 + \frac{r}{A}\right).$$

Corollary 7 *(sufficient condition for optimality of no control). If in case r=0 (no discounting) it is optimal for a complete absence of control, then for r>0 and the same final level of infection p_e it is also optimal for there to be an absence of control.*

Proof. Consider the backward motion along the trajectory from p_e to p_0 for both problems. Control then turns on when the trajectory crosses the switching curve. According to Corollary 5, $\tilde{\mu}(p) < \mu(p)$ for $p < p_e$, therefore $\tilde{\mu}(p)$ must cross the switching curve first. If it has not crossed it till p_0 (when no control is optimal), the same is true for $\mu(p)$ as well. This means that for $r>0$ the absence of control is optimal.

Proposition 3. *For every $0 < p_0 < 1$ and finite positive T_0 there exists $p_e > p_0$, such that $T(p_e)=T_0$.*

Proof. For any $p_0 < p_e < 1$ there exists unique bounded solutions of the optimal control problem for $\mu(p)$, $x(p)$, $s(p)$, and therefore it is possible to define function $T(p_e)$ (24). As the solution of the ODE continuously depends on initial data this function is continuous, and monotonically increases everywhere that it exists. The latter follows from topological considerations. Consider two solutions $\mu_1(p)$ and $\mu_2(p)$, $p_{e1} < p_{e2}$. Since they are trajectories on the plane, they cannot cross, and hence $\mu_1(p) > \mu_2(p)$ for $p_0 < p < p_{e1}$. From the optimality conditions it follows that for the same p values $x_1 \leq x_2$, $s_1 \leq s_2$, and according to (24) $T_1 = T(p_{e1}) < T_2 = T(p_{e2})$.

To demonstrate that the domain of $T(p_e)$ is $p_0 \leq p_e < 1$ and its range is $T > 0$, explicit bounds are determined from below and from above. From Corollary 6 it follows that

$$0 \leq k_1 x + k_2 s \leq \max\left\{\frac{k_1 g}{w_x}\frac{1-p}{p} - 1, \frac{k_2 g}{w_s} - 1\right\} < \max\left\{\frac{k_1 g}{w_x p_S}, \frac{k_2 g}{w_s}\right\} \equiv C .$$

Consider two equations

$$\frac{du}{dt} = Au(1-u), \quad \frac{dv}{dt} = Ae^{-C}v(1-v), \quad u(0) = v(0) = p_0 .$$

Comparing with (7), we obtain that

$$u \geq p > v, \quad \frac{du}{dt}(p) \geq \frac{dp}{dt}(p) > \frac{dv}{dt}(p), \quad p_0 < p < 1 ,$$

therefore from (24)

$$\int_{p_0}^{p_e} \frac{1}{Au(1-u)} du \leq T(p_e) < \int_{p_0}^{p_e} \frac{e^C}{Au(1-u)} du$$

or

$$\frac{1}{A}\ln\left(\frac{p_e(1-p_0)}{p_0(1-p_e)}\right) \leq T(p_e) < \frac{e^C}{A}\ln\left(\frac{p_e(1-p_0)}{p_0(1-p_e)}\right) .$$

Therefore the domain of $T(p_e)$ is $p_0 \leq p_e < 1$, and it takes any positive values as $p_e \to 1$. Then it has an

inverse on $[0,\infty)$, and hence for any finite $T_0>0$ there exists a corresponding $p_0<p_{e0}<1$, such that

$T(p_{e0})=T_0$, and there exists the corresponding optimal control solution.

References

Buchan, L.A. and Padilla, D.K. 1999. "Estimating the probability of long-distance overland dispersal of invading aquatic species," *Ecological Applications* **9**: 254-265.

Brown, C., L.Lynch, and D.Zilberman, 2002. "The economics of controlling insect-transmitted plant diseases," *American Journal of Agricultural Economics*, **84**(2), 279-291.

Clark C.W. 1990. *Mathematical Bioeconomics. The optimal management of renewable resources*.N.Y., Wiley.

Crocker, T.D. and J. Tschirhart. 1992. "Ecosystems, externalities and economies," *Environmental and Resource Economics* **2**, 551-567.

Johnson L.E., Ricciardi A., and Carlton J.T., 2001. "Overland dispersal of aquatic invasive species: a risk assessment of transient recreational boating." *Ecological Applications,* **11**(6) 1789–1799

Laxminarayan R. 2001."Economics of antibiotic resistance: a theory of optimal use." *Journal of Environmental Economics and Management,* **42**, 183–206.

Levins R. 1969. "Some demographic and genetic consequences of environmental heterogeneity for biological control." *Bull. Entomol. Soc. Am.* **15**, 237-240.

O'Neill C.R. Jr. 1997. "Economic impact of zebra mussels." *Grt. Lakes Res.Review*, **3**(1).

Olson L.J., Roy S., 2002. "The economics of invasive species management: the economics of controlling a stochastic biological invasion." *Amer. J. Agr. Econ*, **84**(5) 1311-1316.

Pontryagin L.S. et al. 1964. *Mathematical theory of optimal processes*. Oxford, N.Y., Pergamon Press, 1964.

Rosen, J. B. 1960, "The Gradient Projection Method for Nonlinear Programming", *J. SIAM*, **8**, 181.

Settle, C., T.D. Crocker, and J.F. Shogren. 2002. "On the joint determination of biological and economic systems," *Ecological Economics*, **42**(1-2), 301-11.

Shigesada N. and Kawasaki K. 1997. *Biological invasions: theory and practice.* Oxford University Press: Oxford.

Turchin P. 2003. *Complex population dynamics*. Princeton University Press, Princeton and Oxford.

Van Kooten G. C. and Bulte E. 2000. *The Economics of Nature*, Blackwell Publishers, Malden, MA, 2000.
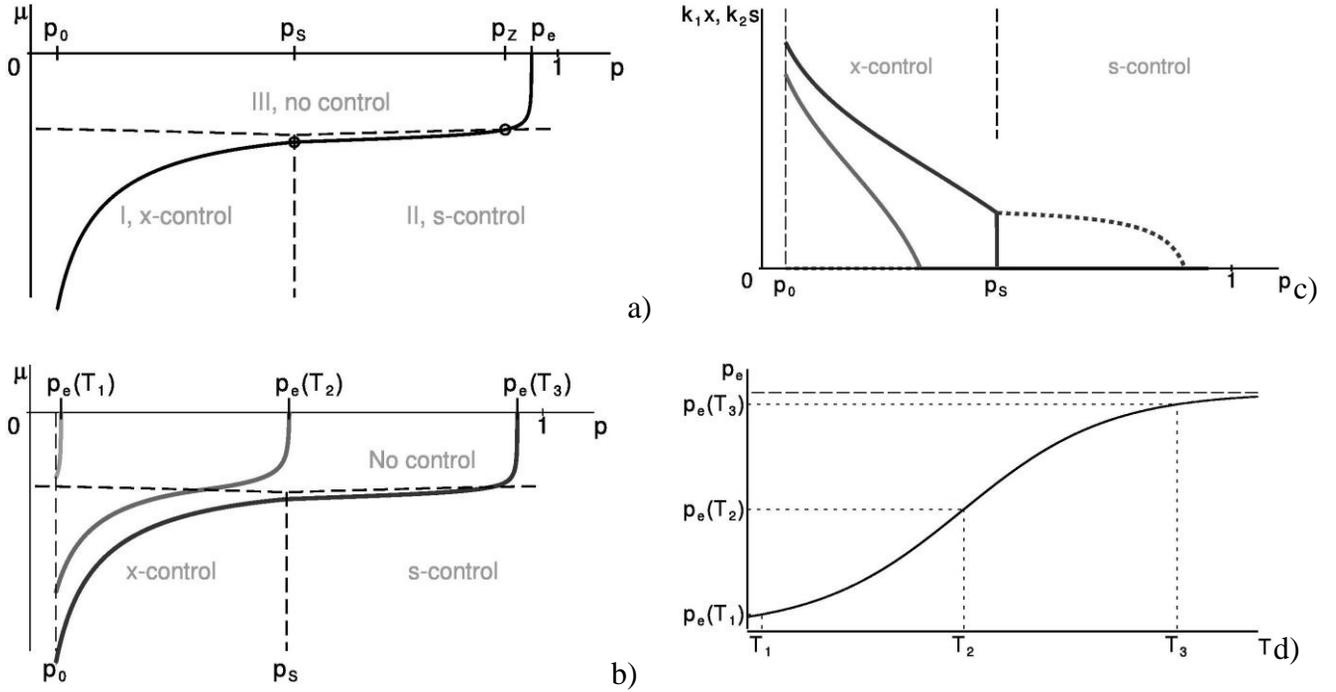
FIGURES

Figure 1. a) Schematically shown are the three regions on the p,μ plane, example trajectory $\mu(p)$, switching points marked by open circles, and characteristic p values — p_0, p_S, p_Z, p_e; b), c) three different control patterns for three different p_e or T (short T_1 has no control; T_2 a period of x-control followed by no control; long T_3 x-control followed by a switch to s-control followed by a switch to no control); d) schematic relation between T and p_e, $p_e \rightarrow 1$ as $T \rightarrow \infty$. In Sect. 6 it is proved that $p_e(T)$ and $T(p_e)$ are monotonically increasing functions.
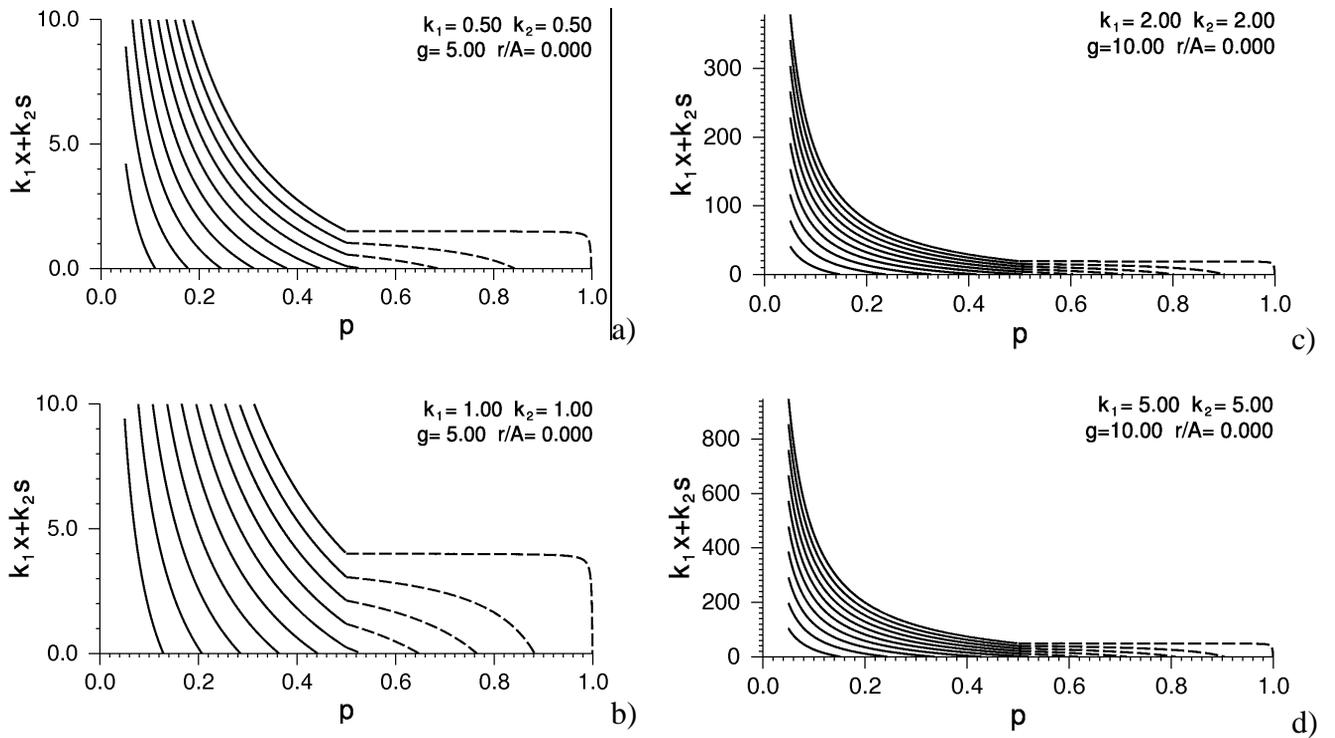
Figure 2. Analytical solutions on the φ-p plane, $\varphi(p)=k_1x(p)+k_2s(p)$. Note that since one of x and s is always zero, the aggregate control shows only the behavior of the other nonzero control variable. Part of the trajectory corresponding to x-control is shown by solid line, the s-control part by dashed line. The trajectories correspond to 10 p_e values between 0.2 and 0.999, p_0=0.05.
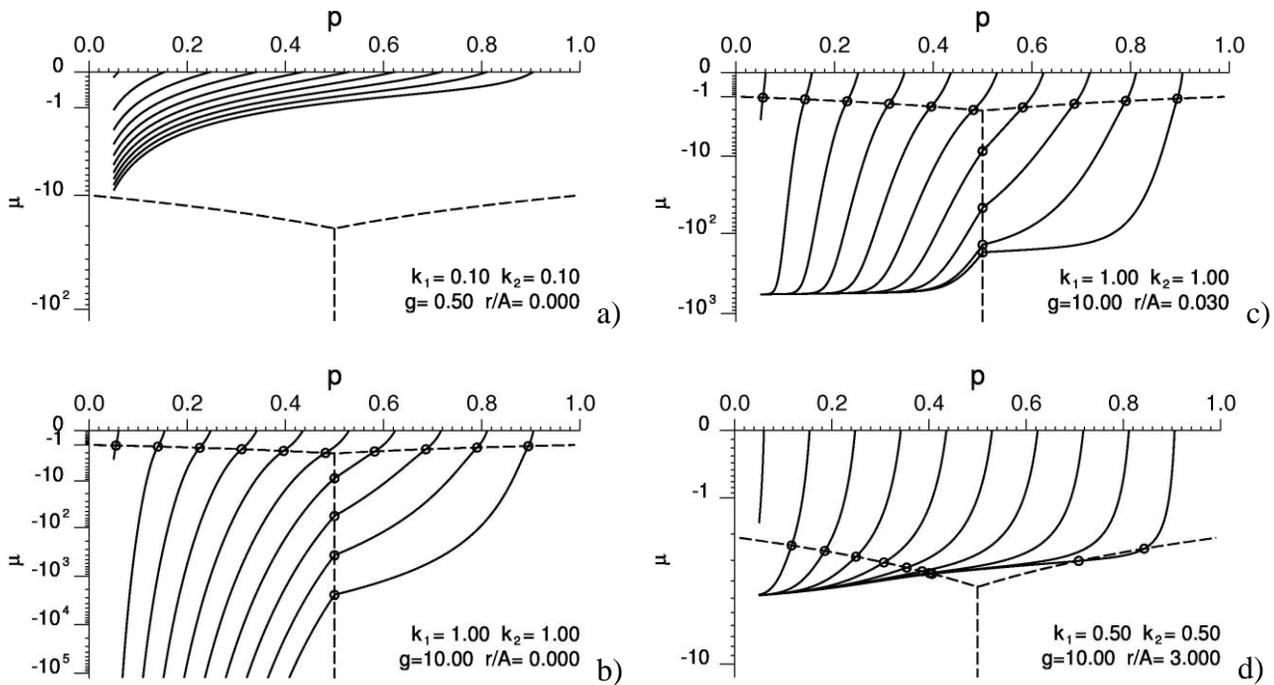
Figure 3. Plots of $\mu(p)$ and switching on the p,μ plane. The trajectories correspond to the values of p_e from

0.1 to 0.9. Control turns on/off when the trajectory crosses the switching curve $\mu_{SW}(p)$ (15) — dashed line.

Switching between x- and s-control occurs when the trajectory crosses the value $p=p_S$. Circles mark

switching points. $p_0=0.05$. a) trajectories did not reach the switching curve, optimal is no control, no

switching. b,c) the trajectories have one or two switching points; d) the trajectory for $p_e=0.9$ has three

switching points. This is a rare control pattern which arises only for large values of r/A: there is x-control for

small p values, then an interval of no control, then an interval of s-control, then no control again.
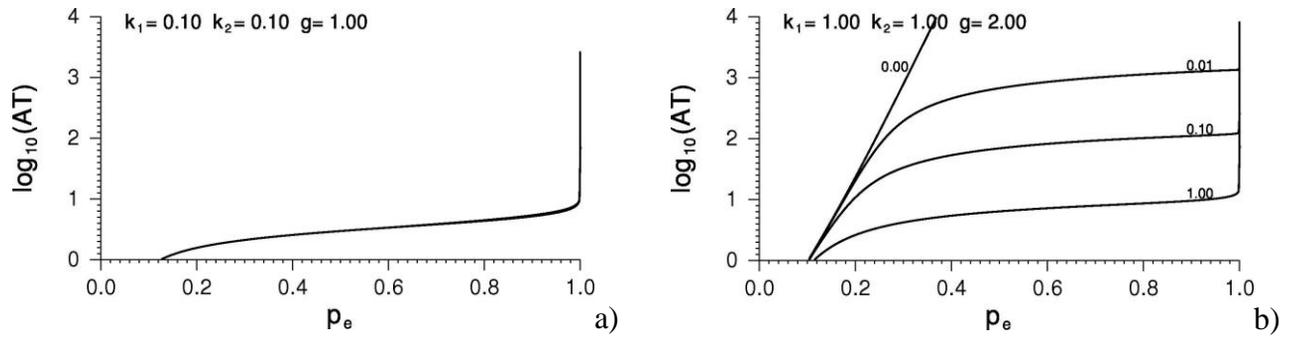
Figure 4. The examples of dependency of $\log_{10}AT$ on p_e for $r/A=0$, 0.01, 0.10, 1. a) Optimal is absence of control, all curves coincide, and can be obtained analytically; b) the greater is r/A, the smaller is T, and the lower is the corresponding curve.
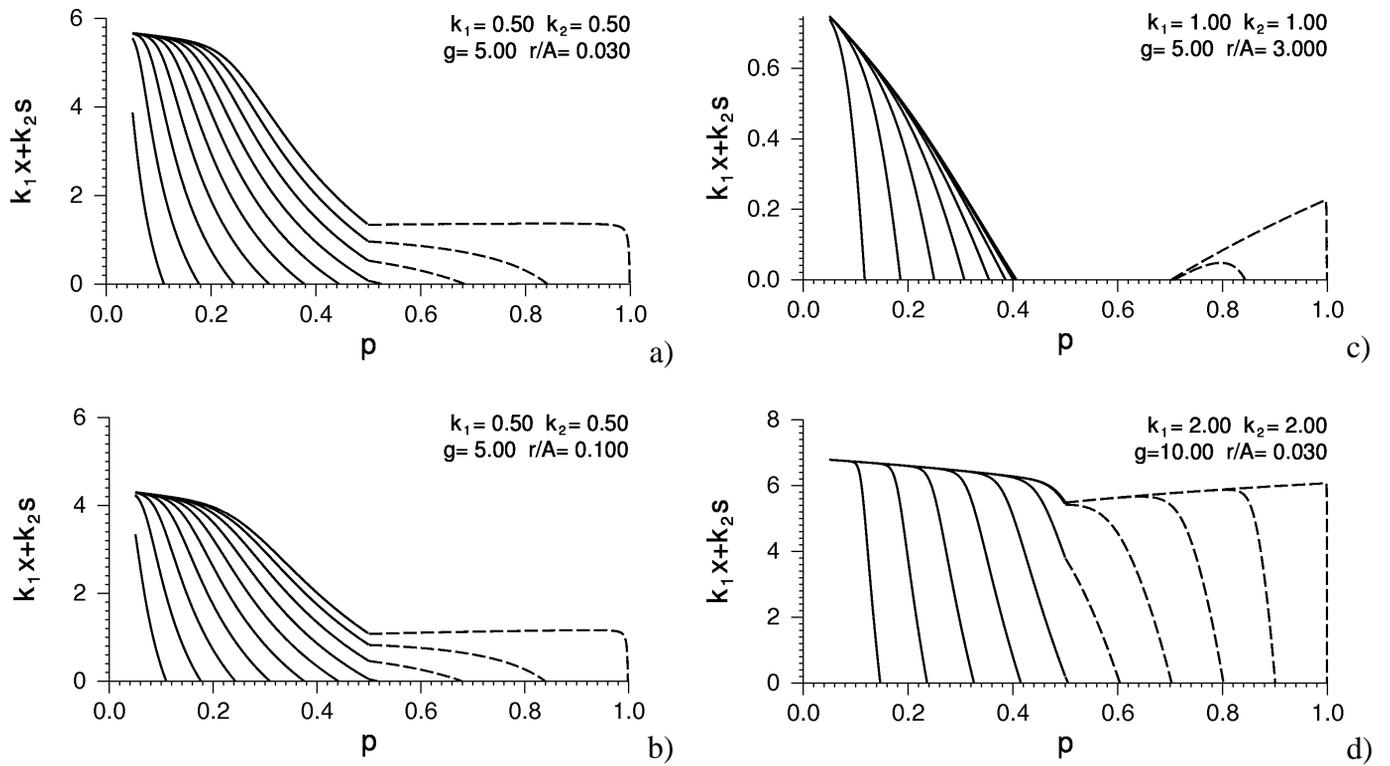
Figure 5. Influence of discounting on behavior of trajectories of the phase plane, compare with Figure 3. a,b) Parameters are those of Figure 3a, but with added discounting of r=0.03 and 0.1; c) An example of separation of x- and s-control at big r/A; d) Same parameters as in Figure 3c, but with discounting r=0.03.
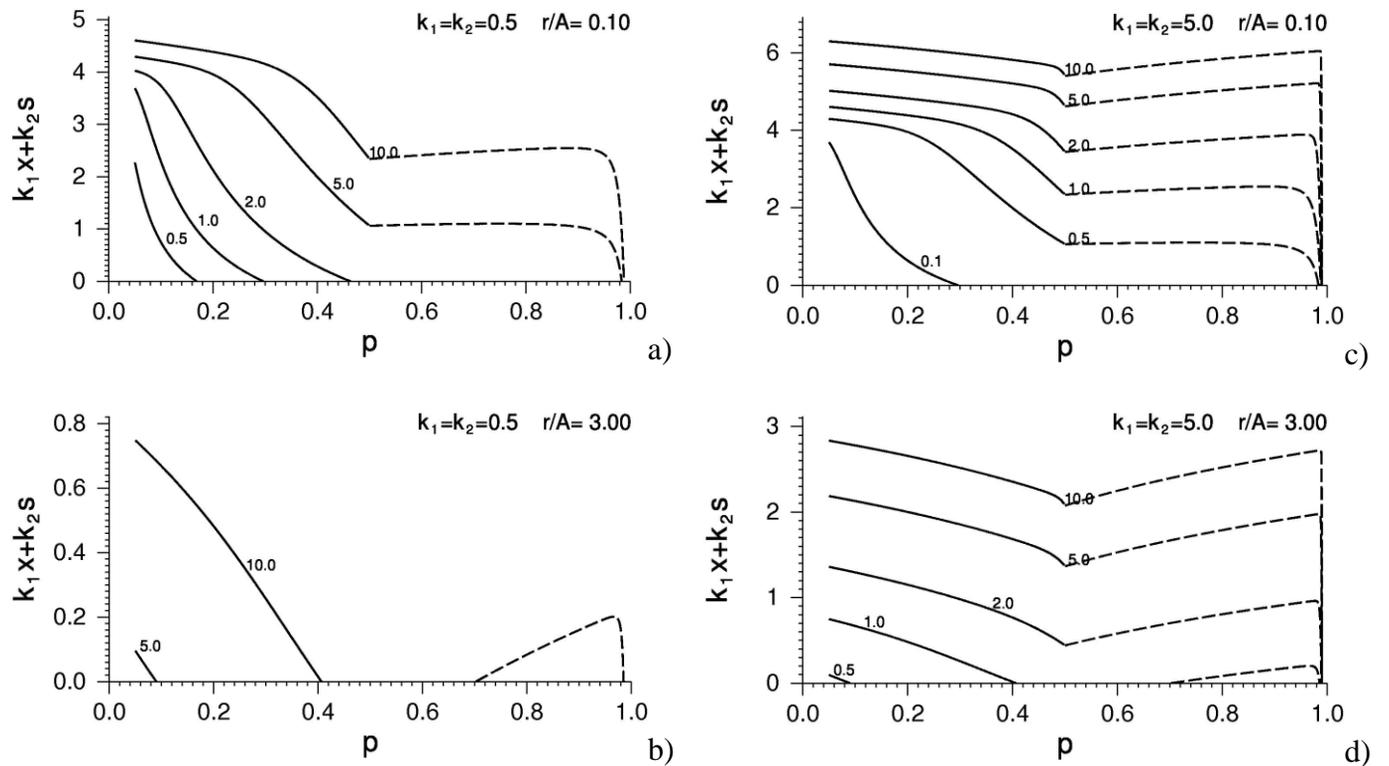
Figure 6. Dependence of control on $k_1=k_2$, g, and r.Shown is the function $\varphi(p)=k_1x(p)+k_2s(p)$, part of the trajectory corresponding to x-control is shown by solid line, the s-control part by dashed line. The trajectories correspond to g=0.1, 0.5, 1.0, 2.0, 5.0, 10.0, the value of g is shown near each trajectory. If a trajectory for some g value is missing, then the absence of control is optimal. Initial invasion level is p_0=0.05.
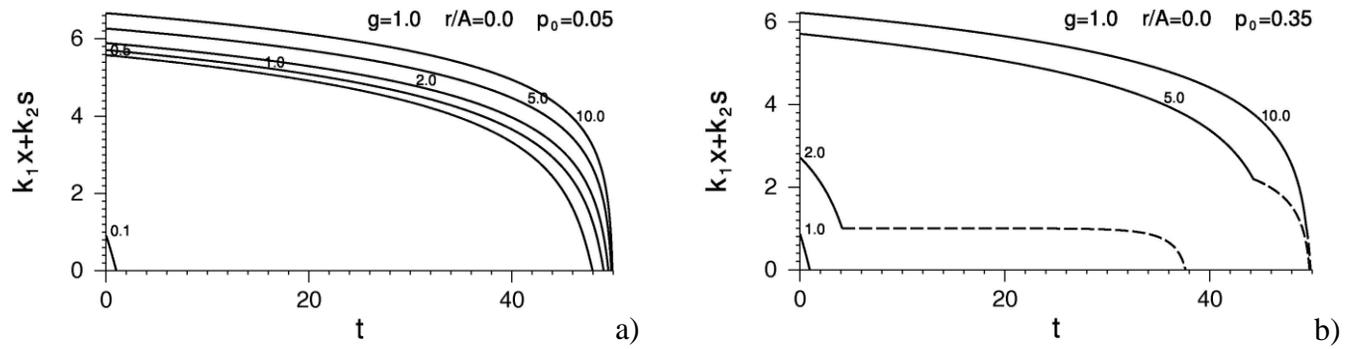
Figure 7. Dependence of control pattern in time on $k_1=k_2=k$, and r for $T=50$. Shown is the function $\varphi(p)=k_1x(p)+k_2s(p)$, part of the trajectory corresponding to x-control is shown by solid line, the s-control part by dashed line. The trajectories correspond to $k=0.1$, 0.5, 1.0, 2.0, 5.0, 10.0, $g=1$. Initial invasion level is $p_0=0.05$ (a) and 0.35 (b).
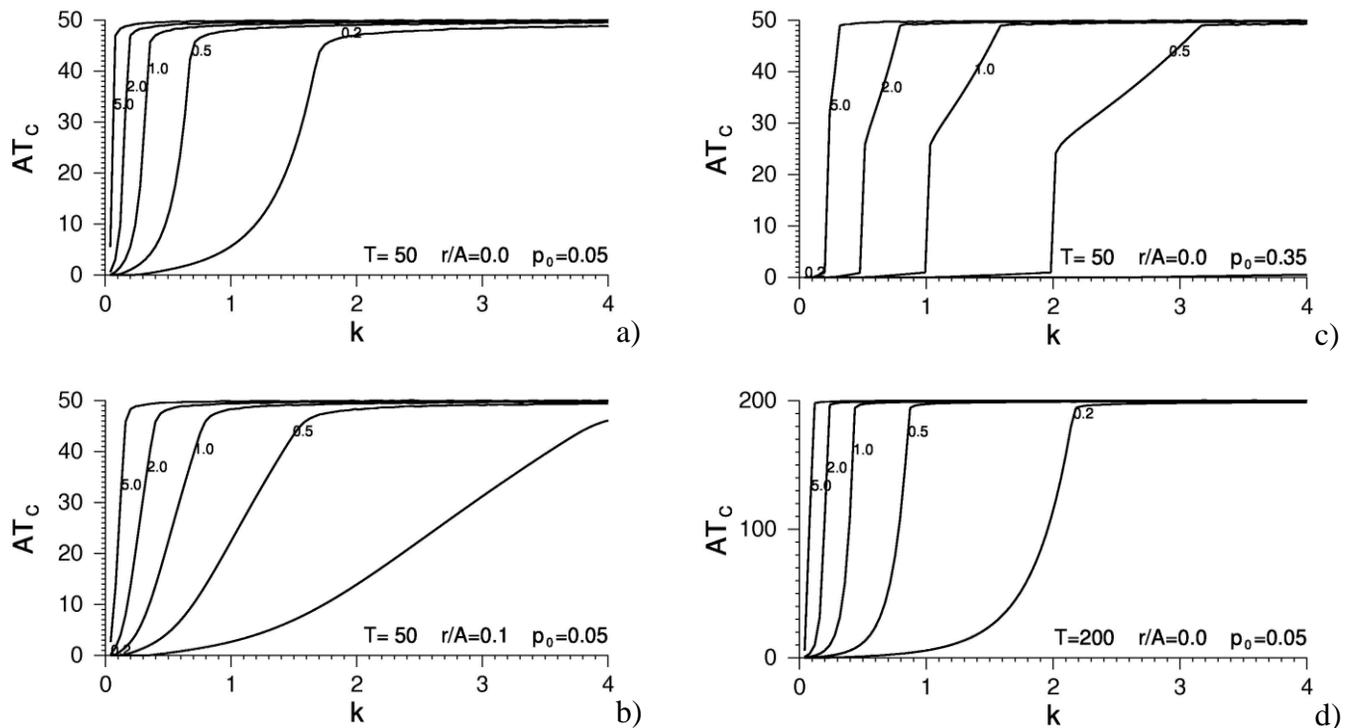

Figure 8. Dependence of control time (when $x+s \neq 0$) on $k_1=k_2=k$, for $g=0.2$, 0.5, 1.0, 2.0, 5.0., values of r, T, and p_0 are shown in the panels.