*J. theor. Biol.* (1997) **185,** 15–29

# Dynamic Multidrug Therapies for HIV: A Control Theoretic Approach

LAWRENCE M. WEIN,† STEFANOS A. ZENIOS‡ AND MARTIN A. NOWAK§

† *Sloan School of Management, M.I.T., Cambridge, MA* 02139, *U.S.A.*, ‡*Graduate School of Business, Stanford University, Stanford, CA 97305, U.S.A. and* § *Department of Zoology, University of Oxford, Oxford OX*1 3*PS, U.K.*

(*Received on* 12 *January* 1996, *Accepted in revised form on* 10 *September* 1996)

Motivated by the inability of current drug treatment to provide long-term benefit to HIV-infected individuals, we derive HIV therapeutic strategies by formulating and analyzing a mathematical control problem. The model tracks the dynamics of uninfected and infected $CD4^+$ cells and free plasma virus, and allows the virus to mutate into various strains. At each point in time, several different therapeutic options are available, where each option corresponds to a combination of reverse transcriptase inhibitors. The controller observes the individual's current status and chooses among the therapeutic options in a dynamic fashion in order to minimize the total viral load. Our initial numerical results suggest that dynamic therapies have the potential to significantly outperform the static protocols that are currently in use; by anticipating and responding to the disease progression, the dynamic strategy reduces the total free virus, increases the uninfected $CD4^+$ count, and delays the emergence of drug-resistant strains.

© 1997 Academic Press Limited

## 1. Introduction

Optimal treatment of the human immunodeficiency virus of type 1 (HIV) infection is the subject of intense research activity. Rapid progress has been made in the development and testing of anti-HIV therapeutic agents, resulting in the approval by the Food and Drug Administration of six reverse transcriptase (RT) inhibitors (AZT, ddI, ddC, d4T, 3TC and nevirapine) and three protease inhibitors (saquinavir, indinavir and ritonavir). These potent drugs inhibit viral replication and lead to a rapid decline in viral abundance, often within days after starting therapy. Unfortunately, RT inhibitors have had only limited success in delaying the onset of AIDS: continual viral replication of HIV, together with the high error rate of reverse transcription of viral RNA into DNA, leads to the emergence of drug-resistant virus strains, typically within weeks or months (depending upon the drug) of treatment initiation (McLeod & Hammer, 1992; Lagakos *et al*., 1993; Volberding *et al*., 1994; Wei *et al*., 1995; Schuurman *et al*., 1995). Moreover, multidrug resistance has been documented in combinations of RT inhibitors (Larder *et al*., 1993; Richman *et al*., 1994; Shafer *et al*., 1994, 1995; Kojima *et al*., 1995) and combinations of protease inhibitors (Condra *et al*., 1995), although Larder *et al*. (1995) and Eron *et al*. (1995) have recently shown that the AZT–3TC combination is able to sustain *in vivo* antiviral effects for at least 1 year.

It would appear that increased effectiveness of HIV therapy could be achieved by developing *dynamic* multidrug approaches, where the combination of drugs received by a patient changes over time in response to the disease progression (e.g. current $CD4^+$ count, viral load, mix of viral strains). In this paper we mathematically model the dynamic multiple drug therapy problem as an optimal control problem that can be informally described as follows: choose an optimal mix of drugs at each point in time in order to minimize the total viral load over some time horizon (e.g. 1 year).

Several mathematical models have been developed that incorporate the effects of therapy on HIV-infected individuals. In a series of papers (Perelson,

0022–5193/97/050015 + 15 $25.00/0/jt960253

© 1997 Academic Press Limited

1989; Perelson *et al*., 1993; Kirschner & Perelson, 1994; Kirschner *et al*., 1995; Kirschner & Webb, 1996), Perelson and Kirschner and their colleagues have studied the timing, frequency and intensity of AZT treatment. Agur (1989) focuses on the optimal tradeoff between the toxicity and efficacy of AZT. The mathematical models considered in these studies do not allow for virus mutations. Nowak *et al*. (1991), McLean & Nowak (1992), Nowak & May (1993), Frost & McLean (1994), Kirschner & Webb (1995) and de Jong *et al*. (1996) analyze descriptive (as opposed to optimization) models for the competitive interaction of AZT-sensitive and AZT-resistant strains of HIV; the papers by Nowak & May, Frost & McLean and Kirschner & Webb also include numerical simulations of alternating and/or combination therapies.

Our deterministic control problem, which is described in Section 2, considers a finite number of virus strains, or quasispecies, and allows mutations from one strain to another. A finite number of therapeutic options are allowed, where each option consists of the simultaneous application of one or more RT inhibitors. The model incorporates uninfected $CD4^+$ T-cells, and infected $CD4^+$ cells and infectious free virus associated with each virus strain; in this context, the RT inhibitors prevent the free virus from successfully infecting uninfected $CD4^+$ cells. Each drug option has a different efficacy against each virus strain, thereby allowing for complex drug–virus interactions.

Because of the high dimensionality of the control problem, we resort to approximation methods in Section 3; more specifically, perturbation methods and the policy improvement algorithm are used to derive a closed form dynamic policy. Several other types of therapeutic approaches, such as protease inhibitors and the reconstitution of the immune system, are considered in Section 4. Results from numerical simulations are reported in Section 5 and discussed in Section 6. Concluding remarks can be found in Section 7.

## 2. Problem Formulation

Our mathematical model incorporates $I$ different strains of HIV. For $i = 1, \ldots, I$, let $y_i(t)$ be the density of $CD4^+$ cells infected by strain $i$ at time $t$, and let $v_i(t)$ denote the density of infectious free virus of strain $i$; non-infectious virions are ignored in our model. If we define $x(t)$ to be the density of uninfected $CD4^+$ cells at time $t$, then the *state* of the system at time $t$ is given by $(x(t), y_1(t), \ldots, y_I(t), v_1(t), \ldots, v_I(t))$, which is denoted by $(x(t), y_i(t), v_i(t))$.

The controller has $J$ therapeutic *options* at time $t$, where each option corresponds to a prespecified *combination* of RT inhibitors, each used in a prespecified dosing schedule. For example, a typical combination might be 100 mg of AZT taken three times daily with 200 mg of ddI taken twice daily. Our control variables $d_j(t)$ satisfy:

$$\sum_{j=1}^{J} d_j(t) \leqslant 1, \tag{1}$$

$$d_j(t) = 0 \text{ or } 1, \tag{2}$$

where $d_j(t) = 1$ if option $j$ is applied at time $t$, and equals zero otherwise. Because the extent of overlapping of toxicity profiles of various drug combinations is extremely difficult to model explicitly, we assume *via* eqns (1) and (2) that at most one drug combination can be used at each point in time; this constraint is without loss of generality, because any number of drug combinations can be incorporated into the model. Although we implicitly assume that the toxicity of each combination remains constant over time, our dynamic policy can be easily adapted to the more realistic setting where the toxicity threshold is an exogenously specified function of either time (e.g. 6 months of intense therapy is alternated with 6 months of light therapy) or the severity of the patient's side effects: we simply define different combinations for different dosage levels of agents, and disallow highly toxic combinations at certain points in time.

We assume that each virus strain has its own infectivity rate, denoted by $\tilde{\beta}_i$, which is the rate that it infects uninfected $CD4^+$ cells; the reason for the "tilde" will become clear in the next section. The RT inhibitors reduce virus infectivity in the following manner. For $i = 1, \ldots, I$ and $j = 1, \ldots, J$, let $p_{ji}$ denote the *efficacy* of drug combination $j$ in blocking new infections by virus strain $i$. Under a generic drug policy $d_j(t)$, the infectivity of virus $i$ is $\tilde{\beta}_i[1 - \Sigma_{j=1}^{J} p_{ji} d_j(t)]$; we assume that the values of $p_{ji}$ are chosen so that the infectivity of each strain is non-negative under all feasible therapeutic strategies.

The dynamics of our system are described by the following set of ordinary differential equations:

$$\dot{x}(t) = \lambda - \left( \mu + \sum_{i=1}^{I} \tilde{\beta}_i v_i(t) \left[ 1 - \sum_{j=1}^{J} p_{ji} d_j(t) \right] \right) x(t), \tag{3}$$

$$\dot{y}_i(t) = \left( \sum_{k=1}^{I} q_{ki} \tilde{\beta}_k v_k(t) \left[ 1 - \sum_{j=1}^{J} \times p_{jk} d_j(t) \right] \right) x(t) - \alpha_i y_i(t), \quad (4)$$

$$v_i(t) = \pi_i y_i(t) - [k_i + \tilde{\beta}_i x(t)] v_i(t). \quad (5)$$

The rate at which uninfected CD4$^+$ cells are invaded by virus strain $i$ at time $t$ is $\tilde{\beta}_i v_i(t) x(t)$, and each of these potential infections leads to a reduction in free virus. The rate of successful infections by strain $i$ is $\tilde{\beta}_i [1 - \Sigma_{j=1}^{J} p_{ji} d_j(t)] v_i(t) x(t)$, and these infections cause a simultaneous decline in uninfected cells $x(t)$ and rise in infected cells $y_i(t)$. The mutation rate $q_{ij}$ in eqn (4) is the fraction of reverse transcriptions of strain $i$ that result in a cell infected by strain $j$. Hence, $\Sigma_{j=1}^{J} q_{ij} = 1$ and the diagonal terms of the mutation matrix are close to one in value, while the off-diagonal terms are nearly zero. Strain $i$ replicates at rate $\pi_i$ after it has infected a CD4$^+$ cell, and thus free virus of strain $i$ is produced at rate $\pi_i y_i(t)$; an alternative interpretation is that $\pi_i/\alpha_i$ virions are produced by a lytic burst during the death of an infected cell. Notice that the quantities $p_{ji}$, $q_{ij}$ and $\pi_i$ can capture a wide range of drug–virus interactions, including both synergistic and antagonistic effects of combination therapy, as observed by St. Clair *et al*. (1991) and Larder *et al*. (1995), and surveyed by Wilson & Hirsch (1995).

Uninfected CD4$^+$ cells increase because of (exponential) proliferation in peripheral tissue compartments (e.g. secondary lymphoid organs) and/or (linear) production from a pool of precursors (e.g. the thymus). For simplicity, we assume a constant production rate $\lambda$. In addition, we let $\mu$, $\alpha_i$ and $k_i$ denote the respective natural death rates of uninfected CD4$^+$ cells, cells infected by strain $i$ and free virus of strain $i$.

Because the primary focus of this paper is on therapeutic regimens and not on natural disease progression, we purposely do not incorporate the human immune response into the model. Hence, we implicitly assume that the strength of the immune response remains constant over the time horizon under study. The model also ignores latently infected cells; although most plasma virus comes from actively infected cells (Ho *et al*., 1995; Wei *et al*., 1995), this does not imply that latently infected cells are unimportant for the emergence of drug resistance. Finally, although the lymph system is the location of considerable production of plasma virus and many new infections (Pantaleo *et al*., 1995), our model focuses on the blood and essentially assumes that the blood and lymph system are in equilibrium.

Let $T$ be the time horizon. Then the mathematical control problem is to choose the binary controls $\{d_j(t), t \geqslant 0\}$ to minimize:

$$\int_0^T \sum_{i=1}^{I} v_i(t) \mathrm{d}t, \quad (6)$$

subject to eqns (1–5).

Our control problem makes two unrealistic demands on the controller. First, we assume that therapy can be changed in a continuous manner, whereas in practice alterations can only be made in a periodic (e.g. weekly) fashion. We also assume that the decision maker, in choosing $d_j(t)$ at time $t$, can observe the current state $(x(t), y_i(t), v_i(t))$. Although recent technology makes it possible to quantify virus load, CD4$^+$ cell counts and virus-infected cells in the blood of infected individuals, these techniques are not currently available for day-to-day treatment of the HIV-infected population at large. A more realistic control problem for current day-to-day treatment would allow the controller to see only a partial observation of the state (e.g. only the total CD4$^+$ count $x(t) + \Sigma_{i=1}^{I} y_i(t)$ and the total viral load $\Sigma_{i=1}^{I} v_i(t)$). The control problem in eqns (1–6) then becomes a nonlinear filtering problem; although partial observations greatly complicate the problem, reasonably good policies (e.g. the linearized Kalman filter on p. 161 of Elliott *et al*., 1994) can be derived. However, these two lines of inquiry are left for future research and hereafter we assume that the controller can observe $(x(t), y_i(t), v_i(t))$ and can change the therapeutic regimen at each time $t$.

## 3. Analysis

The control problem in eqns (1–6) does not appear to admit a closed form solution. Moreover, standard numerical techniques based on Pontryagin's maximum principle (Pontryagin *et al*., 1962) cannot solve a problem of realistic size (e.g. five drug options and 30 virus strains). Hence, we resort to an approximate method, which employs perturbation methods in conjunction with ideas from dynamic programming.

### 3.1. OVERVIEW

Let $V(x, y_i, v_i, t)$ denote the cost incurred [as given in eqn (6)] from time $t$ to time $T$ under the optimal policy, given that the state of the system at time $t$ is $(x, y_i, v_i)$. For ease of notation, we suppress the dependence of $V$ on its arguments. Then

the dynamic programming optimality equation is (Bellman, 1957):

$$-\frac{\partial V}{\partial t} = \sum_{i=1}^{I} v_i + \left( \sum_{i=1}^{I} \frac{\partial V}{\partial v_i} [\pi_i y_i - (k_i + \tilde{\beta}_i x) v_i] \right)$$

$$+ \min_{\{d_j \in \Omega\}} \left\{ \frac{\partial V}{\partial x} \left[ \lambda - \left( \mu + \sum_{i=1}^{I} \tilde{\beta}_i v_i \left[ 1 - \sum_{j=1}^{J} p_{ji} d_j \right] \right) x \right] \right.$$

$$\left. + \sum_{i=1}^{I} \frac{\partial V}{\partial y_i} \left[ \left( \sum_{k=1}^{I} q_{ki} \tilde{\beta}_k v_k \left[ 1 - \sum_{j=1}^{J} p_{jk} d_j \right] \right) x - \alpha_i y_i \right] \right\}, \quad (7)$$

where $\Omega$ denotes the set of admissible controls that satisfy eqns (1) and (2). Our basic approach is as follows: we use asymptotic methods to obtain closed form approximate expressions for the $(2I + 1)$-dimensional process $\{(x(t), y_i(t), v_i(t)\}$ under a static control policy (i.e $d_j(t) = d_j$ for all $t$). From these expressions, we derive a closed form expression for $V$ that has the same interpretation as the $V$ in eqn (7), except that the (approximate) cost is under the static policy, not the optimal policy. Then one iteration of the policy improvement algorithm (Howard, 1960) is performed: we differentiate our approximate closed form expression for $V$ with respect to $x$ and $y_i$, substitute these derivatives into the right side of eqn (7) and perform the embedded minimization in this equation.

This approach yields a closed form dynamic drug control policy: it specifies which drug combination, if any, to use at each point in time in terms of the current state, the current time, and all the problem parameters. If our closed form expression for $V$ was exact, then dynamic programming theory would imply that our proposed policy was better [yields a lower cost in eqn (6)] than the static policy (in fact, if our expression for $V$ was exact then repeated policy improvement iterations would generate a sequence of policies that converges to the optimal policy). However, our expression is not exact, and we cannot draw this conclusion. Nevertheless, this philosophy (finding an approximate value for $V$ and performing one iteration of the policy improvement algorithm) has been used with considerable success in designing dynamic call acceptance/rejection protocols for stochastic models of telephone traffic (e.g. Ott & Krishnan, 1985; Key, 1990).

### 3.2. THE OPTIMAL STATIC POLICY

Because our proposed dynamic policy takes a static policy as its starting point, it is natural to employ the optimal static policy. The static optimization problem is given by the following nonlinear program, where $x$, $y_i$ and $v_i$ are implicit decision variables that are driven by the static drug control $d_j$:

$$\text{minimize}_{x, y_i, v_i, d_j} \sum_{i=1}^{I} v_i \quad (8)$$

$$\text{subject to } \lambda - \left( \mu + \sum_{i=1}^{I} \tilde{\beta}_i v_i \left[ 1 - \sum_{j=1}^{J} p_{ji} d_j \right] \right) x = 0, \quad (9)$$

$$\left( \sum_{k=1}^{I} q_{ki} \tilde{\beta}_k v_k \left[ 1 - \sum_{j=1}^{J} p_{jk} d_j \right] \right) x - \alpha_i y_i = 0, \quad (10)$$

$$\pi_i y_i - [k_i + \tilde{\beta}_i x] v_i = 0, \quad (11)$$

$$\sum_{j=1}^{J} d_j \leqslant 1, \quad (12)$$

$$x, y_i, v_i \geqslant 0, \quad (13)$$

$$d_j = 0 \text{ or } 1. \quad (14)$$

We have been unable to find a closed form solution to eqns (8–14); however, this problem can be solved by direct enumeration of all $J + 1$ feasible static drug policies.

Let $d_j^s$ denote the optimal static policy that solves eqns (8–14). Under this policy, the infectivity of virus $i$ is $\tilde{\beta}_i \bar{d}_i$, where $\bar{d}_i = 1 - \Sigma_{j=1}^{J} p_{ji} d_j^s$. Notice that we have not investigated the possibility of multi-stability or unstable fixed points to eqns (9–11), in which case the static optimization step may yield misleading results. However, it turns out that the policy that emerges from our analysis is quite robust with respect to the starting static policy. This insensitivity can be inferred from eqn (31), where $\bar{d}_i$ appears only in the term that is of order $\epsilon$, and has been borne out by our numerical results (see Section 5).

### 3.3. ASYMPTOTIC ANALYSIS

In this subsection we find a closed form approximate expression for the system trajectory under the optimal static policy. Consider eqns (3–5), with $\bar{d}_i$ taking the place of $1 - \Sigma_{j=1}^{J} p_{ji} d_j(t)$. Keeping in mind that we will use the solution to these equations to derive $V$, let us consider the initial conditions (in this subsection, $s$ is a generic time index and $t$ denotes a specific point in time):

$$x(t) = x, \quad y_i(t) = y_i \quad v_i(t) = v_i. \quad (15)$$

To perform the perturbation analysis, we introduce the parameter $\epsilon$. Without loss of generality, let us

assume that $\tilde{\beta}_1 = \min_{\{1 \leqslant i \leqslant I\}} \tilde{\beta}_i$. Then we let $\epsilon = \tilde{\beta}_1$ and define

$$\beta_i = \frac{\tilde{\beta}_i}{\epsilon}. \tag{16}$$

Although $\beta_1 = 1$, we retain this parameter in the model.

The asymptotic analysis is based on the assumption that $\epsilon$ is small in value (i.e. much less than one). Typical values in the literature for the infectivity parameter $\tilde{\beta}_i$ are roughly $10^{-5}$, which is smaller than the values of the other problem parameters (except the mutation rates); hence, the approximation should perform well.

We assume that our solution is of the form:

$$x(s) = x^{(0)}(s) + \epsilon x^{(1)}(s), \quad y_i(s) = y_i^{(0)}(s) + \epsilon y_i^{(1)}(s), \quad v_i(s) = v_i^{(0)}(s) + \epsilon v_i^{(1)}(s). \tag{17}$$

Although we could define and solve for higher order terms, just deriving the first order terms is quite cumbersome, at least without a computer. We use eqn (16) to replace $\tilde{\beta}_i$ by $\epsilon\beta_i$ and use eqn (17) to substitute in for the system state in eqns (3–5) and eqn (15). For $i = 0$, 1, we collect terms of order $\epsilon^i$ to solve for the unknown processes on the right side of eqn (17). Collecting the constant (i.e. $\epsilon^0$) terms yields the following system of linear differential equations:

$$\dot{x}^{(0)}(s) = \lambda - \mu x^{(0)}(s), \tag{18}$$

$$\dot{y}_i^{(0)}(s) = -\alpha_i y_i^{(0)}(s), \tag{19}$$

$$\dot{v}_i^{(0)}(s) = \pi_i y_i^{(0)}(s) - k_i v_i^{(0)}(s), \tag{20}$$

subject to the initial conditions

$$x^{(0)}(t) = x, \quad y_i^{(0)}(t) = y_i \quad v_i^{(0)}(t) = v_i. \tag{21}$$

It is worth noting that the mathematically perturbed system, eqns (18–21), corresponds precisely to the physical perturbation performed when giving potent RT inhibitors. In Section 4, we use data from Ho *et al*. (1995), Wei *et al*. (1995) and Perelson *et al*. (1996), who used various RT inhibitors and protease inhibitors to perturb the system, to estimate values for the model parameters.

The solution to eqns (18–21) is:

$$x^{(0)}(s) = \frac{\lambda}{\mu} + \left(x - \frac{\lambda}{\mu}\right)e^{-\mu(s-t)}, \tag{22}$$

$$y_i^{(0)}(s) = y_i e^{-\alpha_i(s-t)}, \tag{23}$$

$$v_i^{(0)}(s) = \frac{\pi_i y_i}{k_i - \alpha_i} e^{-\alpha_i(s-t)} + \left(v_i - \frac{\pi_i y_i}{k_i - \alpha_i}\right)e^{-k_i(s-t)}. \tag{24}$$

The order $\epsilon$ terms lead to the following system of differential equations:

$$\dot{x}^{(1)}(s) = -\mu x^{(1)}(s) - x^{(0)}(s) \sum_{i=1}^{I} \beta_i \bar{d}_i v_i^{(0)}(s), \tag{25}$$

$$\dot{y}_i^{(1)}(s) = x^{(0)}(s)\left(\sum_{k=1}^{I} q_{ki}\beta_k \bar{d}_k v_k^{(0)}(s)\right) - \alpha_i y_i^{(1)}(s), \tag{26}$$

$$\dot{v}_i^{(1)}(s) = \pi_i y_i^{(1)}(s) - k_i v_i^{(1)}(s) - \beta_i x^{(0)}(s) v_i^{(0)}(s), \tag{27}$$

subject to the initial conditions

$$x^{(1)}(t) = 0, \quad y_i^{(1)}(t) = 0, \quad v_i^{(1)}(t) = 0. \tag{28}$$

Now we substitute the solution to eqns (22–24) into eqns (25–27) and solve the linear differential eqns (25–28) sequentially using standard methods (Carrier & Pearson, 1991). The solution is given in Appendix A.

### 3.4. THE PROPOSED DYNAMIC POLICY

With the solution to the asymptotic systems in eqns (18–21) and eqns (25–28) in hand, we can estimate the value function $V$ by $\int_t^T \Sigma_{i=1}^I (v_i^{(0)}(s) + \epsilon v_i^{(1)}(s))\mathrm{d}s$. This integration is carried out in Appendix B. The proposed solution is greatly simplified if we assume that the value function is independent of the time horizon. In practice, the time horizon $T$ is very large (i.e. in years) in relation to the time scale of the systems dynamics, which change on a daily basis (Ho *et al*., 1995, Wei *et al*., 1995); hence, this is a natural assumption. Consequently, we set $T = \infty$ and $t = 0$ in eqn (46) to obtain a value function (and hence a policy) that is independent of the time horizon; this approximate, stationary value function is denoted by $V_\epsilon$, to distinguish it from the value function in eqn (7). Differentiating eqn (46) with respect to $x$ and $y_i$ yields (with $T = \infty$ and $t = 0$):

$$\frac{\partial V_\epsilon}{\partial x} = \epsilon \sum_{i=1}^{I} \times \left(\bar{d}_i \sum_{j=1}^{I} \frac{q_{ij}\pi_j}{\alpha_j k_j} - \frac{1}{k_i}\right)\left(\frac{\beta_i}{k_i + \mu}\right)\left(\frac{\pi_i y_i}{\alpha_i + \mu} + v_i\right), \tag{29}$$

$$\frac{\partial V_\epsilon}{\partial y_i} = \frac{\pi_i}{\alpha_i k_i} + \epsilon\left[\pi_i\beta_i\left(\frac{\lambda}{\alpha_i k_i \mu} + \frac{x - \lambda/\mu}{(\alpha_i + \mu)(k_i + \mu)}\right) \times \left(\bar{d}_i \sum_{j=1}^{I} \frac{\pi_j q_{ij}}{\alpha_j k_j} - \frac{1}{k_i}\right)\right]. \tag{30}$$

Now we can substitute these derivatives into eqn (7) and perform the minimization. Because the function

to be minimized is linear in our controls $d_j(t)$, the solution $d_j^*(t)$ can be found in closed form. Let us define the *dynamic quantities*:

$$c_i(t) = \tilde{\beta}_i v_i(t)\left[\sum_{l=1}^{I} \frac{q_{il}\pi_l}{\alpha_l k_l} + \epsilon\beta_l\left(\frac{\pi_l q_{ll}\bar{d}_l}{\alpha_l k_l} - \frac{1}{k_i} + \sum_{j\neq i}^{I} \frac{\pi_j q_{ij}\bar{d}_i}{\alpha_j k_j}\right)\right.$$

$$\times\left(\frac{\pi_l[q_{il}\lambda(\alpha_l + k_l + \mu) + \alpha_l k_l(q_{il}x(t) - y_l(t)]}{\alpha_l k_l(\alpha_l + \mu)(k_l + \mu)}\right.$$

$$\left.\left. - \frac{v_l(t)}{k_l + \mu}\right)\right]. \tag{31}$$

Then the proposed policy is: *apply no therapy* (*i.e.* $d_j^*(t) = 0$ *for* $j = 1, \ldots, J$) *if*:

$$\max_{\{1\leqslant j\leqslant J\}} \sum_{i=1}^{I} p_{ji}c_i(t) \leqslant 0; \tag{32}$$

*otherwise*, *use drug combination* $j^*$ (*i.e.* $d_{j^*}^*(t) = 1$, $d_j^*(t) = 0$ *for* $j \neq j^*$), *where*:

$$j^* = \arg\max_{\{1\leqslant j\leqslant J\}} \sum_{i=1}^{I} p_{ji}c_i(t). \tag{33}$$

We conclude this subsection with several remarks.

The proposed therapeutic strategy is a *dynamic index policy*, where drug combination $j$ has the dynamic index $\Sigma_{i=1}^{I} p_{ji}c_i(t)$, and at each point in time the policy uses the drug combination that possesses the largest index. The quantity $c_i(t)$ essentially represents the marginal increase in the total future viral load if we let one more $CD4^+$ cell get infected by virus strain $i$ at time $t$. The index for drug combination $j$ is computed by weighting the dynamic marginal cost $c_i(t)$ for each virus by the efficiency of drug $j$ for that virus, and summing up over the virus strains. Hence, our dynamic policy uses information on the effectiveness of each drug against each virus and the current potential cost (in terms of total future viral load) of a new infection by each virus strain, and summarizes this information in a succinct manner. An important advantage of index policies is their ease of use: the complexity of the derivation and implementation of an index policy is independent of the problem size; hence, the policy defined by eqns (31–33) can easily be derived for a problem with 20 drug combinations and 150 virus strains. Although this policy is not the optimal solution to the problem shown in eqns (3–6), the optimal solution to dynamic resource allocation problems is often characterized by index policies (e.g. Gittins, 1989). Finally, we note that inequality (32) was never satisfied in any of our numerical runs: the proposed dynamic policy always used one of the drug combinations.

It is possible to refine our optimality procedure by implementing further iterations of the policy improvement algorithm. Let us denote the proposed policy in eqns (31–33) by $d_j^{(1)}(t)$; although this control is state-dependent, it can be expressed solely as a function of time because the system is deterministic. If we define $\bar{d}_i^{(1)}(t) = 1 - \Sigma_{j=1}^{J} p_{ji}d_j^{(1)}(t)$, then this quantity can be used as our starting policy for the next policy improvement iteration. Turning to the asymptotic analysis, we observe that the $\epsilon^0$-order system in eqns (18–21) and its solution in eqns (22–24) are independent of the control. If we substitute $\bar{d}_i^{(1)}(t)$ in for $\bar{d}_i$ in eqns (25) and (26), then the $\epsilon$-order system in eqns (25–28) is still a set of linear differential equations that can be easily solved numerically using the matrix exponential (Golub & Van Loan, 1989). Then we can carry out the calculations shown in Appendix B and eqns (29) and (30) on a computer, and perform the minimization in eqn (7) to get a new policy $d_j^{(2)}(t)$. Of course, higher order terms in the asymptotic expansion can also be performed relatively easily with a computer, and it is conceivable that, with enough higher order terms and enough policy iterations (typically, only a handful of the latter is required to get close to optimality), such a procedure would generate a policy that is very close to optimal. However, because the proposed policy in eqns (31–33) performed well in our numerical study, we have not pursued this computational approach.

In fact, rather than focus on developing more complex refinements of eqns (31–33), we believe that it is more fruitful to seek a *simpler* policy that still performs well. A cruder but more transparent policy can be obtained by only considering the first term of the asymptotic expansion; i.e. by setting $\epsilon = 0$ in eqn (31). Further simplifications to the policy can be achieved by making the reasonable assumptions that the death rates of infected cells and free virus are independent of the viral strains (i.e. $\alpha_i = \alpha$, $k_i = k$). A dramatically simplified policy is obtained by observing that $q_{ii} \gg q_{ij}$: at time $t$, use the drug option $j$ that has the largest value of:

$$\sum_{i=1}^{I} p_{ji}\tilde{\beta}_i\pi_i v_i(t). \tag{34}$$

This policy is very easy to understand, yet still maintains the essential features of a robust dynamic policy: it focuses relatively more attention on strains that have higher infectivity and/or replication rates,

and it redirects its effort towards the virus strain that is most prevalent at each point in time.

## 4. Alternative Therapies

The tedious part of the analysis in Section 3 is the perturbation analysis that leads to the derivatives of the value function for a generic static policy. Now that these derivatives have been estimated, it is a relatively simple matter to consider other types of therapies. Here are two examples; much of the previous notation is reused.

### 4.1. PROTEASE INHIBITORS

Protease inhibitors render newly produced virions non-infectious. Suppose we have $J$ combinations of protease inhibitors, and the controller must decide which, if any, drug combination to use at each point in time. The drugs' effectiveness is given by the matrix $p_{ji}$, and the resulting virus replication rate is $\pi_i[1 - \Sigma_{j=1}^{J} p_{ji} d_j(t)]$. An analysis similar to that in Section 3 yields the dynamic index:

$$\sum_{i=1}^{I} p_{ji} \frac{\partial V_\epsilon}{\partial v_i} \pi_i y_i(t), \tag{35}$$

for drug combination $j$. At each point in time, the proposed therapeutic strategy uses the combination with the largest index if this index is positive; otherwise, no drugs are administered. Note that $\partial V_\epsilon / \partial v_i$ should be positive, so therapy should always be applied. Define $\bar{d}_i = 1 - \Sigma_{j=1}^{J} p_{ji} d_j^s$, where $d_j^s$ solves the static optimization problem that is analogous to the one in Section 3.2. Integrating the approximate value function in Appendix B with respect to $v_i$ gives:

$$\frac{\partial V_\epsilon}{\partial v_i} = \frac{1}{k_i} = \frac{\epsilon \beta_i}{k_i(k_i + \mu)} \left( \sum_{j=1}^{I} \frac{q_{ij} \bar{d}_j \pi_j}{\alpha_j k_j} - \frac{1}{k_i} \right) (k_i x - \lambda), \tag{36}$$

and substitution of this quantity into eqn (35) yields the proposed therapy in terms of the problem parameters and the current state. The simplified policy that is analogous to eqn (34) employs the protease combination that has the largest value of:

$$\sum_{i=1}^{I} p_{ji} \pi_i y_i(t). \tag{37}$$

### 4.2. RECONSTITUTING THE IMMUNE SYSTEM

Consider using a drug, such as interleukin-2 (IL-2), that reconstitutes the immune system. This drug affects our model by increasing $\lambda$, the production rate of uninfected CD4$^+$ cells. Suppose the new production rate is $\lambda + \tilde{\lambda}(t)$, where our control $\tilde{\lambda}(t) \in [0, \bar{\lambda}]$. Then the optimization problem embedded in the dynamic programming optimality equation is simply to minimize $\lambda \partial V / \partial x$. The proposed policy is:

$$\tilde{\lambda}^*(t) = \bar{\lambda} \quad \text{if} \quad \frac{\partial V_\epsilon}{\partial x} < 0, \tag{38}$$

and $\tilde{\lambda}^*(t) = 0$ otherwise, where $\partial V_\epsilon / \partial x$ is given in eqn (29) (with $\bar{d}_i = 1$). Define the constants:

$$c_i = \sum_{j=1}^{I} \frac{q_{ij} \pi_j}{\alpha_j k_j} - \frac{1}{k_i}. \tag{39}$$

If $\tilde{c}_i > 0$ for all $i$ then we never use the drug, and if $\tilde{c}_i \leqslant 0$ for all $i$ then we always use the drug; if neither of these cases hold then a dynamic policy is optimal. The first quantity on the right side of eqn (39) is the expected (with respect to the mutation probabilities $q_{ij}$) value of $\pi/(\alpha k)$. Since the mutation rates are very small, this quantity is nearly equal to $\pi_i/(\alpha_i k_i)$, and if we ignore mutations then the drug is always given if and only if $\pi_i > \alpha_i$ for all $i$. If $\pi_i > \alpha_i$, then the "basic reproductive ratio $R_0$" for the infected cells is greater than unity (that is, each infected cell will produce more than one free virus particle during its lifetime), and so this drug effectively "adds more fuel to the fire". Empirical results (see Section 5) suggest that $\pi_i > \alpha_i$; hence, adding uninfected CD4$^+$ cells in isolation is not desirable; of course, our model has not incorporated an immune response, and thus may be omitting some positive side effects of additional CD4$^+$ cells. Although we do not do so here, a similar analysis can be performed for a therapy that *reduces* CD4$^+$ cell production.

Other therapies that can be analyzed include certain forms of immunotherapy (which would increase the death rate of free virus and/or the death rate of infected cells), *ex vivo* expansion of CD4$^+$ cells (Wilson *et al.*, 1995), which would simultaneously increase $x$ and decrease $y_i$, and dynamic gene therapy (Nabel, 1994), which would increase $v_i$ for certain strains.

Most importantly, we can also consider employing some of these therapies *simultaneously*. For example, if one allowed the simultaneous use of RT inhibitors and production of CD4$^+$ cells (Schwartz *et al.*, 1991), it may turn out to be beneficial to introduce CD4$^+$ cells at times when the infectivity of the virus and the viral load are sufficiently suppressed.

## 5. An Illustrative Example

In this section, the dynamic model is simulated under the proposed policy, as well as under several

simpler policies. No attempt has been made to generate a model of realistic size; rather, we consider only two virus strains and two drug combinations in order to illustrate the nature and the impact of dynamic drug treatment. In a subsequent study, we plan to use data from multidrug clinical trials to analyze larger and more realistic instances of the model.

### 5.1. PARAMETER VALUES

The parameter values for our model are displayed in Table 1. Most of these values were sequentially derived from existing data in the following manner. Ho *et al*. (1995) and Wei *et al*. (1995) estimate a $CD4^+$ production rate of roughly $1.8 \times 10^9$ cells per day. About 2% of the total $CD4^+$ population resides in the peripheral blood, and a human body contains roughly $5 \times 10^6$ mm$^3$. Hence, the $1.8 \times 10^9$ figure is comparable to 7 cells mm$^{-3}$ day$^{-1}$. The death rate $\mu$ was chosen to be 0.007 day$^{-1}$, so that the virus-free equilibrium $CD4^+$ count is $\lambda/\mu = 1000$ cells mm$^{-3}$, which corresponds to the $CD4^+$ count for an uninfected individual.

Now we turn to the death rates $\alpha_i$ of infected cells and $k_i$ of free virus. Ho *et al*. (1995) and Wei *et al*. (1995) estimate the death rate of virus-producing cells to be about 0.35 day$^{-1}$. More recently, using more accurate data, Perelson *et al*. (1996) show that the mean death rate is about 0.49. They were also able to get a rough estimate of 3.07 for the death rate of free virus.

Some of the remaining parameters are derived by considering the quasi-steady state conditions before drug treatment. For typical pre-treatment values, we use the average values (over 20 individuals) in Table 1 of Ho *et al*.; the average pre-treatment $CD4^+$ count was 180 cells mm$^{-3}$ and the average viral load was 134 virions mm$^{-3}$. Only the wild-type virus was present in non-negligible amounts before treatment. Hence, we can consider only the wild-type virus, and set the left side of eqns (3–5) equal to zero (reflecting the quasi-steady state) to obtain a set of four equations [eqns (3–5) and $x + y = 180$] and four unknowns: the pre-treatment number of uninfected cells $x$ and infected cells $y$, the infectivity rate $\tilde{\beta}$ and the replication rate $\pi$. Substituting $180 - y$ for $x$ in eqns (3) and (4) and solving these two equations for $\tilde{\beta}$ and $y$ yields $\tilde{\beta} = (\alpha\lambda - 180\alpha\mu)/(180\alpha v - \lambda v) = 2.58 \times 10^{-4}$ mm$^3$ day$^{-1}$ and $y = 180\tilde{\beta}v/(\alpha + \tilde{\beta}v) = 11.86$ cells mm$^{-3}$. Hence, $x = 168.14$ cells mm$^{-3}$, and the fraction of cells that are infected is $y/(x + y) = 0.066$, which is in close agreement with the estimate of 5% found in Embretson *et al*. (1993). Finally, solving eqn (5) for $\pi$ yields $\pi = v(k + \tilde{\beta}x)/y = 35.18$ virions day$^{-1}$, implying that $\pi/\alpha = 71.8$ virions are produced by an infected cell in our model. Readers should keep in mind that most virus is produced in the lymph system, whereas our estimates for $\pi$ and $\tilde{\beta}$ are based on plasma concentrations.

We use the mutation rate calculated in Mansky & Temin (1995), $q_{12} = q_{21} = 3.4 \times 10^{-5}$. We also assume that drug combination $i$ is targeted at virus strain $i$. More specifically, we let $p_{11} = p_{22} = 0.95$ and $p_{12} = p_{21} = 0.05$, meaning that each drug combination is 95% effective at blocking infections of its own strain, but only 5% effective at blocking infections of the other strain; such a state of affairs might arise if virus 1 is an AZT-resistant strain that is dominant at time zero, and the two drug combinations correspond to two other (i.e. not AZT) RT inhibitors.

Notice that, until now, the parameter values are consistent with the symmetric case introduced

TABLE 1
*Parameter values for the model*

| Variables | | Initial values |
|---|---|---|
| $x$ | Density of uninfected $CD4^+$ cells | 168.08 cells mm $-$ $^3$ |
| $y_1$ | Density of cells infected by strain 1 | 11.88 cells mm$^{-3}$ |
| $y_2$ | Density of cells infected by strain 2 | 0.004 cells mm$^{-3}$ |
| $v_1$ | Density of free virus of strain 1 | 134.25 virions mm$^{-3}$ |
| $v_2$ | Density of free virus of strain 2 | 0.04 virions mm$^{-3}$ |
| Parameters | | Values |
| $\lambda$ | Production rate of uninfected $CD4^+$ cells | 7.0 cells mm$^{-3}$ day$^{-1}$ |
| $\mu$ | Death rate of uninfected $CD4^+$ cells | 0.007 day$^{-1}$ |
| $\alpha_1$, $\alpha_2$ | Death rates of infected $CD4^+$ cells | 0.49 day$^{-1}$ |
| $k_1$, $k_2$ | Death rates of free virus | 3.07 day$^{-1}$ |
| $\tilde{\beta}_1$, $\tilde{\beta}_2$ | Infectivity rates of free virus | $2.58 \times 10^{-4}$ mm$^{-3}$ day$^{-1}$ |
| $\pi_1$ | Replication rate of strain 1 | 35.18 virions day$^1$ |
| $\pi_2$ | Replication rate of strain 2 | 31.66 virions day$^{-1}$ |
| $q_{12}$, $q_{21}$ | Mutation rates (probabilities) | $3.4 \times 10^{-5}$ |
| $p_{11}$, $p_{22}$ | Drug efficacies on their own strain | 0.95 |
| $p_{12}$, $p_{21}$ | Drug efficacies on the other strain | 0.05 |

in Section 3.5. Now we introduce asymmetry by letting $\pi_2 = 0.9\pi_1$; hence, we assume that virus 1 has a higher replication rate than virus 2. The pre-treatment equilibrium state was taken as the starting point of our simulation runs (see Table 1).

### 5.2. SIMULATION RESULTS

For the set of parameter values specified in Subsection 5.1, we computed the proposed dynamic multidrug therapy given by eqns (31–33), and then determined the system behavior under this regimen by simulating eqns (3–5) up to time $T = 6$ months (the system had nearly reached its steady state by this time). The simulation results were generated by the automatic step size Runge–Kutta–Fehlberg method, with a fourth and fifth order pair for higher accuracy. Our results are shown in Fig. 1(a–e). The horizontal axis in all five graphs is time, and the uninfected $CD4^+$ count $x(t)$ is given in Fig. 1(a), the logarithms

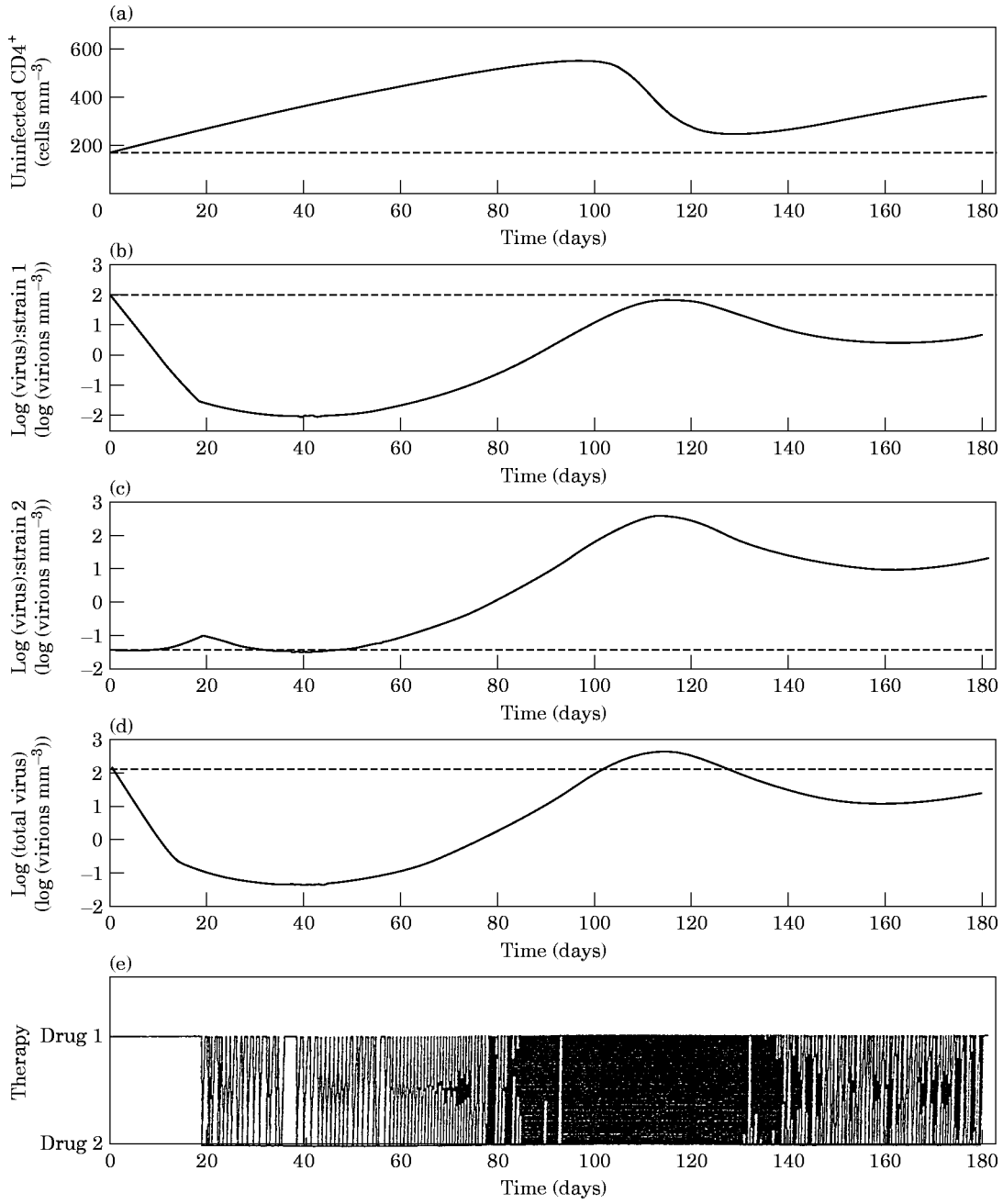


FIG. 1. System behavior under the dynamic policy.

of the viral loads $v_i(t)$ are pictured in Figs 1(b) and 1(c), the logarithm of the total viral load is shown in Fig. 1(d) and the proposed dynamic therapy is displayed in Fig. 1(e). Figure 2 displays the same quantities under the *continuous treatment* policy that continually applies drug 1 [i.e. $d_1(t) = 1$, $d_2(t) = 0$ for all $t$]; this policy was also the optimal static policy [i.e. it was the solution to eqns (8–14)]. Under both policies, the shape of the $v_i$ and $y_i$ curves were very similar, with the free virus $v_i$ lagging behind the infected cell count $y_i$ by several days; hence, the dynamics of the infected cells do not appear in Figs 1 and 2. We also simulated the simplified dynamic policy in eqn (34). Although we do not include figures for this policy, the average (over 1 year) uninfected $CD4^+$ count and average total viral load for the various policies are reported in Table 2.

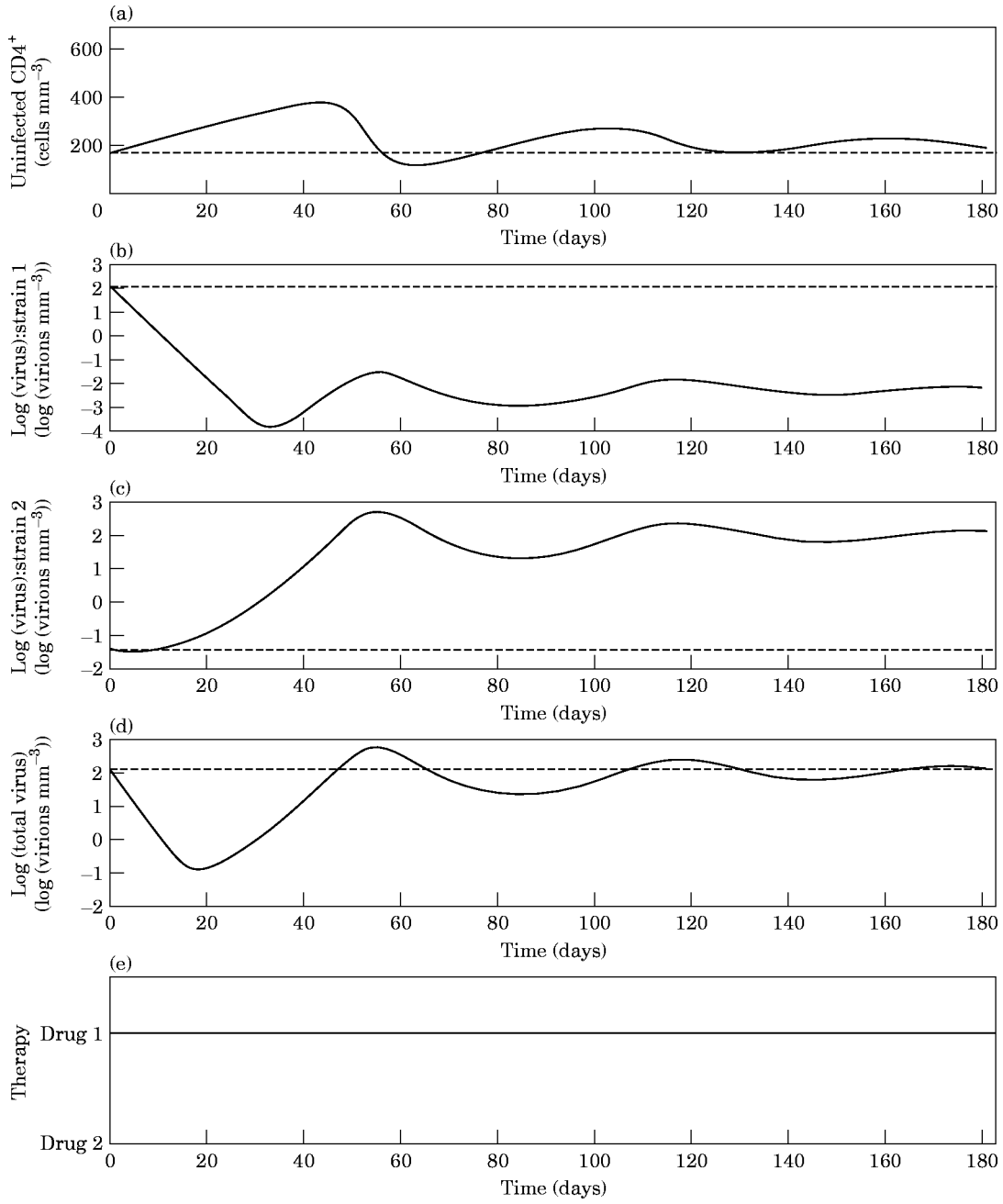


FIG. 2. System behavior under the continuous treatment policy.

TABLE 2
*Summary of numerical results*

| Policy | Average total viral load (virions mm$^{-3}$) | Average uninfected CD4$^+$ count (cells mm$^{-3}$) |
|---|---|---|
| No drugs | 134.29 | 168.08 |
| Continuous treatment | 110.99 | 221.11 |
| Simplified dynamic | 63.26 | 379.37 |
| Dynamic | 61.72 | 379.66 |

## 6. Discussion

It should be stressed that we are considering an illustrative example: the numerical values reported here are merely the result of a particular computer simulation, with a particular choice of parameters.

### 6.1. DYNAMIC POLICY

At time zero, the system is in its drug-free equilibrium, and virus strain 1 is dominant. As expected, the viral load in Fig. 1 drops dramatically throughout the first 20 days, and does not emerge in a significant way until the beginning of the fourth month. There is a concomitant increase, which is roughly linear, in the number of uninfected CD4$^+$ cells, and this quantity peaks at 3 months; the peak corresponds to more than a three-fold increase over the pre-treatment value. Not surprisingly, the dynamic policy initiates treatment with drug 1. The policy first uses drug 2 on day 18, and switches back and forth between the two drugs on average once per day until the end of the third month. During the third and fourth months, the two virus strains feed on the large pool of uninfected cells and simultaneously emerge, peaking with a total viral load that is nearly four times the pre-treatment level. The drug treatment oscillates rapidly between the two options during the fourth month, in an attempt to simultaneously control both viruses. The less fit virus 2 constitutes about 75% of the free virus during the fourth and fifth months, and hence the dynamic policy exerts the majority of its effort in controlling the more fit virus 1.

The large viral load in turn leads to more cell infections, and the uninfected CD4$^+$ count decreases in the fourth month, bottoming out at a level that is higher than the pre-treatment uninfected CD4$^+$ count. This reduction in uninfected cells allows the drugs to bring both viruses back under control during the fifth month. After this time, these oscillations dampen (e.g. the total viral load peaks again at the end of the seventh month with a value of about 210) and a new steady state is slowly reached. Over the first 6 months, the dynamic policy allocated 59.9% of the time to drug 1 and 40.1% of the time to drug 2; hence, the dynamic policy expended more effort than the optimal static policy on the more fit virus, even after accounting for the fact that the policy spends its first 18 days focused on virus strain 1.

### 6.2. CONTINUOUS TREATMENT

Under the continuous treatment policy pictured in Fig. 2, we again see a rapid drop in virus l and a linear rise in the CD4$^+$ count; however, in contrast to the dynamic policy, the continuous treatment policy suppresses virus 1 throughout the first 6 months. Virus 2 emerges during the second month, and reaches a very high level by the end of the month. This high viral load then drops the uninfected CD4$^+$ count below its pre-treatment level, which in turn leads to a reduction of free virus 2 in the third month.

### 6.3. COMPARISON OF POLICIES

Table 2 shows that the dynamic policy performs much better than the continuous treatment policy: the continuous policy achieves a 17.4% reduction in average (over the first 6 months) viral load and a 31.6% increase in uninfected CD4$^+$ count with respect to the drug-free equilibrium, whereas the dynamic policy achieves a 54.0% reduction in viral load and 125.8% increase in uninfected CD4$^+$ count. Moreover, the viral peaks and CD4$^+$ valleys are less pronounced under the dynamic policy than under the continuous treatment policy. Finally, the dynamic policy, by frequently switching between the two drug options during the early months, delays the emergence of virus 2 from the second month until the fourth month. Although the results are not reported here, we also tested an *alternating* policy that uses drug 1 for the first 3 months and then uses drug 2 for the last 3 months. Not surprisingly, this policy did not perform well.

Perhaps the most intriguing observation from our numerical study is that the simplified dynamic policy defined in eqn (34) performed nearly as well as the more complex dynamic policy. The $o(\epsilon)$ term in eqn (31) had little impact on the value of the dynamic index policy in our numerical examples, suggesting that our simplified policy shows promise as a practically implementable procedure.

We should note that the post-treatment steady state values of uninfected CD4$^+$ cells and total virus load under all of the therapeutic strategies are not appreciably different than the pre-treatment steady state values; hence, the therapeutic benefits in our model are achieved during the transient domain. However, the viral dynamics of HIV-infected individuals undergoing dynamic drug treatment would probably exhibit transient behavior for many

years. Therefore, the improvements over the transient domain in our model are more indicative of the improvements that can be realized in clinical practice.

### 6.4. UNEQUAL INFECTIVITY RATES

Because there seems to be some uncertainty about whether virus strains have different infectivity rates, we reduced the infectivity of virus 2 so that $\tilde{\beta}_2 = 0.9\tilde{\beta}_1$. The qualitative results were similar to our base case, although the dynamic policy expended a slightly larger fraction of its effort on virus 1, resulting in a somewhat larger proportion of virus 2 in the viral mix. More generally, numerous other simulation runs were generated by varying the parameters $\pi_i$, $\tilde{\beta}_i$, $\lambda$ and $p_{ij}$, and the qualitative results displayed in Figs 1 and 2 remained intact; hence, the model appears to be robust.

## 7. Conclusions

We have used the control theory paradigm in a HIV therapeutic setting. Our model incorporates different virus strains, and a variety of therapeutic options are available. The approximation method, which uses perturbation analysis and the policy improvement algorithm, gives rise to a *dynamic index policy*: each drug combination has an associated dynamic index, and at each point in time the policy uses the drug combination with the largest index. The dynamic indices succinctly summarize the efficacy of each drug combination on each virus strain and the marginal benefit of blocking a new cell infection by each virus strain; the latter quantity changes over time as a function of an individual's $CD4^+$ count, viral load and viral mix.

Numerical results for a two-virus, two-drug model suggest that dynamic multidrug therapies outperform their static counterparts: the total viral load is reduced, the uninfected $CD4^+$ count is increased, and the emergence of drug resistant strains is delayed. Although the *individualized therapy* that we propose is no doubt more difficult to implement than protocols that are currently in practice, the benefits may outweigh the costs of implementation. These benefits are achieved by frequently changing therapies over time in *response to* and *in anticipation of* the emergence of drug-resistant strains. In addition, the dynamic policy attempts to maintain the more fit virus strains at a relatively low level, while perhaps allowing the less fit strains to partially establish themselves. From a practical perspective, the most important observation from our numerical results was that nearly all of the benefits from a dynamic policy were achieved by a simplified policy [see eqns (34) and (37)] that has considerable intuitive appeal and requires knowledge about only a small portion of the problem parameters and system state.

Although our numerical results focus on the two-virus, two-drug case, the development of more realistic models using data from multidrug studies is planned for the future. It may turn out that the best way to delay the onset of AIDS is *via* the intelligent use of a wide range of therapies (RT inhibitors, protease inhibitors, reconstitution of the immune system, immunotherapy, gene therapy, etc). The model and analysis presented here provides the framework for the development of such therapeutic strategies.

Finally, our approach to this problem circumvents the usual obstacles inherent in analysing high-dimensional nonlinear control problems. This method, which appears to be new, has potential applications in a wide variety of control problems in epidemiology and ecology; the key feature of these problems is the existence of small "infectivity" parameters that allow for the explicit asymptotic analysis of a set of nonlinear differential equations. Besides allowing for mutation among multiple variants of entities (in this case, viruses), the approach can also incorporate discrete age classes (e.g. for optimal vaccinations of measles) and discrete spatial (e.g. lattice) structures (e.g. for dynamic control of spatial epidemics).

LMW was supported by NSF grant DDM-9057297 and EPSRC grant GR/J71786. He gratefully acknowledges a valuable discussion with Peter Whittle during the early course of this research, and thanks Denise Kirschner for a helpful conversation about parameter estimation. MAN was supported by the Wellcome Trust and Keble College. We also thank Rebecca D'Amato for assisting in the numerical computations.

## REFERENCES

AGUR, Z. (1989). A new method for reducing cytotoxicity of the anti-AIDS drug AZT. In: *Biomedical Modeling and Simulation* (Levine, D. S., ed.) pp. 59–61. J. C. Baltzer AG, Scientific Publishing Co. IMACS.

BELLMAN, R. E. (1957). *Dynamic Programming*. Princeton NJ: Princeton University Press.

CARRIER, G. F. & PEARSON, C. E. (1991). *Partial Differential Equatioins*. San Diego, CA: Academic Press.

CONDRA, J. H. SCHLEIF, W. A., BLAHY O. M., GABRYELSKI, L. J., GRAHAM, D. J., QUINTERO A. R. *et al*. (1995). *In vivo* emergence of HIV-1 variants resistant to multiple protease inhibitors. *Nature* **374,** 569–571.

DE JONG, M. D., VEENSTRA, J., STILANAKIS, N. I., SCHUURMAN R., LANGE, J. M. A., DE BOER, R. J. *et al*. (1996). Host–parasite dynamics and outgrowth of virus containing a single K70R amino acid change in reverse transcriptase are responsible for the loss of HIV-1 RNA load suppression by zidovudine. *Proc. Natl Acad. Sci. U.S.A.* **93,** 5501–5506.

ELLIOTT, R. J., AGGOUN, L. & MOORE, J. B. (1994). *Hidden Markov Models: Estimation and Control*. New York: Springer-Verlag.

EMBRETSON, J., ZUPANCIC, M., BENEKE, J., TILL M., WOLINSKI, S., RIBAS, J. L. *et al*. (1993). Analysis of human immunodeficiency virus-infected tissues by amplification and *in situ* hybridization reveals latent and permissive infections at single-cell resolution. *Proc. Natl. Acad. Sci. U.S.A.* **90,** 357–361.

ERON, J. J., BENOIT, S. L., MACARTHUR, R. D., SANATANA, J., QUINN, J. B., KURITSKES D. R. *et al*. (1995). Treatment with lamivudine, zidovudine, or both in HIV-positive patients with 200 to 500 CD4$^+$ cells per cubic millimeter. *New Engl. J. Med.* **333,** 1662–1669.

FROST, S. D. W. & MCLEAN, A. R. (1994). Quasispecies dynamics and the emergence of drug resistance during zidovudine therapy of HIV infection. *AIDS* **8,** 323–332.

GITTINS, J. C. (1989). *Multi-armed Bandit Allocation Indices*. New York: Wiley.

GOLUB, G. H. & VAN LOAN, C. F. (1989). *Matrix Computations*. Baltimore: Johns Hopkins University Press.

HO, D. D., NEUMANN, A. U., PERELSON, A. S., CHEN, W., LEONARD, J. M. & MARKOWITZ, M. *et al*. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373,** 123–127.

HOWARD, R. A. (1960). *Dynamic Programming and Markov Processes*. Cambridge, MA: MIT Press.

KEY, P. B. (1990). Optimal control and trunk reservation in loss networks. *Prob. Engrg. Inf. Sci.* **4,** 203–242.

KIRSCHNER, D. & PERELSON, A. (1994). A model for the immune system response to HIV: AZT treatment studies. In: *Mathematical Populations Dynamics III*, *Theory of Epidemics* (Arino, O., Axelrod, D. & Kimmel, M., eds.) vol. 1, pp. 296–301. Winnipeg Manitoba: Wuerz Publications.

KIRSCHNER, D. & WEBB, G. F. (1995). Effects of drug resistance in chemotherapy strategies in the treatment of HIV infection. *Bull. Math. Biol.* (in press).

KIRSCHNER, D. & WEBB, G. F. (1996). A model for treatment strategy in the chemotherapy of AIDS. *Bull. Math. Biol.* **58,** 367–390.

KIRSCHNER, D., LENHART, S. & SERBIN, S. (1995). Optimal control of the chemotherapy of HIV. *J. Math. Biol.* (in press).

KOJIMA, E., SHIRASAKA T., ANDERSON B. D., CHOKEKIJCHAI S., STEINBERG S. M., BRODER S. *et al*. (1995). Human immunodeficiency virus type 1 (HIV-1) viremia changes and development of drug-related mutations in patients with symptomatic HIV-1 infection receiving alternating or simultaneous zidovudine and didanosine therapy. *J. Infect. Dis.* **171,** 1152–1158.

LAGAKOS, S., PETTINELLI, C., STEIN, D. & VOLBERDING, P. A. (1993). The Concorde Trial. *Lancet* **341,** 1276.

LARDER, B. A., KELLAM, P. & S. D. KEMP. (1993). Convergent combination therapy can select viable multidrug-resistant HIV-1 *in vitro*. *Nature* **365,** 451–453.

LARDER, B. A., KEMP, S. D. & HARRIGAN, P. R. (1995). Potential mechanism for sustained antiretroviral efficacy of AZT-3TC combination therapy. *Science* **269,** 696–699.

MANSKY, L. M. & TEMIN, H. M. (1995). Lower *in vivo* mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J. Virol.* **69,** 5087–5094.

MCLEAN, A. R. & NOWAK, M. A. (1992). Competition between zidovudine sensitive and zidovudine resistant strains of HIV. *AIDS* **6,** 71–79.

MCLEOD, G. X. & HAMMER, S. M. (1992). Zidovudine: 5 years later. *Ann. Int. Med.* **117,** 487–501.

NABEL, G. J. (1994). Gene therapy approaches to AIDS. *AIDS* **8,** S61–S69.

NOWAK, M. A. & MAY, R. M. (1993). AIDS pathogenesis: mathematical models of HIV and SIV infections. *AIDS* **7,** S3–S18.

NOWAK, M. A., ANDERSON, R. M.,, MCLEAN A. R., WOLFS, T. F. W., GOUDSMIT, J. & MAY, R. M. (1991). Antigenic diversity thresholds and the development of AIDS. *Science* **254,** 963–969.

OTT, T. J. & KRISHNAN K. R. (1985). State dependent routing of telephone traffic and the use of separable routing schemes. In: *Proceedings of the* 11*th International Teletraffic Congress* (Akiyama, M., ed.). Amsterdam: Elsevier.

PANTALEO, G., COHEN, O. J., SCHWARTZENTRUBER D. J., GRAZIOSI C., VACCAREZZA M. & FAUCI A. S. (1995). Pathogenic insights from studies of lymphoid tissue from HIV-infected individuals. *J. AIDS Hum. Retrovir.* **10** (**Suppl. 1**), S6–S14.

PERELSON, A. S. (1989). Modeling the interaction of the immune system with HIV. In: *Mathematical and Statistical Approaches to AIDS Epidemiology* (Castillo-Chavez, C., ed.) Lecture Notes in Biomathematics, vol. **83**, pp. 350–370.

PERELSON, A. S., KIRSCHNER D. & DEBOER, R. (1993). The dynamics of HIV infection of CD4$^+$ cells. *Math. Biosci.* **114,** 81–125.

PERELSON, A. S., NEUMANN, A. U., MARKOWITZ, M., LEONARD, J. M. & HO, D. D. (1996). HIV-1 dynamics in vivo: Virion clearance rate, infected cell lifespan, and viral generation time. *Science* **271,** 1582–1586.

PONTRYAGIN, L. S., BOLTYANSKII, V. G., GAMKRELIDZE, R. V. & MISHCHENKO, E. F. (1962). *The Mathematical Theory of Optimal Progesses*. New York: Interscience Publishers.

RICHMAN, D. D., MENG, T., SPECTOR, S. A., FISCHL, M. A., RESNICK, L. & LAI, S. (1994). Resistance to AZT and ddC during long-term combination therapy in patients with advanced infection with human immunodeficiency virus. *J. AIDS* **7,** 135–138.

ST. CLAIR, M. H., MARTIN, J. L., TUDOR-WILLIAMS, G., BACH, M. C., VAVRO, C. L., KING, D. M. *et al*. (1991). Resistance to ddI and sensitivity to AZT induced by a mutation in HIV-1 reverse transcriptase. *Science* **253,** 1557–1559.

SCHUURMAN, R., NIJHUIS, M., VAN LEEUWEN R., SCHIPPER P., DE JONG, D., COLLIS, P., *et al*. (1995). Rapid changes in human immunodeficiency virus type 1 RNA load and appearance of drug-resistant virus populations in persons treated with lamivudine (3TC). *J. Infect. Dis.* **171,** 1411–1419.

SCHWARTZ, D. H., SKOWRON, G. & MERIGAN, T. C. (1991). Safety and effects of interleukin-2 plus zidovudine in asymptomatic individuals infected with human immunodeficiency virus. *J. AIDS* **4,** 11–23.

SHAFER, R. W., KOZAL, M. J., WINTERS, M. A., IVERSON, A. K. N., KATZENSTEIN, D. A., RAGNI M. V. *et al*. (1994). Combination therapy with zidovudine and didanosine selects for drug-resistant human immunodeficiency virus type 1 strains with unique patterns of *pol* gene mutations. *J. Infect. Dis.* **169,** 722–729.

SHAFER, R. W., IVERSON, A. K. N., WINTERS, M. A., AGUINIGA, E., KATZENSTEIN, D. A., MERIGAN, T. C. *et al*. (1995). Drug resistance and heterogeneous long-term virologic responses of human immunodeficiency virus type 1-infected subjects to zidovudine and didanosine combination therapy. *J. Infect. Dis.* **172,** 70–78.

VOLBERDING, P. A., LAGAKOS, S. W., GRIMES, J. M., STEIN, D. S., BALFOUR, J. J. REICHMAN, R. C. *et al*. (1994). The duration of zidovudine benefit in persons with asymptomatic HIV infection. *JAMA* **272,** 437–442.

WEI, X., GHOSH, S. K., TAYLOR, M. E. JOHNSON, V. A., EMINI, E. A., DEUTSCH P. *et al*. (1995). Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373,** 117–123.

WILSON, C. C. & HIRSCH, M. S. (1995). Combination antiretroviral therapy for the treatment of human immunodeficiency virus type-1 infection. *Proc. Ass. Am. Phys.* **107,** 19–27.

WILSON, C. C., WONG, J. T., GIRARD, D. D. MERRILL, D. P., DYNAN, M., AN D. D. *et al*. (1995). Ex vivo expansion of CD4 lymphocytes from human immunodeficiency virus type 1-infected persons in the presence of combination antiretroviral therapy. *J. Infect. Dis.* **172,** 88–96.

## APPENDIX A

The solution to eqns (25–28) is:

$$x^{(1)}(s) = C_x e^{-\mu(s-t)} - \sum_{i=1}^{I} \times \left( \beta_i \bar{d}_i \frac{\lambda}{\mu} \left( \frac{\pi_i y_i e^{-\alpha_i(s-t)}}{(k_i - \alpha_i)(\mu - \alpha_i)} + \left( v_i - \frac{\pi_i y_i}{k_i - \alpha_i} \right) \right. \right.$$

$$\times\frac{e^{-k_i(s-t)}}{\mu-k_i}\Bigg)+\beta_i\bar{d}_i\Bigg(x-\frac{\lambda}{\mu}\Bigg)\Bigg(\frac{\pi_i y_i e^{-(\alpha_i+\mu)(s-t)}}{\alpha_i(k_i-\alpha_i)}$$

$$+\Bigg(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\Bigg)\frac{e^{-(k_i+\mu)(s-t)}}{k_i}\Bigg)\Bigg), \tag{A.1}$$

$$y_i^{(1)}(s)=C_{y_i}e^{-\alpha_i(s-t)}+q_{ii}\beta_i\bar{d}_i\frac{\lambda}{\mu}\Bigg(\frac{\pi_i y_i s e^{-\alpha_i(s-t)}}{k_i-\alpha_i}$$

$$+\Bigg(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\Bigg)\frac{e^{-k_i(s-t)}}{\alpha_i-k_i}\Bigg)-q_{ii}\beta_i\bar{d}_i$$

$$\times\Bigg(x-\frac{\lambda}{\mu}\Bigg)\Bigg(\frac{\pi_i y_i e^{-(\alpha_i+\mu)(s-t)}}{\mu(k_i-\alpha_i)}-\Bigg(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\Bigg)$$

$$\times\frac{e^{-(k_i+\mu)(s-t)}}{\alpha_i-k_i-\mu}\Bigg)+\sum_{j\neq i}^{I}\Bigg(q_{ji}\beta_j\bar{d}_j\frac{\lambda}{\mu}\Bigg(\frac{\pi_j y_j e^{-\alpha_j(s-t)}}{(k_j-\alpha_j)(\alpha_i-\alpha_j)}$$

$$+\Bigg(v_j-\frac{\pi_j y_j}{k_j-\alpha_j}\Bigg)\frac{e^{-k_j(s-t)}}{\alpha_i-k_j}\Bigg)+q_{ji}\beta_j\bar{d}_j\Bigg(x-\frac{\lambda}{\mu}\Bigg)$$

$$\times\Bigg(\frac{\pi_j y_j e^{-(\alpha_j+\mu)(s-t)}}{(k_j-\alpha_j)(\alpha_i-\alpha_j-\mu)}$$

$$+\Bigg(v_j-\frac{\pi_j y_j}{k_j-\alpha_j}\Bigg)\frac{e^{-(k_j+\mu)(s-t)}}{\alpha_i-k_j-\mu}\Bigg)\Bigg), \tag{A.2}$$

$$v_i^{(1)}(s)=\frac{C_{y_i}\pi_i e^{-\alpha_i(s-t)}}{k_i-\alpha_i}+C_{v_i}e^{-k_i(s-t)}-\beta_i\frac{\lambda}{\mu}$$

$$\times\Bigg(\frac{\pi_i y_i e^{-\alpha_i(s-t)}}{(k_i-\alpha_i)^2}+\Bigg(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\Bigg)s e^{-k_i(s-t)}\Bigg)$$

$$-\beta_i\Bigg(x-\frac{\lambda}{\mu}\Bigg)\Bigg(\frac{\pi_i y_i e^{-(\alpha_i+\mu)(s-t)}}{(k_i-\alpha_i)(k_i-\alpha_i-\mu)}$$

$$-\Bigg(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\Bigg)\frac{e^{-(k_i+\mu)(s-t)}}{\mu}\Bigg)$$

$$+\pi_i q_{ii}\beta_i\bar{d}_i\frac{\lambda}{\mu}\Bigg(\frac{\pi_i y_i e^{-\alpha_i(s-t)}}{(k_i-\alpha_i)^2}+\Bigg(s-\frac{1}{k_i-\alpha_i}\Bigg)$$

$$+\Bigg(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\Bigg)\frac{s e^{-k_i(s-t)}}{\alpha_i-k_i}\Bigg)$$

$$-\pi_i q_{ii}\beta_i\bar{d}_i\Bigg(x-\frac{\lambda}{\mu}\Bigg)\Bigg(\frac{\pi_i y_i e^{-(\alpha_i+\mu)(s-t)}}{\mu(k_i-\alpha_i)(k_i-\alpha_i-\mu)}$$

$$+\Bigg(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\Bigg)\frac{e^{-(k_i+\mu)(s-t)}}{\mu(\alpha_i-k_i-\mu}\Bigg)$$

$$+\sum_{j\neq 1}^{I}\Bigg(\pi_i q_{ji}\beta_j\bar{d}_j\frac{\lambda}{\mu}\Bigg(\frac{\pi_j y_j e^{-\alpha_j(s-t)}}{(k_j-\alpha_j)(\alpha_i-\alpha_j)(k_i-\alpha_j)}$$

$$+\Bigg(v_j-\frac{\pi_j y_j}{k_j-\alpha_j}\Bigg)\frac{e^{-k_j(s-t)}}{(\alpha_i-k_j)(k_i-k_j)}\Bigg)$$

$$+\pi_i q_{ji}\beta_j\bar{d}_j\Bigg(x-\frac{\lambda}{\mu}\Bigg)$$

$$\Bigg(\frac{\pi_j y_j e^{-(\alpha_j+\mu)(s-t)}}{(k_j-\alpha_j)(\alpha_i-\alpha_j-\mu)(k_i-\alpha_j-\mu)}$$

$$+\Bigg(v_j-\frac{\pi_j y_j}{k_j-\alpha_j}\Bigg)\frac{e^{-(k_j+\mu)(s-t)}}{(\alpha_i-k_j-\mu)(k_i-k_j-\mu}\Bigg)\Bigg), \tag{A.3}$$

where the constants appearing in these equations are

$$C_x=\sum_{i=1}^{I}\Bigg(\frac{\beta_i\bar{d}_i\lambda}{\mu(\mu-k_i)}\Bigg(v_i-\frac{\pi_i y_i}{\mu-\alpha_i}\Bigg)$$

$$-\frac{\beta_i\bar{d}_i}{k_i}\Bigg(x-\frac{\lambda}{\mu}\Bigg)\Bigg(\frac{\pi_i y_i}{\alpha_i}+v_i\Bigg)\Bigg), \tag{A.4}$$

$$C_{y_i}=-q_{ii}\beta_i\bar{d}_i\frac{\lambda}{\mu}\Bigg(\frac{\pi_i y_i t}{k_i-\alpha_i}+\Bigg(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\Bigg)\frac{1}{\alpha_i-k_i}\Bigg)$$

$$+q_{ii}\beta_i\bar{d}_i\Bigg(x-\frac{\lambda}{\mu}\Bigg)\Bigg(\frac{\pi_i y_i}{\mu(k_i-\alpha_i)}-\Bigg(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\Bigg)$$

$$\times\frac{1}{\alpha_i-k_i-\mu)}\Bigg)+\sum_{j\neq i}^{I}\Bigg(q_{ji}\beta_j\bar{d}_j\frac{\lambda}{\mu}\Bigg(\frac{\pi_j y_j}{(\alpha_j-k_j)(\alpha_i-\alpha_j)}$$

$$-\Bigg(v_j-\frac{\pi_j y_j}{k_j-\alpha_j}\Bigg)\frac{1}{(\alpha_i-k_j)}\Bigg)+q_{ji}\beta_j\bar{d}_j\Bigg(x-\frac{\lambda}{\mu}\Bigg)$$

$$\times\Bigg(\frac{\pi_j y_j}{(k_j-\alpha_j)(\alpha_j+\mu-\alpha_i)}$$

$$-\Bigg(v_j-\frac{\pi_j y_j}{k_j-\alpha_j}\Bigg)\frac{1}{\alpha_i-k_j-\mu)}\Bigg)\Bigg), \tag{A.5}$$

$$C_{v_i} = -\frac{C_{y_i}\pi_i}{k_i-\alpha_i} + \beta_i\frac{\lambda}{\mu}\left(\frac{\pi_i y_i}{(k_i-\alpha_i)^2}\right) + \left(v_i - \frac{\pi_i y_i}{k_i-\alpha_i}\right)t\Bigg)$$

$$+\beta_i\left(x-\frac{\lambda}{\mu}\right)\left(\frac{\pi_i y_i}{(k_i-\alpha_i)(k_i-\alpha_i-\mu)}\right.$$

$$\left.-\left(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\right)\frac{1}{\mu}\right) - \pi_i q_{ii}\beta_i\bar{d}_i\frac{\lambda}{\mu}\left(\frac{\pi_i y_i}{(k_i-\alpha_i)^2}\right.$$

$$\left.\times\left(t-\frac{1}{k_i-\alpha_i}\right)+\left(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\right)\frac{1}{\alpha_i-k_i}\right)$$

$$+\pi_i q_{ii}\beta_i\bar{d}_i\left(x-\frac{\lambda}{\mu}\right)\left(\frac{\pi_i y_i}{\mu(k_i-\alpha_i)(k_i-\alpha_i-\mu)}\right.$$

$$\left.+\left(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\right)\frac{1}{\mu(\alpha_i-k_i-\mu)}\right)$$

$$+\sum_{j\neq i}^{I}\left(\pi_i q_{ji}\beta_j\bar{d}_j\frac{\lambda}{\mu}\left(-\frac{\pi_j y_j}{(k_j-\alpha_j)(\alpha_i-\alpha_j)(k_i-\alpha_j)}\right.\right.$$

$$\left.-\left(v_j-\frac{\pi_j y_j}{k_j-\alpha_j}\right)\frac{1}{(\alpha_i-k_j)(k_i-k_j)}\right)$$

$$-\pi_i q_{ji}\beta_j\bar{d}_j\left(x-\frac{\lambda}{\mu}\right)$$

$$\times\left(\frac{\pi_j y_j}{(k_j-\alpha_j)(\alpha_i-\alpha_j-\mu)(k_i-\alpha_j-\mu)}\right.$$

$$\left.\left.+\left(v_j-\frac{\pi_j y_j}{k_j-\alpha_j}\right)\frac{1}{(\alpha_i-k_j-\mu)(k_i-k_j-\mu}\right)\right). \quad \text{(A.6)}$$

## APPENDIX B

The approximate value function $V_\epsilon$ is given by:

$$\int_t^T \sum_{i=1}^{I}(v_i^{(0)}(s)+\epsilon v_i^{(1)}(s))ds$$

$$=\sum_{i=1}^{I}\left(\frac{\pi_i y_i(1-e^{-\alpha_i(T-t)})}{\alpha_i(k_i-\alpha_i)}+\left(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\right)\right.$$

$$\times\frac{1-e^{-k_i(T-t)}}{k_i}+\epsilon\left[\frac{C_{y_i}\pi_i(1-e^{-\alpha_i(T-t)})}{\alpha_i(k_i-\alpha_i)}\right.$$

$$+\frac{C_{v_i}(1-e^{-k_i(T-t)})}{k_i}+\frac{\beta_i\lambda}{\mu}\left(\frac{\pi_i y_i(e^{-\alpha_i(T-t)}-1)}{\alpha_i(k_i-\alpha_i)^2}\right.$$

$$\left.+\left(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\right)\frac{(k_iT+1)e^{-k_i(T-t)}-(k_it+1)}{k_i^2}\right.$$

$$+\beta_i\left(x-\frac{\lambda}{\mu}\right)\left(\frac{\pi_i y_i(e^{-(\alpha_i+\mu)(T-t)}-1)}{(k_i-\alpha_i)(k_i-\alpha_i-\mu)(\alpha_i+\mu)}\right.$$

$$\left.+\left(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\right)\frac{1-e^{-(k_i+\mu)(T-t)}}{\mu(k_i+\mu)}\right)$$

$$-\pi_i q_{ii}\beta_i\bar{d}_i\frac{\lambda}{\mu}\left(\frac{\pi_i y_i}{(k_i-\alpha_i)^2}\right.$$

$$\times\left(\frac{(\alpha_iT+1)e^{-\alpha_i(T-t)}-(\alpha_it+1)}{\alpha_i^2}+\frac{1-e^{-\alpha_i(T-t)}}{\alpha_i(k_i-\alpha_i)}\right)$$

$$\left.+\left(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\right)\frac{(k_iT+1)e^{-k_i(T-t)}-(k_it+1)}{k_i^2(\alpha_i-k_i)}\right)$$

$$+\pi_i q_{ii}\beta_i\bar{d}_i\left(x-\frac{\lambda}{\mu}\right)\left(\frac{\pi_i y_i(e^{-(\alpha_i+\mu)(T-t)}-1)}{\mu(k_i-\alpha_i)(k_i-\alpha_i-\mu)(\alpha_i+\mu)}\right.$$

$$\left.+\left(v_i-\frac{\pi_i y_i}{k_i-\alpha_i}\right)\frac{e^{-(k_i+\mu)(T-t)}-1}{\mu(\alpha_i-k_i-\mu)(k_i+\mu)}\right)$$

$$+\sum_{j\neq i}^{I}\left(\pi_i q_{ji}\beta_j\bar{d}_j\frac{\lambda}{\mu}\left(\frac{\pi_j y_j(e^{-\alpha_j(T-t)}-1)}{\alpha_j(k_j-\alpha_j)(\alpha_i-\alpha_j)(\alpha_j-k_i)}\right.\right.$$

$$\left.-\left(v_j-\frac{\pi_j y_j}{k_j-\alpha_j}\right)\frac{e^{-k_j(T-t)}-1}{k_j(\alpha_i-k_j)(k_i-k_j)}\right)$$

$$-\pi_i q_{ji}\beta_j\bar{d}_j\left(x-\frac{\lambda}{\mu}\right)$$

$$\times\left(\frac{\pi_j y_j(e^{-(\alpha_j+\mu)(T-t)}-1)}{(k_j-\alpha_j)(\alpha_i-\alpha_j-\mu)(k_i-\alpha_j-\mu)(\alpha_j+\mu)}\right.$$

$$+\left(v_j-\frac{\pi_j y_j}{k_j-\alpha_j}\right)$$

$$\left.\left.\left.\left.\frac{e^{-(k_j+\mu)(T-t)}-1}{(\alpha_i-k_j-\mu)(k_i-k_j-\mu)(k_j+\mu)}\right)\right)\right]\right). \quad \text{(B.1)}$$