Optimal drug pricing, limited use conditions and stratified net benefits for Markov models of disease progression

Gregory S. Zaric
Richard Ivey School of Business, University of Western Ontario, London, Canada
Tel: 1.519.661.3415, email: gzaric@ivey.uwo.ca

Many health payers use cost effectiveness analysis as part of their decisions to fund new drugs. Limited use conditions are a method of directing treatment with new drugs to those populations where they will be most effective. In this paper we investigate how a drug manufacturer could determine pricing and limited use conditions to maximize profits. We assume that the payer makes funding decisions on the basis of net monetary benefits, that the disease can be modeled using a Markov model of disease progression, and that the drug reduces the probability of progression between states of the Markov model.

Introduction

The annual cost of the anti-cancer drug Avastin (bevacizumab) can exceed $100,000 [1]. According to a newspaper article, the manufacturers justified this price by "citing the inherent value of life-sustaining therapies." [1] Referring to the drug's price, an executive for the manufacturer was quoted as saying "right now the health economics hold up, and therefore I don't see any reason to be touching them." [1] Prior to this, manufacturers often justified drug prices by high research costs rather than by using value for money arguments. In this paper we explore the impact of setting prices and use conditions by giving explicit consideration to health economic factors.

Cost effectiveness analysis (CEA) may be used to evaluate the merits of new medical technologies such as drugs, devices, policies and procedures. With cost effectiveness analysis the incremental costs associated with an intervention are divided by the incremental health benefits to produce the incremental cost effectiveness ratio (ICER). The ICER can then be compared with a willingness to pay (WTP) threshold to determine if the intervention is worth funding. The societal WTP can be explicit or inferred from past decisions. For example, a recent review of recommendations made by the UK National Institutes of Clinical Excellence (NICE) found an inferred WTP of £30,000/QALY gained [2].

An alternative to CEA is the net monetary benefits (NMB) approach [3]. The NMB is calculated by multiplying the incremental health benefits by the WTP and subtracting the incremental costs. If the NMB is positive then the intervention is considered worth funding. If the same terms are included in the calculation of incremental costs and benefits, then NMB

and CEA will yield the same result. In particular, a positive NMB is equivalent to an ICER less than the WTP.

Many public healthcare payers have attempted to incorporate health economic considerations into funding and formulary decisions. NICE routinely commissions CEAs of new interventions. The Ontario Drug Benefits plan has a template for formulary submission that includes a request for an estimate of the ICER of the new drug [4]. The Canadian Expert Drug Advisory Committee (CEDAC) often includes economic data as part of the rationale for their recommendations for formulary managers.

Not all interventions are equally effective among all groups of patients. Consequently, not all interventions will be equally cost effective in all groups of patients. For instance, the cost effectiveness of joint replacement surgery varies by age and gender [5]; the cost effectiveness of preventive strategies for women with BRCA1 or BRCA2 mutations may vary with age [6]; and several other examples can be found. If funding or formulary decisions are made on the basis of cost effectiveness (or NMB), then conclusions regarding the appropriateness of funding may be different for different groups.

Many formularies classify listed drugs as being “general benefit” or “limited use”. General benefit drugs will be reimbursed for all uses, while limited use drugs will only be reimbursed for specific conditions. A limited use condition (LUC) is a rule that specifies the conditions under which a drug will be reimbursed. Coyle et al. discussed the possibility of formally using stratified cost effectiveness to set LUCs [7]. They defined a model in which the ICER is calculated for all relevant population subgroups, and they made assumptions about adherence to the LUC in each subgroup. The objective was to use this information to define an optimal LUC using an NMB approach. They presented an example, based on a study of thrombolytic treatments for acute myocardial infarction [8], in which the LUC can be set on the basis of age and location of infarction.

Although Coyle et al. [7] discussed LUCs as a tool that could be used by payers to limit improper use of new technologies, LUCs could be developed by other parties. In particular, this could be done by drug manufacturers in two ways. A manufacturer could only request formulary listing for a restricted subset of patients. Alternately, if a manufacturer knew that a payer would set LUCs for its product on the basis of stratified cost effectiveness, then the manufacturer could request general formulary listing but set its price in anticipation of LUCs imposed by the payer.

In this paper we consider a manufacturer attempting to achieve formulary listing with a third-party payer and we address the following question: If the manufacturer can set the

price and the LUC for the new drug, how should the manufacturer set prices and determine the LUC to maximize profits? In the next sections we develop a Markov model of disease progression and formulate the manufacturer's profit maximization problem. We show that the manufacturer's problem can be decomposed to a single variable optimization problem. We use analytical and numerical techniques to illustrate a number of properties of the optimal profit, drug price, and proportion of the population that receives the drug. We then offer some directions for future research and concluding comments.

The Model

In this section we develop a model that a drug manufacturer could use to set a price and LUCs for some new drug. We assume that the payer is known to make funding decisions on the basis of NMB, that the disease is modeled using a Markov model of disease progression, and that the drug manufacturer sets prices and LUCs to maximize profits. We assume that the population is heterogeneous with respect to the probability of disease progression. We assume that the manufacturer can observe different patient subtypes represented by the differing transition probabilities, and that the manufacturer will consider this when setting limited use conditions.

Markov Model of Disease Progression

We assume that a disease is modeled using the three-stage Markov model (Figure 1), similar to a model used elsewhere [9, 10]. We assume that the disease progresses from stage 1 to stage 2 to death with no recovery. Let $p_{i,j}$ be the one-step probability of moving from stage i to stage j; i = 1, 2; D, j = 1, 2, D. Since there is no recovery $p_{2,1} = 0$. All notation is summarized in Table 1.

Let q_i be the number of quality adjusted life years of survival (QALYs) accrued by an individual during each unit of time spent in stage i. Let c_i be the cost per unit time spent in stage i. We assume that c_i includes all costs except the cost of a the new drug. We assume that $q_D=c_D=0$. Let d be the cost of the new drug per unit time spent in stage 1. Let α be the one-period discount rate applied to all costs and health benefits, $0\leq\alpha\leq1$.

We assume that the new drug reduces the probability of transition from stage 1 to stage 2 from $p_{1,2}$ to $r\times p_{1,2}$, $0<r<1$, and that the new drug is only given to individuals in stage 1. We assume that the new drug has no impact on the probability of death, $p_{1,D}$. We assume that the drug has no impact on any other model parameters. Let $Q_i(p_{1,2})$ and $C_i(p_{1,2})$ be the discounted expected future QALYs and costs, respectively, accrued by an individual

currently in stage i when the transition probability from stage 1 to stage 2 is $p_{1,2}$, i = 1, 2. $Q_1(p_{1,2})$ and $Q_2(p_{1,2})$ are determined by solving the following equations [11]:

$$Q_2\left(p_{1,2}\right) = q_2 + \alpha\left(1 - p_{2,D}\right)Q_2\left(p_{1,2}\right) \quad ,$$

and

$$Q_1\left(p_{1,2}\right) = q_1 + \alpha p_{1,2}Q_2\left(p_{1,2}\right) + \alpha\left(1 - p_{1,2} - p_{1,D}\right)Q_1\left(p_{1,2}\right) \quad .$$

This yields

$$Q_2\left(p_{1,2}\right) = \frac{q_2}{1 - \alpha\left(1 - p_{2,D}\right)} \qquad (1)$$

and

$$Q_1\left(p_{1,2}\right) = \frac{q_1 + \alpha p_{1,2}Q_2\left(p_{1,2}\right)}{1 - \alpha\left(1 - p_{1,2} - p_{1,D}\right)} \quad . \qquad (2)$$

Expressions for $Q_1(rp_{1,2})$ and $Q_2(rp_{1,2})$, the expected future QALYs accrued by individuals receiving the new drugs, are obtained by replacing $p_{1,2}$ with $rp_{1,2}$ in (1) and (2). The corresponding equations for $C_1(p_{1,2})$, $C_2(p_{1,2})$, $C_1(rp_{1,2})$, $C_2(rp_{1,2})$ are obtained by replacing q_i with c_i. Since there is no recovery from stage 2, $Q_2(rp_{1,2})=Q_2(p_{1,2})$ and $C_2(rp_{1,2})=C_2(p_{1,2})$.

Let $D_i(p_{1,2})$ be the discounted expected future cost of the new drug for an individual in stage i, i = 1, 2. Since the new drug is only given to patients in stage 1, $D_2(p_{1,2}) = 0$ and

$$D_1\left(p_{1,2}\right) = \frac{d}{1 - \alpha\left(1 - p_{1,2} - p_{1,D}\right)} \quad . \qquad (3)$$

Net Monetary Benefit of the New Drug

Let λ be the willingness to pay (WTP) for QALYs. Then the net monetary benefit (NMB) of the new drug for an individual whose transition probability from stage 1 to stage 2 is $p_{1,2}$ is given by:

$$NMB\left(p_{1,2}\right) = \lambda\left(Q_1\left(rp_{1,2}\right) - Q_1\left(p_{1,2}\right)\right) - \left(C_1\left(rp_{1,2}\right) - C_1\left(p_{1,2}\right)\right) - D_1\left(rp_{1,2}\right)$$

To keep the notation compact, define $K=1-\alpha(1-p_{1,2}-p_{1,D})$, $K_r=1-\alpha(1-rp_{1,2}-p_{1,D})$, $J_Q=\alpha q_1-\alpha Q_2(p_{1,2})+\alpha^2(1-p_{1,D})Q_2(p_{1,2})$, and $J_C=\alpha c_1-\alpha C_2(p_{1,2})+\alpha^2(1-p_{1,D})C_2(p_{1,2})$. Using this notation, NMB can be rewritten as

$$NMB\left(p_{1,2}\right) = \frac{\left(1-r\right)p_{1,2}}{KK_r}\left(\lambda J_Q - J_C\right) - \frac{d}{K_r} \qquad (4)$$

$$= \frac{1}{KK_r}\left[\left(1-r\right)p_{1,2}\left(\lambda J_Q - J_C\right) - dK\right] \qquad (5)$$

The first term in (4) is the NMB if the new drug was free, and the second term is the discounted expected future cost of the drug for an individual in stage 1. Since d≥0 and K≥0, a payer that funds on the basis of NMB≥0 will only fund a new drug if it would have positive NMB if the drug had no cost.

The Manufacturer's Profit Maximization Problem

We assume that the population is heterogeneous with respect to the transition probability $p_{1,2}$. Let g() and G() be the probability density function (p.d.f.) and cumulative density function (c.d.f.) of $p_{1,2}$. We assume that the manufacturer can identify and use the heterogeneity in $p_{1,2}$ in setting limited use conditions. We assume that the manufacturer sets LUCs by choosing values L and U such that only individuals with $L \le p_{1,2} \le U$ would receive the new drug. We use the term "targeting" to refer to the selection of L and U.

Let m≥0 be the cost to the manufacturer associated with producing enough drug for one patient for one unit of time. The expected revenue to the manufacturer per unit sold is equal to the cost to the payer, given by (3). Let π(d,L,U) be the manufacturer's profit. Then

$$\pi(d,L,U) = (d-m)\int_L^U \frac{1}{K_r} g(p_{1,2})\, dp_{1,2}$$

$$= (d-m)\int_L^U \frac{1}{1-\alpha+\alpha p_{1,D}+\alpha r p_{1,2}} g(p_{1,2})\, dp_{1,2} \quad . \qquad (6)$$

We assume that the payer will only fund a new drug if it has positive NMB for all individuals who may receive the drug. Thus, in order for the drug to be approved by the payer, the following constraint must be satisfied:

$$NMB \ge 0 \ \forall\ p_{1,2} \in [L,U] \quad . \qquad (7)$$

Since K,K_r≥0, (7) can be replaced by

$$(1-r)\, p_{1,2}(\lambda J_Q - J_C) - dK \ge 0 \ \forall\ p_{1,2} \in [L,U] \quad . \qquad (8)$$

Note that (8) requires the drug to be cost effective for everyone as opposed to cost effective on average.

The manufacturer's optimization problem can thus be written as follows:

(M) $$\max_{d,L,U} \pi(d,L,U) = (d-m)\int_L^U \frac{1}{1-\alpha+\alpha p_{1,D}+\alpha r p_{1,2}} g(p_{1,2})\, dp_{1,2} \qquad (6)$$

s.t. $$(1-r)\, p_{1,2}(\lambda J_Q - J_C) - dK \ge 0 \ \forall\ p_{1,2} \in [L,U] \qquad (8)$$

$$d \ge 0 \qquad (9)$$

$$0 \le L \le U \le 1 \qquad (10)$$

Let d*, L*, and U* be the optimal solution to (M), and let π^* be the resulting optimal profit.

It is clear that $d\pi/dU \ge 0$ and $d\pi/dL \le 0$. The derivative of the left hand side of (8) with respect to $p_{1,2}$ is

$$\frac{d}{dp_{1,2}} \lambda(1-r)\, p_{1,2} J_Q - (1-r)\, p_{1,2} J_C - dK = (1-r)\left(\lambda J_Q - J_C\right) - \alpha d \quad . \qquad (11)$$

Let $d' = (1-r)\left(\lambda J_Q - J_C\right)/\alpha$. If $d \ge d'$ then (11) is negative or zero, and if $d \le d'$ then (11) is positive or zero. These observations allow problem (M) to be decomposed into two sub-problems, (M1) and (M2).

(M1) $$\max_{d,L} \pi(d,L) = (d-m)\int_L^1 \frac{1}{1-\alpha+\alpha p_{1,D}+\alpha r p_{1,2}} g\left(p_{1,2}\right) dp_{1,2} \qquad (12)$$

$$(1-r)L\left(\lambda J_Q - J_C\right) - d\left(1-\alpha+\alpha p_{1,D}\right) - \alpha dL \ge 0 \qquad (13)$$

$$d \le (1-r)\left(\lambda J_Q - J_C\right)/\alpha \qquad (14)$$

$$d \ge 0 \qquad (15)$$

$$0 \le L \le 1 \qquad (16)$$

(M2) $$\max_{d,U} \pi(d,U) = (d-m)\int_0^U \frac{1}{1-\alpha+\alpha p_{1,D}+\alpha r p_{1,2}} g\left(p_{1,2}\right) dp_{1,2} \qquad (17)$$

$$(1-r)U\left(\lambda J_Q - J_C\right) - d\left(1-\alpha+\alpha p_{1,D}\right) - \alpha dU \ge 0 \qquad (18)$$

$$d \ge (1-r)\left(\lambda J_Q - J_C\right)/\alpha \qquad (19)$$

$$d \ge 0 \qquad (15)$$

$$0 \le U \le 1 \qquad (20)$$

The optimal solution to (M) can then be found by taking the maximum of the optimal solutions to (M1) and (M2).

Consider problem (M2). There are two cases: $\lambda J_Q - J_C < 0$ and $\lambda J_Q - J_C \ge 0$. First, consider the case $\lambda J_Q - J_C < 0$. Constraint (18) can be rearranged as follows:

$$U\left\{(1-r)\left(\lambda J_Q - J_C\right) - \alpha d\right\} \ge d\left(1-\alpha+\alpha p_{1,D}\right)$$

$$U \le \frac{d\left(1-\alpha+\alpha p_{1,D}\right)}{(1-r)\left(\lambda J_Q - J_C\right) - \alpha d} < 0 \qquad (21)$$

which violates (20). The second inequality in (21) follows from the fact that the numerator is positive and the denominator is negative. Second, consider the case where $\lambda J_Q - J_C \ge 0$. Since 1-

$\alpha-\alpha p_{1,D}=1-\alpha(1-p_{1,D})\geq 0$, the left hand side of (18) is strictly decreasing in d. Thus, the largest value of the left hand side of (18) occurs at the smallest value of d, which occurs when (19) holds as an equality. Then (18) becomes

$$(1-r)U\left(\lambda J_Q - J_C\right)-\left(1-\alpha+\alpha p_{1,D}\right)(1-r)\left(\lambda J_Q - J_C\right)/\alpha-(1-r)U\left(\lambda J_Q - J_C\right)$$

$$=-\left(1-\alpha+\alpha p_{1,D}\right)(1-r)\left(\lambda J_Q - J_C\right)/\alpha \leq 0 \quad . \qquad (22)$$

The inequality in (22) is strict if $\lambda J_Q - J_C > 0$, in which case (18) is violated and there is no feasible solution. Equation (22) holds as an equality when $\lambda J_Q - J_C = 0$, in which case d=0 and $\pi^* \leq 0$. Thus, the only feasible solution to (M2) is d=0 which results in non-positive optimal profit. We therefore ignore (M2) and focus on (M1).

Consider (M1). The feasible region for (M1) is shown in Figure 2. Constraint (14) is redundant given (13). Since $\pi(d,L)$ is increasing in d and decreasing in L, (13) will hold as an equality in the optimal solution. We substitute (13) as an equality into (12). The manufacturer then solves the following one-variable optimization problem:

$$\text{(M1-1)} \max_L \pi(L) = \left(\frac{L(1-r)\left(\lambda J_Q - J_C\right)}{1-\alpha+\alpha p_{1,D}+\alpha L} - m \right) \int_L^1 \frac{1}{1-\alpha+\alpha p_{1,D}+\alpha r p_{1,2}} g\left(p_{1,2}\right) dp_{1,2} \qquad (23)$$

$$0 \leq L \leq 1 \qquad (16)$$

It is clear from (23) that $\pi \leq 0$ when $L \leq 0$ or $L \geq 1$. Thus, (16) can be ignored (if a value of $\pi(L)>0$ cannot be found for $L \in [0,1]$ then the manufacturer would choose to not produce).

Since (23) involves an integral with respect to $p_{1,2}$ and $p_{1,2}$ appears in the denominator of the integrand it is difficult to obtain an algebraic expression for the optimal solution even for the simplest case of $p_{1,2}$ being uniformly distributed in [0,1]. However, we are able to identify a number of properties of the optimal solution, summarized below. The proofs are shown in the appendix.

Properties of the optimal solution to (M1-1): The following are true when $\lambda J_Q - J_C > 0$ and $(1-r)(\lambda J_Q - J_C)/(1+\alpha p_{1,D}) > m$:

i. An optimal solution exists with $\pi^* \geq 0$.

ii. $d^* < \dfrac{(1-r)\left(\lambda J_Q - J_C\right)}{1+\alpha p_{1,D}}$

iii. $\frac{d\pi^*}{dq_1} \geq 0$, $\frac{d\pi^*}{dq_2} \leq 0$, $\frac{d\pi^*}{dc_1} \leq 0$, $\frac{d\pi^*}{dc_2} \geq 0$, $\frac{d\pi^*}{dr} \leq 0$, $\frac{d\pi^*}{dm} \leq 0$, $\frac{d\pi^*}{d\lambda} \geq 0$; $\frac{d\pi^*}{dp_{2,D}}$ has the same sign as $\lambda q_2 - c_2$; $\frac{d\pi^*}{dp_{1,D}}$ may be positive or negative.

iv. Let $L^*(x)$ be the optimal value of L as a function of parameter *x*. When m=0,

$$\frac{dL^*(q_1)}{dq_1} = \frac{dL^*(q_2)}{dq_2} = \frac{dL^*(c_1)}{dc_1} = \frac{dL^*(c_2)}{dc_2} = \frac{dL^*(\lambda)}{d\lambda} = \frac{dL^*(p_{2,D})}{dp_{2,D}} = 0.$$

v. The payer's total NMB is positive at d*, L*. ■

Parts i. and ii. state that a solution exists and that it is optimal for the manufacturer to place an upper limit on price if the payer makes funding decisions based on NMB.

Part iii. states that, for all but two parameters, the optimal profit is a monotonic function of each model parameter. In most instances the relationship is intuitive. For example, $d\pi^*/d\lambda \geq 0$, meaning that the manufacturer's profits increase as the payer's WTP increases; $d\pi^*/dr \leq 0$, meaning that as the impact of the drug increases (i.e., r decreases) the manufacturer's profit increases. Two non-obvious results are $d\pi^*/dq_2 \leq 0$ and $d\pi^*/dc_1 \leq 0$. The first occurs because, as q_2 increases, the health benefit of preventing disease progression from stage 1 to stage 2 decreases. The second occurs because, as c_1 increases, the relative monetary benefit of preventing disease progression decreases. In both cases there is less potential monetary benefit to be gained by introducing the new drug.

Part iv. states that when the marginal cost of producing one unit of the new drug is zero, then the optimal value of L (i.e., the LUC) does not vary in q_1, q_2, c_1, c_2, λ or $p_{2,D}$. Thus, a profit-maximizing manufacturer can determine the proportion of the population that will receive the new drug without knowing any of these terms. The optimal price and profit do, however, depend on these terms.

Part v. states that, even if the payer funds on the basis of net benefit (or cost effectiveness), and if the payer makes its WTP value, λ, openly known to the manufacturer, it is not optimal for the manufacturer to set a price that will extract all WTP from the manufacturer. That is, the optimal decision by the manufacturer will always result in the payer having positive NMB (or, equivalently, ICER below the threshold WTP).

Constraint (13) makes the tradeoff between price and access explicit. The manufacturer is able to charge the highest possible price when L=1. However, the number of people receiving the drug at this price, and hence the manufacturer's profits, would be low.

As the drug is made available to greater numbers of individuals the manufacturer must charge a lower price in order for it to remain cost effective and attractive to the payer.

Numerical Examples

We illustrate with a number of examples. We assume that $p_{1,2}$ is beta-distributed in the interval [0,1]. We use the notation B(a,b) to represent a beta distribution with parameters a and b. The base case data for these examples is shown in Table 2. The calculations were performed in a spreadsheet. The integral in the manufacturer's objective function was evaluated using Simpson's rule [12] with 1,000 intervals in [0,1]. For the case of a uniform distribution of $p_{1,2}$, where the integral can be evaluated analytically, the error of approximation was less than 0.5%. Profit is expressed per person entering stage 1 of the Markov model.

Let d_{max} be the upper bound on d given in part ii of the properties of the solution to (M1-1). In Figure 3 we show the performance of π, d and d_{max} as functions of L for different assumptions about the distribution of $p_{1,2}$. We considered four cases: B(1,1) (which is equivalent to a uniform distribution), B(1,10), B(10,10), and B(10,1). In each row of Figure 3 the left graph shows the distribution of $p_{1,2}$ and the right graph shows the values of π, d and d_{max}. Since d_{max} does not depend on the parameters of the distribution of $p_{1,2}$, d_{max} = \$1,664 for all four distributions of $p_{1,2}$. In all four cases there is a unique optimal solution $L^*>0$, and in all four cases L^* is less than the average of $p_{1,2}$. For B(1,10), in which the average value of $p_{1,2}$ throughout the population is 0.091, $L^*=0.031$. For B(10,1), in which the average value of $p_{1,2}$ is 0.909, $L^*=0.58$. Although L^* is different for each distribution, the differences are not due solely to differences in the average value of $p_{1,2}$: the average value of $p_{1,2}$ is 0.5 for both B(1,1) and B(10,10) but they have different values of L^*.

The optimal profit varies from \$1,838 for B(10,1) to \$4,444 for B(1,10). High profits do not necessarily correspond with high prices. For B(1,10), which has the highest optimal profit, the optimal price is \$746 (45% of the maximum), whereas for B(10,1), which has the lowest optimal profit, the optimal price is \$1618 (98% of the maximum). The extent to which π^* is sensitive to changes in L varies according to the shape of the distribution of $p_{1,2}$. For B(10,1), π^* is relatively insensitive to changes in L, whereas for B(1,1), π^* is quite sensitive to small changes in L. The proportions of the population targeted by the optimal LUC are 91.4%, 73.0%, 98.9% and 99.6% for B(1,1), B(1,10), B(10,10) and B(10,1), respectively. Thus, profits are not necessarily increasing in the proportion of the population that is targeted.

Figure 4 shows the behavior of π^*, d^* and L^* as a function of r for $p_{1,2} \sim B(10,10)$. π^*, d^* and L^* are all decreasing in r. For smaller r (i.e., a more effective drug) the optimal profit increases (as shown in the Properties of the solution to (M1-1)) and the optimal price also increases (Figure 4a). The optimal profit increases significantly, and at a faster rate than the increase in price, as r gets smaller. Although the optimal price increases, d_{max} also increases reflecting the fact that smaller r represents a more beneficial drug. For all values shown in Figure 4, the ratio of d^*/d_{max} is between 89.7% and 90.7%. That is, the upper bound on price is close to the optimal price.

Figure 5 shows the impact of the variance of the distribution of $p_{1,2}$ on L^*, π^* and d^*. The variance in $p_{1,2}$ is a measure of the heterogeneity of the population: high variance corresponds to a highly heterogeneous population, while low variance corresponds to a largely homogeneous population. We varied the variance of the distribution of $p_{1,2}$ by setting both parameters of the underlying beta distribution equal to each other and varying the size of each simultaneously. For all cases the average value of $p_{1,2}$ was 0.5. d^* and π^* decrease approximately linearly in the variance of $p_{1,2}$ (Figure 5a), while L^* decreases at a faster than linear rate (Figure 5b). Thus, for large variance in $p_{1,2}$ (i.e., a relatively heterogeneous population), it is optimal for the manufacturer to set L^* low and target a relatively high proportion of the population, but to compensate by setting price relatively low. For low variance (i.e., a relatively homogeneous population) it is optimal for the manufacturer to be much more restrictive in targeting the new drug. When there is not much heterogeneity the manufacturer is able to charge a higher price and earn a higher profit than when the variance is high.

Figure 6 shows the average incremental cost effectiveness ratio (ICER) of the new drug for the entire population targeted as a function of the variance in $p_{1,2}$. The average ICER is the total incremental cost for all individuals with $p_{1,2} \geq L^*$ divided by the total incremental QALYs for all individuals with $p_{1,2} \geq L^*$. The average ICER is decreasing in the variance of $p_{1,2}$. For very small variance, the average ICER is approximately equal to the WTP for health benefits. Thus, when there is not much heterogeneity, the manufacturer can extract the payer's entire WTP for benefits. However, when there is a high level of patient heterogeneity, it is not optimal for the manufacturer to do so.

A condition in Part iv. of the properties of (M1-1) was that m=0. In extensive numerical testing we were not able to find any instances where these relationships were false for m>0.

Special Case: No Targeting

If targeting is impossible, unethical or otherwise impractical, then L=0 and U=1. The net benefits constraint can be rewritten so that, rather than requiring NMB to be positive everywhere, the total expected NMB is required to be positive. The manufacturer's optimization problem is rewritten as follows:

$$\text{(S)} \qquad \max_d \pi(d) = (d-m)\int_0^1 \frac{1}{1-\alpha+\alpha p_{1,D}+\alpha r p_{1,2}} g(p_{1,2})\,dp_{1,2}$$

$$= (d-m)E\left[\frac{1}{1-\alpha+\alpha p_{1,D}+\alpha r p_{1,2}}\right] = (d-m)E\left[\frac{1}{K_r}\right] \qquad (24)$$

$$\text{s.t.} \qquad (1-r)\left(\lambda J_Q - J_C\right)E\left[\frac{p_{1,2}}{KK_r}\right] - dE\left[\frac{1}{K_r}\right] \geq 0 \qquad (25)$$

$$d \geq 0$$

Since (24) is strictly increasing in p, the optimal solution is

$$d^* = \frac{(1-r)\left(\lambda J_Q - J_C\right)E\left[p_{1,2}/(KK_r)\right]}{E\left[1/K_r\right]} \quad . \qquad (26)$$

We note that total NMB=0 when there is no targeting since (25) is always a binding constraint.

Since L*=0 in (S) and L*≥0 in (M), the case of no targeting always results in more individuals receiving the drug than the case of targeting. For the case of m=0 it is straightforward to show that the optimal profit in (S) is greater than the optimal profit in (M1-1). The optimal profit in (S) can be written in terms of L*, the optimal target LUC, as follows:

$$\pi^* = (1-r)\left(\lambda J_Q - J_C\right)\int_0^1 \frac{p_{1,2}}{KK_r} g(p_{1,2})\,dp_{1,2}$$

$$= (1-r)\left(\lambda J_Q - J_C\right)\left\{\int_0^{L^*} \frac{p_{1,2}}{KK_r} g(p_{1,2})\,dp_{1,2} + \int_{L^*}^1 \frac{p_{1,2}}{1-\alpha+\alpha p1,D+\alpha p_{1,2}} \times \frac{1}{K_r} g(p_{1,2})\,dp_{1,2}\right\}$$

$$\geq (1-r)\left(\lambda J_Q - J_C\right)\left\{\int_0^{L^*} \frac{p_{1,2}}{KK_r} g(p_{1,2})\,dp_{1,2} + \frac{L^*}{1-\alpha+\alpha p1,D+\alpha L^*}\int_{L^*}^1 \frac{1}{K_r} g(p_{1,2})\,dp_{1,2}\right\}. \qquad (27)$$

The first term in the curly brackets in (27) is strictly positive. The second term in the curly brackets is the optimal profit for the case with targeting. Thus, changing the payer's NMB constraint so that the drug must be beneficial on average rather than for everyone results in greater profit for the manufacturer. Note that this does not necessarily hold for m>0.

Extension: Imperfect Targeting

We assumed that targeting would be done on the basis of $p_{1,2}$ and that $p_{1,2}$ could be directly observed by the manufacturer. In reality, the manufacturer may not be able to directly observe this transition probability, but would observe instead some patient characteristic that is correlated with this probability. For example, the transition probability from stage 1 to stage 2 may be correlated with age, so that the distribution of $p_{1,2}$ throughout the population would be determined by the distribution of age throughout the population.

Rather than setting LUCs on the basis of transition probability values, the manufacturer would choose upper and lower ages. Let t be the "targeting characteristic". Let T_L and T_U be the lower and upper limits on t, such that the drug is only available to those individuals with $T_L \leq t \leq T_U$. Let h(t) be the p.d.f. of t, and let $g(p_{1,2}|t)$ be the p.d.f. of $p_{1,2}$ given t. Then the manufacturer's profit function would be rewritten as:

$$\pi\left(d, T_L, T_U\right) = \left(d - m\right) \int_{T_L}^{T_U} \left(\int_0^1 \frac{1}{1 - \alpha + \alpha p_{1,D} + \alpha r p_{1,2}} g\left(p_{1,2} \mid t\right) dp_{1,2} \right) h(t) dt \quad .$$

The NMB, (4), would also be rewritten to reflect the fact that t, and not $p_{1,2}$, is being used for targeting:

$$E\left[NMB(t)\right] = \int_0^1 \left(\frac{(1-r)\, p_{1,2}}{KK_r} \left(\lambda J_Q - J_C\right) - \frac{d}{K_r} \right) g\left(p_{1,2} \mid t\right) dp_{1,2} \quad .$$

Constraint (7) would also be rewritten as follows:

$$E\left[NMB(t)\right] \geq 0 \;\forall\; t \in \left[T_L, T_U\right] \quad .$$

This could easily be extended to the case where t correlates with more than one heterogeneous parameter by using the joint distribution of a vector of parameters given t.

Discussion

In this paper we developed a model that drug manufacturers could use to optimally set prices and LUCs when payers make funding decisions on the basis of NMB and the payer's WTP is known. We formulated the manufacturer's profit maximization problem as a three variable optimization problem and decomposed it into an equivalent unconstrained one-variable problem. We identified a number of properties of the optimal solution, both analytically and numerically.

The model is easily extended to the case of more than 3 stages with no recovery. Regardless of the number of stages, the expected future costs and QALYs in stage 2 do not depend on $p_{1,2}$. Thus, the expression for Q_2, C_2 and D_2 for such a model could be substituted into their equivalents in the current model without fundamentally changing the analysis.

We assumed that there was a single source of patient heterogeneity, and that this heterogeneity was reflected in the probability of disease progression in the Markov model. In reality, there may be multiple sources of patient heterogeneity. For example, there may be heterogeneity in preferences for health states [13], and numerous studies have demonstrated variability in costs and resource utilization .

We assumed that LUCs would be adhered to perfectly and at no cost. Both of these assumptions may be violated in practice. Incorporating a cost of enforcement would be straightforward. Coyle et al. [7] discussed the concept of "leakage", in which a certain percentage of individuals not targeted would receive the drug anyway. Future modeling efforts could incorporate the impact of leakage.

Many payers are concerned not only with cost effectiveness but also with total budget. It would be relatively straightforward to add a budget constraint to the manufacturer's model, although a manufacturer would likely possess a great deal of uncertainty about the size of a payer's budget for any one particular drug. We assumed that the production cost m was constant regardless of the number of people who receive the drug. Future research could model m as a function of L*.

Although this work was presented from the manufacturer's perspective, it is also important to payers. Understanding how manufacturer's optimally set prices can give payers a sense of the impact of changes in policy. In some areas sophisticated contracts involving risk sharing schemes are being negotiated between payers and manufacturers. A model such as the one in this paper may be beneficial to payers when designing such contracts.

Table 1: Summary of Notation

Manufacturer's Decision Variables	
d	Price of the new drug per unit time
L	Lower limit on $p_{1,2}$
U	Upper limit on $p_{1,2}$
Manufacturer's Objective Function	
$\pi(d,L,U)$	Profit as a function of the manufacturer's decisions
m	Marginal production cost of the new drug
Probabilities	
$p_{i,j}$	Probability of progressing from stage i to stage j in one time step
r	Relative probability of progressing from stage 1 to stage 2 for an individual receiving the new drug, $0<r<1$
Costs and QALYs	
q_i	Number of QALYs accrued per time step in stage i, i = 1, 2
c_i	cost per time step accrued in stage i, i = 1, 2
$Q_i(p_{1,2})$	Discounted expected future QALYs accrued by an individual in stage i when the transition probability from stage 1 to 2 is $p_{1,2}$
$C_i(p_{1,2})$	Discounted expected future costs, excluding the new drug costs, accrued by an individual in stage i when the transition probability from stage 1 to 2 is $p_{1,2}$
$D_i(p_{1,2})$	Discounted expected future costs of the new drug accrued by an individual in stage i when the transition probability from stage 1 to 2 is $p_{1,2}$
α	Discount rate per time step
Net Monetary Benefits	
λ	Willingness to pay for QALYs
NMB	Net monetary benefits
Defined Quantities	
K	$K=1-\alpha(1-p_{1,2}-p_{1,D})$
K_r	$K_r=1-\alpha(1-rp_{1,2}-p_{1,D})$
J_Q	$J_Q=\alpha q_1-\alpha Q_2(p_{1,2})+\alpha^2(1-p_{1,D})Q_2(p_{1,2})$
J_C	$J_C=\alpha c_1-\alpha C_2(p_{1,2})+\alpha^2(1-p_{1,D})C_2(p_{1,2})$

Table 2: Base case values for numerical illustrations.

Parameter	Value
$p_{1,D}$	0.01
$p_{2,D}$	0.05
r	0.95
α	0.97
λ	\$50,000
m	0
q_1	1
q_2	0.8
c_1	\$1,200
c_2	\$12,000
$f(p_{1,2})$	B(10,10)

Figure 1: Markov model of disease progression.

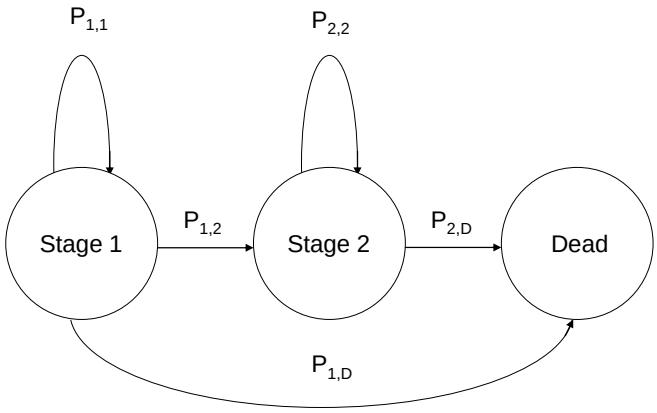


Figure 2: Feasible region for (M1). The arrows represent numbered constraints. The shaded area represents the feasible region.

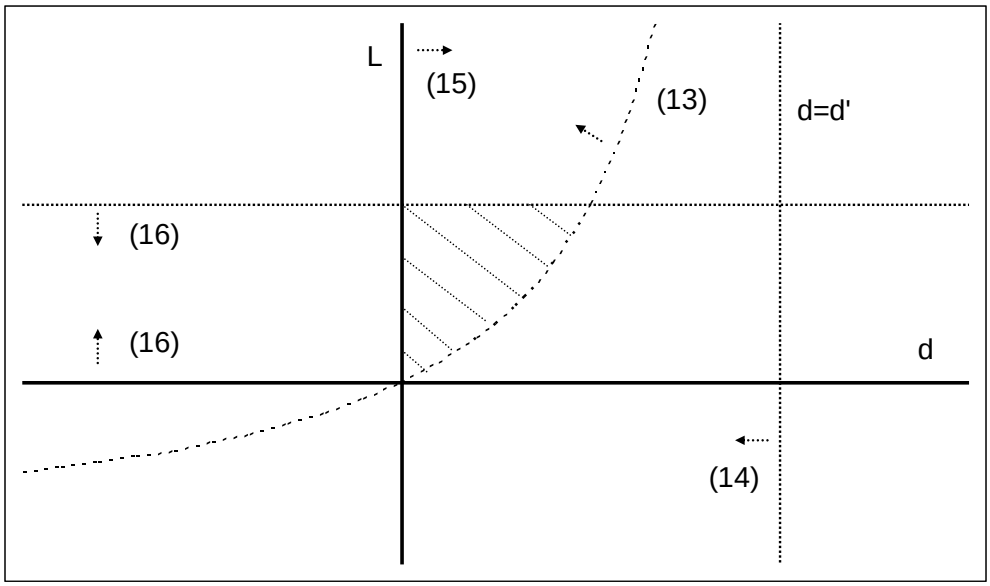
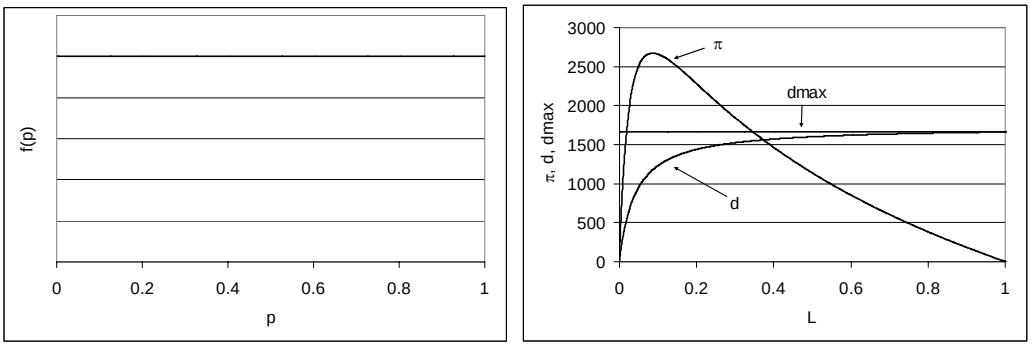
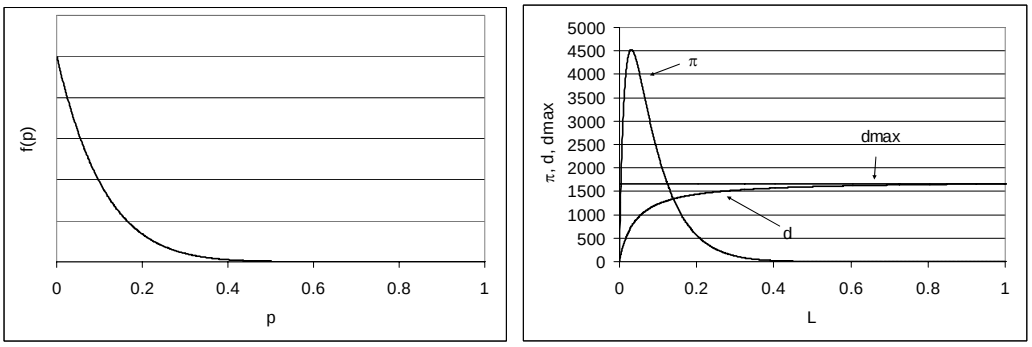


Figure 3: Optimal profit (π) and drug cost (d) as a function of L for different assumptions about the distribution of L.

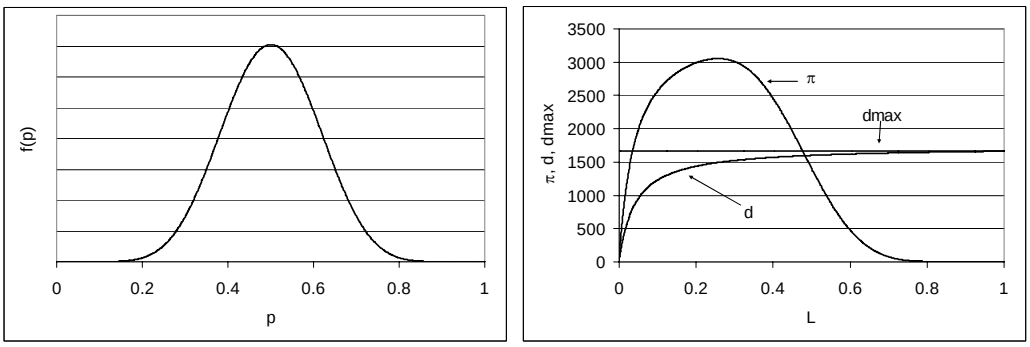
B(1,1)



B(1,10)



B(10,10)



B(10,1)

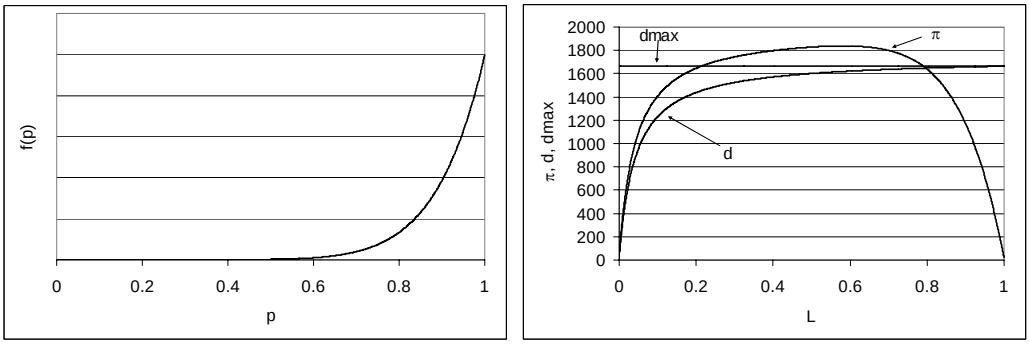


Figure 4: π* and d* as a function of r.

Figure 4a

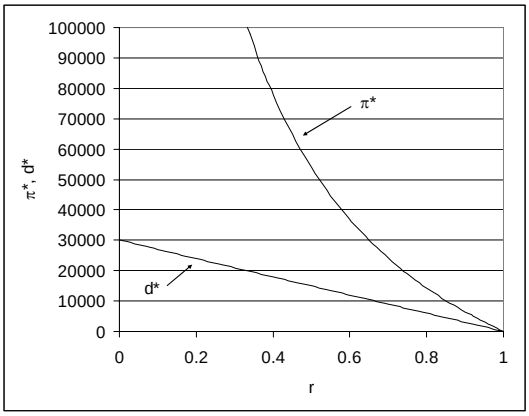


Figure 4b

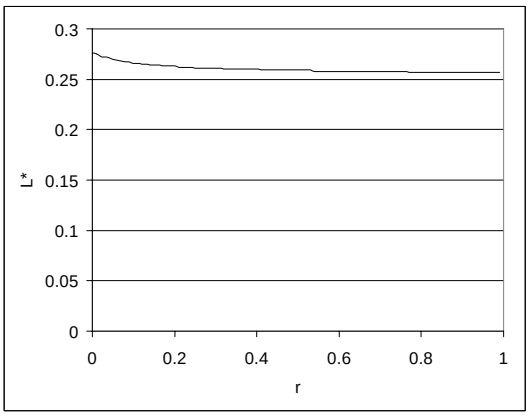


Figure 5: L*, π*, d*/d_{max} and ICER* as a function of the variance of L.

Figure 5 a

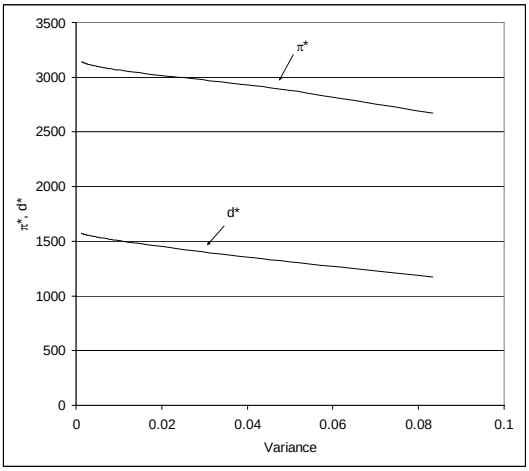


Figure 5b

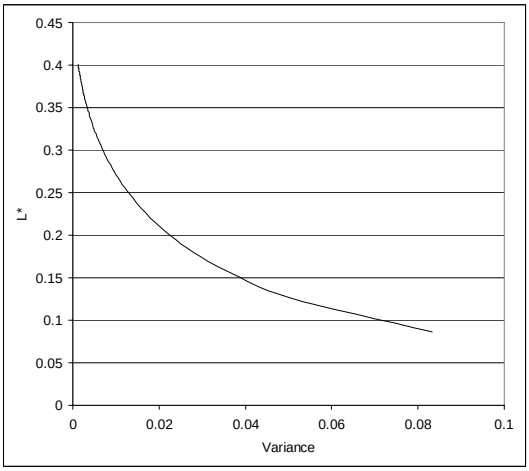
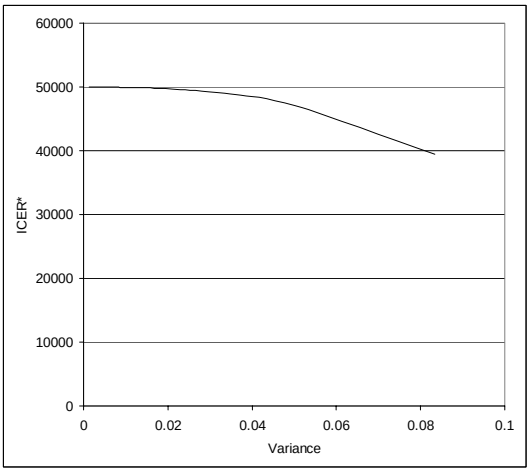


Figure 6: Average incremental cost effectiveness ratio as a function the variance of the distribution of $p_{1,2}$.



References

1. Berenson, A., *A cancer drug's big price rise is cause for concern*, in *New York Times*. 2006: New York.
2. Devlin, N. and D. Parkin, *Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis.* Health Econ, 2004. **13**(5): p. 437-52.
3. Tambour, M., N. Zethraeus, and M. Johannesson, *A note on confidence intervals in cost effectiveness analysis.* Int J Tech Assess Health Care, 1998. **14**(3): p. 467-71.
4. Detsky, A.S., *Guidelines for economic analysis of pharmaceutical products: a draft document for Ontario and Canada.* Pharmacoeconomics, 1993. **3**(5): p. 354-61.
5. Chang, R.W., J.M. Pellisier, and G.B. Hazen, *A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip.* JAMA, 1996. **275**(11): p. 858-65.
6. Anderson, K., et al., *Cost-effectiveness of preventive strategies for women with a BRCA1 or a BRCA2 mutation.* Ann Intern Med, 2006. **144**(6): p. 397-406.
7. Coyle, D., M.J. Buxton, and B.J. O'Brien, *Stratified cost-effectiveness analysis: a framework for establishing efficient limited use criteria.* Health Econ, 2003. **12**(5): p. 421-7.
8. Mark, D.B., et al., *Cost effectiveness of thrombolytic therapy with tissue plasminogen activator as compared with streptokinase for acute myocardial infarction.* N Engl J Med, 1995. **332**(21): p. 1418-24.
9. Kuntz, K.M. and S.J. Goldie, *Assessing the sensitivity of decision-analytic results to unobserved markers of risk: Defining the effects of heterogeneity bias.* Med Decis Making, 2002. **22**: p. 218-227.
10. Zaric, G.S., *The impact of ignoring population heterogeneity when Markov models are used in cost effectiveness analysis.* Medical Decision Making, 2003. **23**(5): p. 379-96.
11. Taylor, H.M. and S. Karlin, *An introduction to stochastic modeling. Revised Edition*. 1994, San Diego: Academic Press, Inc. 566.
12. Adams, R.A., *Single variable calculus*. 1986, Don Mills, Ontario: Addison-Wesley Publishers Limited. 643.
13. Sculpher, M. and A. Gafni, *Recognizing diversity in public preferences: The use of preference sub-groups in cost effectiveness analysis.* Health Economics, 2001. **10**(4): p. 317-324.
14. Mas-Colell, A., M.D. Whinston, and J.R. Green, *Microeconomic theory*. 1995, New York: Oxford University Press.

Appendix

i. True because (M1) involves optimizing a continuous function over a closed set.

ii. The inequality is found by substituting L=1 into (13). The inequality is strict because $\pi=0$ when L=1 and $\pi>0$ for some $L\in[0,1)$.

iii. The remaining proofs are derived from straightforward applications of the envelope theorem [14].

Expressions for these terms are shown below. Let $H=1-\alpha+\alpha p_{1,D}$.

$$\frac{d\pi^*}{dq_1}=\left(\frac{L^*}{H+\alpha L^*}\right)(1-r)\lambda\alpha\int_{L^*}^{1}\frac{1}{H+\alpha r p_{1,2}}g\left(p_{1,2}\right)dp_{1,2}\geq 0$$

$$\frac{d\pi^*}{dq_2}=\left(\frac{L^*}{H+\alpha L^*}\right)(1-r)(\lambda)(-\alpha)\frac{1-\alpha\left(1-P_{1,D}\right)}{1-\alpha\left(1-P_{2,D}\right)}\int_{L^*}^{1}\frac{1}{H+\alpha r p_{1,2}}g\left(p_{1,2}\right)dp_{1,2}\leq 0$$

$$\frac{d\pi^*}{dc_1}=\left(\frac{L^*}{H+\alpha L^*}\right)(1-r)(-\alpha)\int_{L^*}^{1}\frac{1}{H+\alpha r p_{1,2}}g\left(p_{1,2}\right)dp_{1,2}\leq 0$$

$$\frac{d\pi^*}{dc_2}=\left(\frac{L^*}{H+\alpha L^*}\right)(1-r)(-1)(-\alpha)\frac{1-\alpha\left(1-P_{1,D}\right)}{1-\alpha\left(1-P_{2,D}\right)}\int_{L^*}^{1}\frac{1}{H+\alpha r p_{1,2}}g\left(p_{1,2}\right)dp_{1,2}\geq 0$$

$$\frac{d\pi^*}{dr}=\left(\frac{L^*}{H+\alpha L^*}\right)(-1)\left(\lambda J_Q-J_C\right)\int_{L^*}^{1}\frac{1}{H+\alpha r p_{1,2}}g\left(p_{1,2}\right)dp_{1,2}+\left(\frac{L^*G}{H+\alpha L^*}-m\right)\left[\int_{L^*}^{1}\frac{-\alpha p}{\left(H+\alpha r p_{1,2}\right)^2}g\left(p_{1,2}\right)dp_{1,2}-\frac{}{H}\right.$$

$$\frac{d\pi^*}{dm}=-\int_{L^*}^{1}\frac{1}{H+\alpha r p_{1,2}}g\left(p_{1,2}\right)dp_{1,2}\leq 0$$

$$\frac{d\pi^*}{d\lambda}=\left(\frac{L^*}{H+\alpha L^*}\right)(1-r)J_Q\int_{L^*}^{1}\frac{1}{H+\alpha r p}g\left(P_{1,2}\right)dp\geq 0$$

The proof that $d\pi^*/dp_{1,D}$ may be positive or negative is easily shown numerically and omitted here.

iv. First, consider q_1. The first order necessary condition (FONC) for optimal L is $d\pi/dL=0$. Let $L^*(q_1)$ be the value of L, as a function of the parameter q_1, that satisfies this condition. Then the optimal profit can be written as $\pi^*(L^*(q_1), q_1)$. Applying the chain rule to the FONC yields

$$\frac{dL^*\left(q_1\right)}{dq_1}=\left.\left(d^2\pi/dLdq_1\right)\middle/\left(d^2\pi/dL^2\right)\right|_{L=L^*}$$

Let $G=(1-r)(\lambda J_Q-J_C)$ and let $H=1-\alpha+\alpha p_{1,D}$. The FONC at the optimality is:

$$\frac{d\pi}{dL}=\left(\frac{GH}{\left(H+\alpha L\right)^2}-m\right)\int_{L}^{1}\frac{1}{H+\alpha r p_{1,2}}g\left(p_{1,2}\right)dp_{1,2}-\left(\frac{LG}{H+\alpha L}-m\right)\left(\frac{1}{H+\alpha L}\right)g\left(L\right)=0\quad,$$

which becomes

$$\frac{d\pi}{dL}=G\left\{\left(\frac{H}{\left(H+\alpha L\right)^2}\right)\int_{L}^{1}\frac{1}{H+\alpha r p_{1,2}}g\left(p_{1,2}\right)dp_{1,2}-\left(\frac{L}{H+\alpha L}\right)\left(\frac{1}{H+\alpha L}\right)g\left(L\right)\right\}=0$$

when m=0. Since $\lambda J_Q-J_C>0$ and $0<r<1$, $G>0$. Thus, the term inside the curly brackets must equal zero. The terms q_1 only appears in G. Thus, $\frac{d^2\pi}{dLdq_1}=\frac{dG}{dq_1}\times 0=0$.

The argument is the same for q_2, c_1, c_2, λ and $p_{2,D}$ as they also only appear in G.

v. This is true because (13) is a binding constraint in the optimal solution and NMB≥0 for all $p_{1,2}\geq L^*$.

■