Valuation of Biodiversity for Use in New Product Research in a Model of Sequential Search

R. David Simpson
Roger A. Sedjo

Discussion Paper 96-27

July 1996

Resources for the Future
1616 P Street, NW
Washington, DC 20036
Telephone 202-328-5000
Fax 202-939-3460

© 1996 Resources for the Future. All rights reserved.
No portion of this paper may be reproduced without permission of the authors.

Discussion papers are research materials circulated by their authors for purposes of information and discussion. They have not undergone formal peer review or the editorial treatment accorded RFF books and other publications.

Valuation of Biodiversity for Use in New Product Research in a Model of Sequential Search

R. David Simpson and Roger A. Sedjo

Abstract

We develop a model of search in which a researcher chooses the size of sequential batches of samples to test. While earlier work has considered similar questions, the contribution of this paper is to use the search model to place a value on the marginal research opportunity. The valuation of such opportunities may be of little interest or relevance in many of the contexts in which search models are employed, but we apply our analysis to an area of considerable societal interest: the valuation of biological diversity for use in new product research. While data from which to make inferences are limited, we find that, using plausible estimates of relevant parameters, the value of biodiversity in these applications is negligible.

Keywords: biodiversity, search, sequential, conservation incentives

JEL Classification Nos.: D83; Q29

Table of Contents

List of Figures and Tables

Valuation of Biodiversity for Use in New Product Research in a Model of Sequential Search

R. David Simpson and Roger A. Sedjo[1]

INTRODUCTION

At a formal level, this paper is about optimal search. We imagine a researcher who can choose to pursue any subset or all of *n* separate "research opportunities" in pursuit of an objective. We assume that each research opportunity can either be a "success" or a "failure." As additional "successes" would be redundant, search is suspended when a success is achieved. In the initial period of search, and in each period following a period in which search has been unsuccessful, the researcher must choose how many of the remaining research opportunities open to her she will pursue. Pursuing a greater number of research opportunities simultaneously will increase the probability that at least one success will be achieved. The more research opportunities are pursued simultaneously, however, the more likely it is that two or more successes will be recorded. As additional successes are redundant, costs incurred to achieve them are, at least *ex post*, wasteful.[2]

---

[1] The authors are, respectively, Fellow and Senior Fellow, Energy and Natural Resources Division, Resources for the Future. This work has been supported by subcontract DE-FG02-95ER62113 from the United States Department of Energy Office of Energy Research's Global Change Integrated Assessment Research Program. We thank seminar participants at Montana State University for helpful comments on an earlier version of this work. We are grateful to Gar Ragland for help with calculations and figures. All opinions expressed are those of the authors only.

[2] If we relaxed the assumption that each sample evaluation is an independent, identical Bernoulli trial, subsequent "successes" might generate incremental benefits, but the basic principle remains the same: once one product of value is found, additional "successes" are of less value. See the discussion in Section III.

What we have described is a model of optimal search intensity. This subject received attention in the early literature on optimal search. Stigler's seminal paper (1961) asks how large a fixed sample ought to be collected before making the choice of which alternative to accept. Other work (e.g., Benhabib and Bull, 1983; Morgan and Manning, 1985) considers the choice of how large a set of opportunities to search over at any given time. The focus of this paper, however, is not on the optimal choice of batch size in conducting such search, but rather, on an issue that does not come up in the contexts to which search models are typically applied. We ask what value the researcher would assign to having an additional research opportunity: what would she be willing to pay to preserve an n+1*st* opportunity that she may pursue if search is unsuccessful in all other attempts?

This question is not of much interest in many of the economic contexts to which search models have been applied. In some instances (in models of labor markets and job acceptance, for example), it is assumed that the rate at which "research opportunities" (or job offers, or whatever) arrive is exogenous. In other cases, the number of "research opportunities" (or whatever the analogue may be in the particular application) may be assumed to be either infinite or fixed. In the former case the value of the "marginal research opportunity" must be zero, while in the latter the preservation of the option to pursue the "marginal research opportunity" is typically regarded as being beyond the researchers' control.

This paper, however, is motivated by an important real-world problem in which the number of research opportunities is finite and, it is argued, the preservation of these research opportunities is very much within society's control. Some natural scientists believe that biological diversity may decline rapidly in the near future (see, e.g., Ehrlich and Ehrlich, 1981;

Myers, 1986; and Simberloff, 1986). One of the reasons for which biodiversity is felt to be valuable is as a source of new products. Natural organisms have evolved elaborate chemical mechanisms to enhance growth, attract mates, capture prey, avoid predators, and resist infection. If these chemical leads could be adapted for commercial production, they could be of great value in industrial, agricultural, and, particularly, pharmaceutical applications. This search for commercial products among wild genetic resources has been labeled "biodiversity prospecting." A number of authors have argued that, if the proceeds of biodiversity prospecting were dedicated to the maintenance of the natural habitats in which endangered species are found, further incentives for the preservation of such endangered species would be generated (see, e.g., Eisner, 1992; Wilson, 1992; Rubin and Fish, 1995). A number of economists have begun to look at the magnitudes of the values generated by biodiversity prospecting; recent studies include those of Pearce and Puroshothamon (1992), Aylward (1993), Artuso (1994), Mendelsohn and Balick (1995), and Polasky and Solow (1995).

Most existing estimates of the value of biodiversity for use in new product research and development have examined the total value generated by biodiversity prospecting.[3] In an earlier paper (Simpson, Sedjo, and Reid, 1996) we argue that the economically relevant question concerns not the total value of a collection of species, or the value of one collection of species in isolation from all others, but rather the value of the "marginal species." Conservation incentives

---

[3] Because little evidence is available on the terms of market transactions, there is little data from which directly to infer the value of biodiversity *in situ* for new product research. Such deals as are observed generally have secret royalty provisions. In addition, compensation is typically paid not just for access to biological resources *in situ*, but also for the provision of collection and processing labor and expertise. For a more complete discussion, see Simpson, Sedjo, and Reid, 1996.

do not typically involve questions of whether or not to eradicate biodiversity entirely (even in a specific region), but rather, whether incrementally to decrease the size of a natural habitat, and, hence incrementally decrease the number of species the habitat supports.[4] The relevant issue with respect to the conservation incentives engendered by biodiversity prospecting is whether a hectare of land will yield greater returns as habitat for endangered species used in new product research (in addition to whatever other earnings arise from its continuing existence in a more-or-less "natural" state), or converted to other applications (in which, presumably, its biodiversity would be substantially reduced). The value of the marginal species in commercial bioprospecting operations is, then, an extremely important question in the formulation of conservation policy.[5]

In our earlier paper we argued that the value of the marginal species for use in pharmaceutical research is likely to be small. This conclusion was based largely on a consideration of the number of species available for testing for their pharmaceutical potential; a sort of diamonds-and-water paradox of value arises. If any species tested at random might be expected to yield a success, it would also be likely that two or more species in a large collection

---

[4] This statement requires some explanation. While much is made of the concern that incremental reductions in habitat size or quality can trigger discontinuous effects on biodiversity, commonly used models of biogeography and species distribution presume an approximately (abstracting from the fact that the number of species is necessarily confined to integer values) continuous, smooth relationship between habitat area and numbers of species supported (see, e.g., MacArthur and Wilson, 1967), and these models are cited even by those who argue most forcefully for the preservation of species (see, e.g., Myers, 1988; 1990; Wilson, 1992).

[5] This question may be, in fact, more important than an economist might generally believe. Some influential advocates of conservation argue for investments in developing-country biodiversity prospecting operations as a means of augmenting conservation incentives. To the extent that one cannot “make something from nothing–” that is, that rents are only earned on scarce resources–an investigation into the relative scarcity of biodiversity in new product research should have very important implications for conservation policy.

will yield the product sought. Thus, the expected value of the marginal species would be low since the probability of redundancy rises as the size of the collection increases. Conversely, if the probability of success in any single trial were low, the probability of redundancy would also be low, but so would be the probability of encountering *any* successes. We found in our earlier paper that there exists a probability of success at which the value of the marginal species reaches an upper bound (given a collection of species of a fixed size from which to search), and that this upper bound declines as the size of the collection of species over which search is to be conducted increases.

In order to calibrate the model in our earlier paper to real-world data, however, it was necessary to make one *ad hoc* assumption: that the overall rate of profit (expected revenues divided by expected costs) is no greater than some hypothesized level. While such an assumption may seem reasonable, it seems reasonable precisely because one might expect that, if a high rate of return could be realized by investing in biodiversity prospecting, this rate of return would be lowered by increasing the investment. There is nothing inconsistent about large rents arising from scarce resources, however. Thus, in order to defend the assertion that the value of the marginal species is not great, we must consider how that value is affected by profit-maximizing expenditures on new product research.

Important omissions from our earlier model were any consideration of the time required to evaluate samples or the intensity of effort undertaken in testing. We incorporate these considerations in this paper. Again using a simple specification of the sample evaluation process as a series of independent Bernoulli trials, we suppose that each test requires a certain period of time to perform. Each test is also costly to perform. The supplier of genetic resources (in

practice, the country or organization within a country controlling access to biological diversity) makes a decision as to how many species to subject to testing (either testing them itself or selling access to others to test) in each time period.

There is a tradeoff inherent in increasing the number of species tested simultaneously. On one hand, by testing more species at once, it becomes more likely that the product sought will be discovered sooner rather than later, and the corresponding benefits received sooner. On the other hand, the more species tested at any one time, the more likely it is that unnecessary costs will be incurred in developing redundant products. The optimal testing strategy is that which balances these considerations. Evidence concerning the rates at which biodiversity prospecting is now being conducted suggests that the value of the marginal species is negligible. There are enough good leads available for pharmaceutical research that the additional opportunities represented by as yet untested (and, perhaps, undiscovered or undescribed) species are negligible.

In the remainder of this paper we develop a model of biodiversity prospecting with simultaneous sample evaluation. The next section introduces and develops the formal model of optimal research intensity. In the second section we derive an upper bound on the value of the marginal species and present some numerical results concerning the likely magnitude of that upper bound. We then conduct some sensitivity analyses. We find that the value of the marginal species under any plausible conjectures about the rapidity with which testing is conducted is relatively small.

It may be wise for us now, at the beginning of the paper, to disabuse readers of any impression that we are going to present an accurate and realistic description of the use of natural

products in pharmaceutical research. We are not. It is simply too difficult to create a realistic yet tractable model. We will, however, make two related arguments. First, if the world worked as our simple model describes it, the value of biodiversity for use in pharmaceutical research would be negligible. Second, the many simplifying assumptions we make in developing a tractable model are likely to bias our estimates of value upward rather than downward. The third section discusses limitations of the present model, its possible extensions, and the reasons for which we believe the model, while admittedly limited and schematic, nevertheless provides important insights. A final section briefly concludes. Some formal details of the analysis are somewhat tedious and are relegated to an appendix.

I. A MODEL WITH OPTIMAL INTENSITY OF SIMULTANEOUS TESTING

We will model species as independent identical Bernouli trials with respect to their potential as the source of a new product. That is, each species that might be tested will be presumed to be equally likely to be the source of any particular product for which search is undertaken. Once one "hit" is made, additional hits are completely superfluous with respect to that application, so search is suspended.

In what follows we are going to ask how an optimal prospecting plan would proceed. That is, how many species would be chosen for evaluation as a function of the number of species remaining to be tested? Denote by $V[n, m]$ the net present expected value of prospecting when n species remain to be sampled, m species are chosen for testing in the present period, and, if no successful product is discovered among the species tested in the present period, sampling and

testing continues in optimal fashion. Denote by $V^*(n)$ the net present expected value of a prospecting program when n species remain to be tested and m is chosen optimally.

We assume that each species available for testing represents an independent identical Bernoulli trial with probability p of success. Success is rewarded with profits R, which are assumed to be net of production and marketing expenses, but gross of research and development costs. Costs of research and development are c; in other words, c is what it costs to determine whether or not a species is worth R or nothing in the application for which it is being tested.

The expected present-period payoff to sampling all species in a collection of size m is

$$R\left[1 - (1-p)^m\right] - mc\,. \tag{1}$$

That is, it is the probability that *at least one* success is encountered, times the payoff in the event of a success, less the cost of testing the m species. We might justify (1) in a number of ways. It is the payoff a single researcher would expect in the simultaneous evaluation of m independent samples. It is also the aggregate expected payoff m researchers would earn if each evaluated one sample and all successful researchers share in the resultant payoff equally, or, equivalently, each successful researcher has an equal probability of being the sole recipient of an exclusive patent on the product. The equivalence of these two interpretations is demonstrated in the appendix. Finally, and less rigorously, we might think of (1) as representing a possible outcome if each successful researcher were to develop a differentiated product. Equation (1) would then embody the assumption that the dissipation in profits resulting from interproduct competition is balanced by the rents arising from product differentiation in exactly such a way as to keep total profits constant regardless of the number of successes.

We can use (1) to construct the expected net present value V(*n*, *m*) using the notation we introduced above:

$$V(n,\, m) \;=\; R\left[1-(1-p)^m\right] \;-\; mc \;+\; \delta(1-p)^m V^*(n-m). \qquad (2)$$

We are going to proceed by differentiating with respect to *m* and imposing the first-order condition for maximization. This procedure presumes that expression (2) is, in fact, differentiable[6] with respect to *m* (and, more to the point, that what we have labeled as V*(*n*-*m*) is differentiable with respect to *n* and *m*). We will also assume throughout that second-order conditions are satisfied. The justification of these assumptions here would fragment the exposition, so we refer interested readers to the appendix for details.

Differentiating (2) with respect to *m*, we have

$$\frac{\partial V(n,\, m)}{\partial m} = -R\ln(1-p)(1-p)^m - c + \delta\ln(1-p)(1-p)^m V^*(n-m) - \delta(1-p)^m \frac{\partial V^*(n-m)}{\partial n}$$

(because $\frac{\partial V^*(n-m)}{\partial n} \;=\; \frac{-\partial V^*(n-m)}{\partial m}$).

We will, in all that follows, assume that *p* is relatively small. This assumption can be justified on two grounds. First, the probability of success in any given sample evaluation in a biodiversity prospecting program is small; experts' estimates of the probability of finding a commercial product in any given sample evaluation are typically on the order of one in ten

---

[6] By talking about "differentiability," we are also abstracting from the fact that the number of species, and the number of species tested at any given time, is necessarily an integer. Inasmuch as *n* is certainly a large number, and *m* will be a relatively large number if the value of the marginal species is nonnegligible (if *m* were a small number, it would mean that few species are tested at any one time, and hence that the marginal species would not be tested until the distant future, meaning that its discounted present value would be negligible), the simplifying assumption of differentiability seems innocuous.

thousand (as reported, for example, by Roberts, 1992). Second, our emphasis in what follows will be on the valuation of the "marginal species." What is it worth, we ask, to have one more species available for testing? If p were not a very small number, the value of the marginal species would necessarily be negligible: the number of species available for testing is large, so the probability with which the marginal species would prove redundant with others approaches one. There is another consideration in this model. If p were not small and c were nonnegligible, an optimal strategy would involve testing relatively few species simultaneously, so as to avoid redundant costs. Under these conditions, testing large numbers of species simultaneously would likely involve excessive expense incurred in replicating "hits." Hence, the testing of the marginal species might be deferred to the distant future, and, due to discounting, the value of the marginal species would again be negligible. Under the assumption that p is small, the approximation $\ln(1-p) \approx -p$ is very close.[7] Thus, we will use this approximation in all subsequent expressions in which $\ln(1-p)$ would appear.[8] Restate the above derivative, then, as

$$\frac{\partial V(n, m)}{\partial m} = pR(1-p)^m - c - \delta p(1-p)^m V^*(n-m) - \delta(1-p)^m \frac{\partial V^*(n-m)}{\partial n}. \qquad (3)$$

Let $\mu(n)$ represent the optimal number of species to test in the current period when n species remain untested. When m is chosen optimally, i.e., when $m = \mu(n)$, the left hand side of

---

[7] By "small," we typically mean a probability of discovery on the order of one in ten thousand or less (Roberts, 1992). To nine decimal places, $\ln(1-0.0001) = -0.000100005$.

[8] Abusing notation slightly, but in hopes of reducing eye strain, we will use equal signs, rather than approximation signs, in all expressions in which the approximation is employed.

(3) is zero,[9] so, rearranging, we have

$$\frac{\partial V^*(n-\mu(n))}{\partial n} = \frac{Rp(1-p)^{\mu(n)} - c}{\delta(1-p)^{\mu(n)}} - pV^*(n-\mu(n)). \quad (4)$$

We will refer to the left-hand side of (4) as the "value of the marginal species" when n - $\mu(n)$ species remain untested. Consider next the case in which m is chosen optimally, and now differentiate (2) with respect to n. Doing so, we obtain

$$\frac{dV(n,\,\mu(n))}{dn} = \left[pR(1-p)^{\mu(n)} - c - \delta p(1-p)^{\mu(n)} V^*(n-\mu(n)) - \delta(1-p)^{\mu(n)} \frac{\partial V^*(n-\mu(n))}{\partial n} \right] \frac{\partial \mu(n)}{\partial n}$$
$$+ \delta(1-p)^{\mu(n)} \frac{\partial V^*(n-\mu(n))}{\partial n}$$

Since μ is chosen optimally, the expression in square brackets is zero, so

$$\frac{dV^*(n)}{dn} = \delta(1-p)^{\mu(n)} \frac{\partial V^*(n-\mu(n))}{\partial n}. \quad (5)$$

We can generate explicit, or at least computable, expressions for $V^*(n)$ and $\partial V^*(n)/\partial n$ by first determining the value of a collection of such a size that the optimal strategy would be to test all remaining species in one final batch, then using this result to determine the value of a collection of species when the optimal strategy would be to test that number of species such that

---

[9] We have found it somewhat easier to assume differentiability and set derivatives equal to zero for the satisfaction of first-order conditions. An alternative approach would have been to have taken the difference $V^*(n, m) - V^*(n, m\text{-}1)$, in which case replacing $\ln(1\text{-}p)$ with p would have been exact. Under this assumption, however, we would have to make the approximation that this difference is approximately zero when m is chosen optimally. When n and m are both relatively large, either approximation seems appropriate. Empirically, n is large. If m were not large, the value of the marginal species must be negligible, as it would not be sampled until far into the future. So, in all cases in which we might entertain the hope that the value of the marginal species would be nonnegligible, the assumption of differentiability ought not to be troublesome.

the optimal strategy in the final period--if testing in the second-to-last period is unsuccessful--would be to test all species then remaining, and so forth.

Consider a decision taken to sample all remaining species simultaneously. Suppose that n_1 species remain to be sampled and the supplier is trying to decide how many, m, of them to test in the coming period (by our subsequent definition of terms, the optimal choice must be to set $m = n_1$, but we will derive the condition for this to be the case). If we assume that no more than one additional period of testing is to occur, the supplier's problem is to

$$\max_{m} \; R\left[1-(1-p)^{m}\right] - mc + \delta(1-p)^{m}\left(R\left[1-(1-p)^{n_1-m}\right] - (n_1-m)c\right) \tag{6}$$

Differentiating with respect to m, we have

$$p(1-p)^{m}R - c - \delta p(1-p)^{m}\left(R\left[1-(1-p)^{n_1-m}\right] - (n_1-m)c\right) - \delta(1-p)^{m}\left[pR(1-p)^{n_1-m} - c\right]$$

When n becomes small enough that it is optimal to evaluate all remaining species in a single final period (i. e., $m = n$), the above derivative must be nonnegative (value cannot be increased by evaluating fewer samples in the purportedly final period), so

$$p(1-p)^{\mu}R - c - \delta(1-p)^{\mu}(pR-c) \geq 0. \tag{7}$$

We will use the notation $\mu_1 = \mu(n_1) = n_1$ for that value of n such that (7) holds as an equality.

When $n \leq \mu_1$, the value of the collection is necessarily

$$V^*(n) = R\left[1-(1-p)^{n}\right] - nc, \tag{8}$$

and

$$\frac{\partial V^*(n)}{\partial n} = p(1-p)^n R - c; \tag{9}$$

but, when $n = \mu_1$, we can use (7) to solve for $(1-p)^{\mu_1}$ in (9) and state

$$\frac{\partial V^*(\mu_1)}{\partial n} = \frac{\delta(pR-c)c}{pR - \delta(pR-c)}. \tag{10}$$

We can also use (7) to find

$$\mu_1 = \frac{\ln[pR - \delta(pR-c)] - \ln c}{p}. \tag{11}$$

Thus (8), (10) and (11) serve to establish μ_1, $V^*(\mu_1)$, and $\partial V^*(\mu_1)/\partial n$.

The general first-order condition, (3) may be rearranged as

$$\mu = \frac{\ln\left(pR - \delta p V^*(n-\mu) - \delta \frac{\partial V^*(n-\mu)}{\partial n}\right) - \ln c}{p}. \tag{12}$$

Expressions (2), (5) and (12) form the bases of an algorithm to compute other values of $\mu(n)$, $V^*(n)$, and $\partial V^*(n)/\partial n$ in one period once we have calculated their values in a subsequent period (recall that we are proceeding by backward induction, so we compute later values of μ, V^*, etc., first).

Let us now define some additional notation. Let $\mu_2 = \mu(n_2)$ be the number of species sampled in the second-from-last period of potential testing (potential, as opposed to actual, testing, as testing may terminate upon the identification of a successful product), when n_2 is the total number of species remaining to be tested, and the optimal choice of batch size in the last period if testing is unsuccessful in the second-to-last period is μ_1. That is, $n_2 = \mu_1 + \mu_2$ is the largest number of remaining species for which the optimal testing strategy calls for the

completion of testing in no more than two periods. More generally, let $n_t = \sum_{i=1}^{t} \mu_i$ when t potential periods of testing remain and μ_i species are optimally tested in each period i (conditional on no successful tests being conducted before the final period). Then n_t is the largest number of species such that testing will be completed in no more than t periods.

Having computed μ_1 from (11), $V^*(\mu_1)$ from (8), and $\partial V^*(\mu_1)/\partial n$ from (9), we can now also compute $n_2 = \mu_1 + \mu_2$ using (11), $V^*(n_2)$ from (8), and $\partial V^*(n_2)/\partial n$ from (9). By iterating, we can generate n_t, $V^*(n_t)$, and $\partial V^*(n_t)/\partial n$ for arbitrarily large t. Note in particular that a simple form emerges for the value of the marginal species, $\partial V^*(n_t)/\partial n$. When there are n_1 species remaining to be tested,

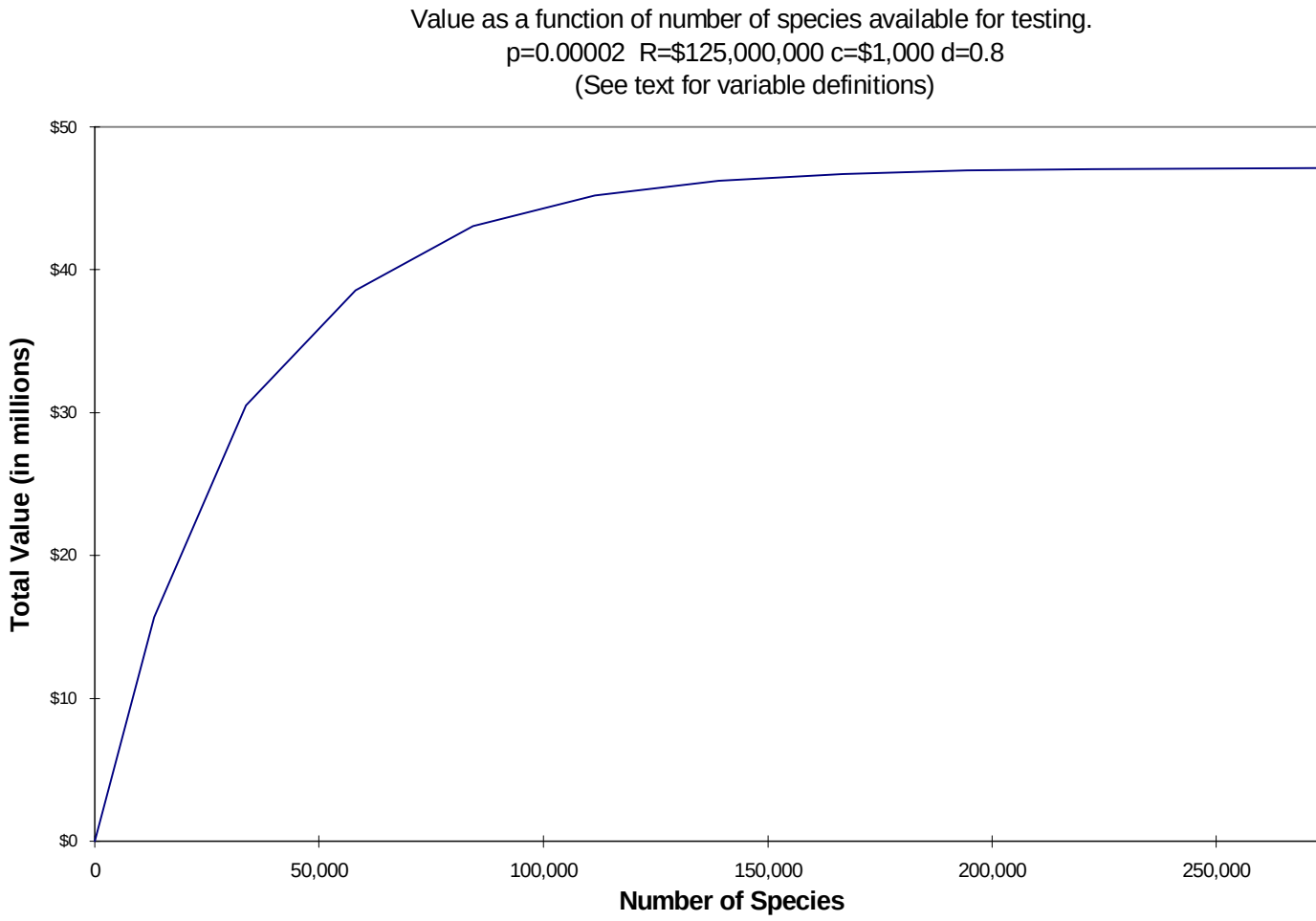
$$\frac{\partial V^*(n_1)}{\partial n} = \delta(1-p)^{n_1}(pR-c), \tag{14}$$

Using this result with (5) yields

$$\frac{\partial V^*(n_t)}{\partial n} = \delta^t(1-p)^{n_t}(pR-c). \tag{15}$$

Since we demonstrate that $V^*(n)$ is differentiable in the appendix, and monotonic and increasing at a decreasing rate by (9) and (5), we can connect the values of V^* defined at n_1, n_2, etc. with smooth curves so as to approximate the entire function. An example is illustrated in **Figure 1**.

Figure 1.


Value as a function of number of species available for testing.
p=0.00002 R=$125,000,000 c=$1,000 d=0.8
(See text for variable definitions)
Total Value (in millions)
$50
$40
$30
$20
$10
$0
0
50,000
100,000
150,000
200,000
250,000
Number of Species

II. AGGREGATION AND NUMERICAL EXAMPLES

In this section we extend the model we have described above to incorporate aggregation over a series of new product demands and calibrate the model with data from the pharmaceutical industry. Regrettably, data on biodiversity prospecting is very sparse. Companies guard proprietary information closely, and it can be very difficult to infer, for example, natural materials testing costs from overall firm R&D costs. Finally, we have, at best, very rough estimates of the probability with which species tested at random yield "hits" (and, of course, this probability differs over the application for which a product is sought).

Thus, we are going to proceed on the basis of some indirect inferences. In the next subsection we discuss issues of aggregation. Following that, we demonstrate how we can come up with a "best case" estimate of the value of the marginal species for use in new pharmaceutical product research. This "best case," or upper bound, estimate is derived by identifying those values of p and c for which the value of the marginal species is greatest, subject to other parameter values inferred from available data. In the final subsection following we demonstrate the sensitivity of this "best case" estimate to choices of other parameter values.

Aggregation

In the section above we derived the value of biodiversity in the search for a single new product. In the real world, the set of natural organisms might be evaluated in the search for a great many different products; Mendelsohn and Balick [1995] report that some 500 different tests are now being performed on natural organisms.

It also seems reasonable to suppose that different new product demands are identified over time. Twenty years ago, the AIDS virus, for example, was unknown. In addition, existing drugs[10] lose their efficacy as the disease organisms they are used against mutate and develop resistance. Finally, changing demography affects new product demand. An older, wealthier, and vainer population demands more treatments for skin wrinkles, obesity, and baldness.

Regrettably, there is no direct evidence on the rate at which new product demands are identified. We do, however, have two related pieces of information. The first we have already mentioned: that pharmaceutical researchers perform some 500 screens on natural materials. The second is that on average about 25 new drugs are approved for use every year (PMA, 1982-94; Simpson, Sedjo, and Reid, 1996).

We can, then, estimate the process of demand generation in the following way. Suppose that the expected number of new product demands identified in any given period is an independently, identically distributed random variable (determined, for example, by a Poisson process). Suppose further that the number of new product demands is also statistically independent of the probability of success in any single test of a species' efficacy for pharmaceutical use. Let Z be the expected number of new product demands identified per period of time. Let T be the expected number of products for which tests are undertaken in any period of time. Then we can write

---

[10] Taken literally, this would imply that we should model current demands for new products as a function of past successes. We have not included this consideration, in order to keep the model tractable.

$$T = Z + (1-p)^{\mu_t} Z + (1-p)^{\mu_t+\mu_{t-1}} Z + \ldots + (1-p)^{\mu_t+\mu_{t-1}+\ldots+\mu_2} Z$$
$$= Z\left(1 + \sum_{i=0}^{t-2} (1-p)^{\sum_{j=0}^{i} \mu_{t-j}}\right) \quad (16)$$

In words, the expected number of products for which natural organisms are being tested is equal to the expected number of new product demands identified in the current period, plus the expected number of new product demands identified in the previous period for which no suitable lead was identified among the μ_t organisms sampled in the first period of testing, plus the expected number of new product demands identified two periods previously for which no suitable lead was found among the $\mu_t + \mu_{t-1}$ organisms tested in the first two periods of testing, etc. Note that the summation in the exponent stops with the second-to-last, rather than the last, period of testing, as once all organisms have been tested, no more tests are possible.

Similarly, we can describe the expected number of "hits"--that is, the expected number of new product demands met--as

$$H = Z\left[1 - (1-p)^{\mu_t}\right] + Z(1-p)^{\mu_t}\left[1 - (1-p)^{\mu_{t-1}}\right] + Z(1-p)^{\mu_t+\mu_{t-1}}\left[1 - (1-p)^{\mu_{t-2}}\right]$$
$$+ \ldots + Z(1-p)^{\mu_t+\mu_{t-1}+\ldots+\mu_2}\left[1 - (1-p)^{\mu_1}\right]$$

$$= Z\left[1 - (1-p)^{n_t}\right]. \quad (17)$$

In words, the expected number of "hits" is the expected number of new product demands identified in the current period times the probability of making a "hit" among the μ_t species tested in the first period of testing, plus the expected number of new product demands identified in the previous period for which a suitable lead was not found in the first period, times the probability of making a "hit" among the next testing batch of μ_{t-1} species, etc., until all species have been tested. The second line of expression (17) is intuitively straightforward: in a steady

state, the expected number of "hits" in any given time period is the expected number of new product demands identified times the aggregate probability that a "hit" will be recorded in *any* period of testing.

Again, we cannot observe Z. Combining (16) and (17), however, we can relate the number of "hits" (on which we have data) to the number of tests (on which we also have some data). Doing so, we have that the expected number of tests is equal to

$$T = \frac{H\left(1+\sum_{i=0}^{t-2}(1-p)^{\sum_{j=0}^{i}\mu_{t-j}}\right)}{1-(1-p)^{n}}. \qquad (18)$$

Let us, for simplicity, restrict attention now to instances in which the number of species, n_t, is the largest number of species such that all species will be tested in no more than t periods.[11] Then the aggregate expected value of the marginal species, summed over all future expected new product demands and discounted to the present is

$$H\frac{\delta^{t}}{1-\delta}\frac{(1-p)^{n_t}(pR-c)}{1-(1-p)^{n_t}}. \qquad (19)$$

Constructing a "best case" scenario from available data

Let us quickly summarize the available information and the gaps. First, we adopt Mendelsohn and Balick's (1995) estimate of the value of a new pharmaceutical product of $125

[11] Inasmuch as casual empiricism suggests that testing proceeds in a relatively large number of stages (and this observation is, as we show below, consistent with the other evidence we observe), we do not believe that we are losing a great deal of accuracy by imposing this admittedly artificial restriction--especially inasmuch as we only do so in order to generate a scenario we can then subject to sensitivity analyses.

million[12] for our *R*. Mendelsohn and Balick also report that some 500 screens are conducted for new drugs at any given, time, and we adopt this figure for our *T*. *H*, the expected number of new product finds in any given time period, is also a random variable, but evidence on the number of new drug approvals from the U. S. Food and Drug Administration indicates that an average value of about 25 per year is reasonable (PMA, 1982-94). The number of species, *n*, available for testing will depend on the taxa to which research attention is devoted, and we will consider three alternatives below. Testing typically takes place over a number of years (it is often ten or more years from the time a natural material is first tested until a derived product is approved for use and then marketed). We thus assume a discount rate of twenty percent *per period*, which might be regarded as the result of an approximately four percent rate compounded annually over a five year testing period.[13]

Thus, we have data on *R*, *n*, δ, *T*, and *H*. Less reliable information is available concerning the probability of making a discovery in any single trial, *p*, and the cost of testing and evaluation, *c*. Rather than employing the very spotty data from which *p* and *c* might be inferred separately, we take an indirect approach.

We derive an *upper bound* on the value of the marginal species, subject to condition that the implied value we derive for *T* is (approximately) 500. We will derive the formal conditions for the procedure we are about to describe momentarily, but it may be helpful if we first give an

---

[12] Mendelsohn and Balick report a value of $125 million before, and $96 million after, corporate taxes. As they note some disagreement concerning this amount, all results are linear in *R*, and incorporating a bias toward higher values seems prudent, we proceed with the pretax figure.

[13] Since we are considering a testing period of this duration, we will also assume that the number of new product identified "per period" is 5 years per period x 25 new products per year = 125 new products per period.

overview. Given any values for R, δ, and n, the value of the marginal species will vary for different values of p and c. We begin by arbitrarily choosing the cost of testing, c and t, the maximum number of periods over which testing can occur. We next choose p so as to maximize the value of the marginal species–that is, we require that p satisfy a first-order condition set out in expression (20) below. Choosing p in this way generally induces a change in the number of species optimally sampled per period, however. This, in turn, means that testing may no longer be concluded in a maximum of t periods. Thus, we then change c so as to make the number of species remaining to be sampled in the t-th from final period equal to n. We then choose p again to maximize the value of the marginal species, then choose c again to set the total number of species remaining in the t-th from final period equal to n, and iterate in this fashion until we have arrived at values of p and c such that the value of the marginal species is maximized given that all n remaining species will be tested in no more than t periods. Finally, after going through this exercise, we compute the implied value of expression (18) (that is, the implied number of new products for which testing is being performed). If the implied value of T is greater than 500, it must be because we have chosen t to be too large (and c to be to great), and we repeat the entire procedure described above for a smaller value of t. Of course, if the implied value of T is less than 500, we must make the opposite correction.

It is important to be clear about what we are doing here. The probability with which success is achieved in a sample evaluation, p, is not a choice variable for any economic agent. We are not asking how profits are maximized with respect to p (if p were a choice variable, the obvious answer to this question would be to set $p = 1$). We are, rather, trying to come up with some estimate of the value of the marginal species which is consistent with what limited data we

do have. We could, in principle, choose any combination of *p* and *c* that satisfy the condition that the implied value of the number of tests, *T*, be close to the reported value of 500. By requiring in addition that *p* be chosen to maximize the value of the marginal species, we are attempting to build in an upward bias in our estimate. That is, if we find that even an upper bound on the value of the marginal species is relatively modest, we can be reasonably confident of policy prescriptions based on this finding.

We now turn to the formal conditions under which the value of the marginal species reaches a maximum. We begin by differentiating (19) with respect to *p*, imposing the first-order condition that the derivative be equal to zero, and evaluating (19) when this condition is imposed; while we do not present the derivations here, it is easy to verify that the second-order condition for an upper bound holds. It may not be immediately apparent that the value of the marginal species is differentiable in *p*, since some changes in *p* may also induce changes in *t*, the maximum number of periods over which testing can occur. Rather than fragmenting the exposition further by digressing into a proof of this result, we have relegated this proof to the appendix as well.

Differentiating (19) with respect to *p*, we have

$$H\frac{\delta^t}{1-\delta}\frac{\left[1-(1-p)^{n_t}\right]\left[R(1-p)^{n_t}-n_t(pR-c)(1-p)^{n_t-1}\right]-n_t(pR-c)(1-p)^{n_t}(1-p)^{n_t-1}}{\left[1-(1-p)^{n_t}\right]^2}$$

$$= H\frac{\delta^t}{1-\delta}\frac{R(1-p)^{n_t}\left[1-(1-p)^{n_t}\right]-n_t(pR-c)(1-p)^{n_t-1}}{\left[1-(1-p)^{n_t}\right]^2}.$$

Setting this expression equal to zero, we have

$$H\frac{\delta^t}{1-\delta}\frac{R}{n_t}(1-p)^{n_t+1} = H\frac{\delta^t}{1-\delta}\frac{(pR-c)(1-p)^{n_t}}{1-(1-p)^{n_t}}. \quad (20)$$

Note that the right side of expression (20) is the value of the marginal species, from expression (19), and so the left side of (20) must also be the value of the marginal species when *p* is chosen to maximize that value. While it is not generally possible to solve expression (20) analytically, it is easily solved numerically.

Table 1: "Best case" estimate of the value of the marginal species
(See text for definition of variables)

Number of species	Value of the marginal species	p	c	Maximum years of testing
250,000	$2,618.18	0.000010	$814	50
1,000,000	$0.000647	0.000011	$907	200
5,000,000	$7.409x10^{-38}	0.000011	$907	965

In Table 1 we report the results of three numerical exercises. In the first row of the table we assume that 250,000 species are available for testing. This is a rough estimate of the number of higher plants in the world (Wilson, 1992), these species may be particularly likely sources of new pharmaceutical products (Joffe and Thomas, 1988). In the second row we report results when there are 1,000,000 species available for testing. This figure might be regarded as a rough estimate of the number of described insect species (Wilson, 1992). The results reported in the third row are based on 5,000,000 species available for testing. This might be regarded as a very conservative estimate of the total number of species now extant in the world (Wilson, 1992).

In Table 1 we report results following the procedure we have described above. As can be seen there, a reasonable upper bound on the value of the marginal species when there are

250,000 species would be about $2,500. When there are 1,000,000 species, the value of a million-and-first species would be less than a tenth of a cent. With 5,000,000 species, the value of a five-million-and-first species would be on the order of $\$10^{-37}$.

Note that the values of p and c do not vary greatly between the "best case" scenarios with one and five million species. Heuristically, when there are a million species remaining to test, $\mu(n)$ is chosen to be a relatively small number. The optimal search strategy calls for batches of roughly the same size over the first many periods of testing. Since values in the distant future are heavily discounted, the optimal testing strategy does not differ by much from the strategy that would be pursued if there were an infinite number of species available for testing. Given that the testing strategy does not differ greatly from that which would be pursued if there were infinite options, it is not surprising that the value of the marginal species is also not greatly different from the value of the marginal species among an infinite set: i. e., zero.

Sensitivity of the estimates

The "best case" scenarios we have assembled above are just that: estimates of the value of the marginal species under favorably chosen circumstances. We can consider the sensitivity of these estimates to alternative choices of parameter values. In **Tables 2 and 3** we consider the value of the 250,000th and the one millionth species, respectively, for various values of p and c, fixing R and δ at $125 million and 0.8 as above.

Table 2: Sensitivity analysis of the value of a 250,000th species[a]
(Maximum years of testing in parentheses)

cost probability	$0.00	$200.00	$400.00	$600.00	$800.00	$1,000.00	$1,200.00	$1,400.00	$1,600.00
0.000001	$65,711.31 (b)	*	*	*	*	*	*	*	*
0.000002	$102,352.00 (b)	$6,918.53 (25)	*	*	*	*	*	*	*
0.000003	$119,567.65 (b)	$28,648.97 (15)	*	*	*	*	*	*	*
0.000004	$124,159.06 (b)	$41,086.31 (15)	$4,134.19 (45)	*	*	*	*	*	*
0.000005	$120,868.83 (b)	$46,175.73 (15)	$12,231.73 (30)	$1.45 (185)	*	*	*	*	*
0.000006	$112,959.14 (b)	$46,588.80 (15)	$17,258.43 (30)	$1,933.00 (60)	*	*	*	*	*
0.000007	$102,634.61 (b)	$44,226.74 (15)	$19,480.66 (25)	$4,955.01 (45)	$30.06 (130)	*	*	*	*
0.000008	$91,350.59 (b)	$40,373.22 (15)	$19,744.94 (25)	$7,012.34 (40)	$812.50 (70)	*	*	*	*
0.000009	$80,036.73 (b)	$35,865.68 (15)	$18,806.09 (25)	$8,005.54 (35)	$1,928.04 (60)	$30.23 (130)	*	*	*
0.000010	$69,258.35 (b)	$31,230.30 (15)	$17,203.56 (25)	$8,203.73 (35)	$2,754.36 (50)	$321.26 (85)	$0.00 (350)	*	*
0.000011	$59,332.15 (b)	$26,781.01 (15)	$15,299.85 (25)	$7,881.61 (35)	$3,196.67 (50)	$727.30 (70)	$17.79 (140)	*	*
0.000012	$50,408.50 (b)	$22,690.03 (15)	$13,326.98 (25)	$7,258.17 (35)	$3,323.65 (45)	$1,052.01 (65)	$122.07 (100)	$0.03 (260)	*
0.000013	$42,529.56 (b)	$19,038.07 (20)	$11,425.28 (25)	$6,487.92 (35)	$3,231.62 (45)	$1,243.54 (60)	$267.85 (85)	$8.50 (150)	0.00 (1100)
0.000014	$35,669.79 (b)	$15,848.36 (20)	$9,672.82 (25)	$5,673.22 (35)	$3,004.85 (45)	$1,314.64 (55)	$392.65 (80)	$45.11 (115)	0.007 (240)
0.000015	$29,763.81 (b)	$13,103.97 (20)	$8,106.59 (25)	$4,877.88 (35)	2,706.93 (45)	$1,296.34 (55)	$473.02 (70)	$96.82 (100)	$3.65 (160)
0.000016	$24,725.32 (b)	$10,772.69 (20)	$6,737.54 (25)	$4,138.63 (35)	$2,381.96 (45)	$1,219.51 (55)	$509.02 (70)	$143.83 (90)	$16.33 (130)
0.000017	$20,459.52 (b)	$8,812.54 (20)	$5,560.75 (25)	$3,473.92 (35)	$2,058.55 (45)	$1,109.28 (55)	$509.66 (65)	$176.56 (85)	$34.48 (115)
0.000018	$16,871.11 (b)	$7,178.12 (20)	$4,562.43 (25)	$2,890.34 (35)	$1,753.89 (45)	$984.00 (55)	$485.72 (65)	$193.44 (80)	$51.88 (105)
0.000019	$13,869.12 (b)	$5,825.14 (20)	$3,724.49 (30)	$2,387.13 (35)	$1,477.24 (45)	$856.16 (55)	$446.69 (65)	$196.80 (80)	$64.85 (100)
0.000020	$11,369.72 (b)	$4,711.51 (20)	$3,027.10 (30)	$1,959.24 (35)	$1,232.53 (45)	$733.61 (50)	$399.96 (65)	$190.16 (75)	$72.34 (95)

a: Upper diagonal elements are not defined; when $pR-c < 0$, the expected payoff is less than the cost of sampling, and no research takes place.

b. When $c = 0$, all species are sampled in a single batch.

Table 3: Sensitivity analysis of the value of a 1,000,000th species[a]
(Maximum years of testing in parentheses)

	$0.00	$200.00	$400.00	$600.00	$800.00	$1,000.00	$1,200.00	$1,400.00	$1,600.00
0.000001	$31,039.81	*	*	*	*	*	*	*	*
	(b)								
0.000002	$22,837.78	$203.13	*	*	*	*	*	*	*
	(b)	(70)							
0.000003	$12,602.29	$830.93	*	*	*	*	*	*	*
	(b)	(45)							
0.000004	$6,181.48	$595.51	$4.08	*	*	*	*	*	*
	(b)	(45)	(130)						
0.000005	$2,842.54	$308.14	$18.09	0.00	*	*	*	*	*
	(b)	(45)	(95)	(680)					
0.000006	$1,254.85	$139.51	$15.79	$0.06	*	*	*	*	*
	(b)	(45)	(85)	(190)					
0.000007	$538.57	$58.83	$9.11	$0.31	$0.00	*	*	*	*
	(b)	(45)	(80)	(145)	(465)				
0.000008	$226.43	$23.75	$4.39	$0.34	$0.00	*	*	*	*
	(b)	(50)	(80)	(130)	(245)				
0.000009	$93.71	$9.31	$1.93	$0.23	$0.00	$0.00	*	*	*
	(b)	(50)	(80)	(120)	(195)	(470)			
0.000010	$38.30	$3.58	$0.80	$0.12	$0.01	$0.00	$0.00	*	*
	(b)	(50)	(80)	(115)	(175)	(305)	(1335)		
0.000011	$15.50	$1.35	$0.32	$0.06	$0.00	$0.00	$0.00	*	*
	(b)	(55)	(80)	(115)	(165)	(250)	(505)		
0.000012	$6.22	$0.51	$0.12	$0.02	$0.00	$0.00	$0.00	$0.00	*
	(b)	(55)	(85)	(115)	(160)	(225)	(365)	(990)	
0.000013	$2.48	$0.19	$0.05	$0.01	$0.00	$0.00	$0.00	$0.00	$0.00
	(b)	(55)	(85)	(115)	(155)	(210)	(305)	(550)	(4340)
0.000014	$0.98	$0.07	$0.02	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
	(b)	(60)	(85)	(115)	(150)	(200)	(275)	(420)	(915)
0.000015	$0.39	$0.03	$0.01	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
	(b)	(60)	(90)	(120)	(150)	(195)	(255)	(360)	(600)
0.000016	$0.15	$0.01	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
	(b)	(65)	(90)	(120)	(150)	(190)	(245)	(330)	(480)
0.000017	$0.06	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
	(b)	(65)	(95)	(120)	(150)	(190)	(235)	(305)	(420)
0.000018	$0.02	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
	(b)	(65)	(95)	(120)	(150)	(190)	(230)	(290)	(380)
0.000019	$0.01	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
	(b)	(70)	(95)	(125)	(155)	(185)	(230)	(280)	(355)
0.000020	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00	$0.00
	(b)	(70)	(100)	(125)	(155)	(185)	(225)	(275)	(340)

a: Upper diagonal elements are not defined; when $pR-c < 0$, the expected payoff is less than the cost of sampling, and no research takes place.
b. When $c = 0$, all species are sampled in a single batch.

Not surprisingly, the value of the marginal species is uniformly lower when there are more species. If the universe over which testing can occur consists of a million or more species, the value of the marginal species is very low under any reasonable specification of probabilities and costs. Also not surprisingly, the value of the marginal species is higher the lower is the cost of testing. More interestingly, the value of the marginal species is a single-peaked function of the probability of success in any single trial. The value of the marginal species is necessarily zero at $p = c/R$. It increases in p over an interval, but soon begins to decline. When the probability of success is high, so is the probability of redundancy.

The greatest values in Tables 2 and 3 are, then, found in those instances in which the probability of success in any given trial is modest and the costs of sample evaluation are low. The numbers in parentheses below the value of the marginal species give the maximum number of years (i. e., five times the maximum number of periods, under the assumptions we laid out above) over which testing can occur. It can be seen that the larger estimates of the value of the marginal species are obtained only under assumptions which also imply what seem to be implausibly rapid rates of testing. Given that many organisms even in particularly attractive taxa (higher plants for example) have never been tested for their pharmaceutical potential, we might well suppose that any combination of parameters implying the completion of testing in less than, say, fifty years is empirically implausible.

III. REALISM AND RELEVANCE

We want to disabuse the reader of any lingering impression that we intend our model to be a realistic depiction of the process of new product research among natural organisms. What we have described above is a drastic and schematic simplification of the real process. We have

not constructed a realistic model. We do believe, however, that we have constructed a relevant one. We doubt that a more realistic model would reverse our conclusion that the marginal species is simply not of much value.

Let us defend this assertion by considering some possible objections to our approach. First, our assumption of identical Bernoulli trials is obviously unrealistic. In the real world, some things are better leads than others, even if more than one shows promise with respect to a particular application. Even when a usable product is developed for one purpose, the search continues for other products for the same purpose (consider, for example, what seems to be unending progress in generating new headache remedies). The fact remains, however, that the vast majority of materials tested for pharmaceutical uses are valueless. It is true that some things work better than others, but most things do not work at all. Assuming a binary outcome is not a bad approximation.

The assumption that the outcomes of different sample evaluations are statistically independent is more problematic than is the assumption that they are identically distributed. The pragmatic reason for adopting the independence assumption is that it dramatically simplifies the search problem: the order in which species are tested does not depend on the outcome of tests of earlier species. While we know that simplifying our mathematics does not simplify the real-world phenomena we seek to explain, there are several reasons for which we believe that assuming independence does not impart a downward bias to our results. First, the great majority of species have been subjected to little or no chemical or genetic analysis (Farnsworth, 1988). In fact, many biologists believe that the great majority of species have not yet been

discovered (Wilson, 1992). To regard the pharmaceutical payoff from these unexplored species as independently and identically distributed, at least *ex ante*, does not seem unreasonable.

The second reason for which the iid assumption seems innocuous--or that it does not bias our estimates *downward*, at least--is that correlation between the values of species would seem to imply that the value of the marginal species must be lower than one would assume under independence. Consider the limiting case. If the values of different species were perfectly correlated, it would be a case of "seen one, seen 'em all." That is, either every species would yield the useful commercial product sought, or none would. The incremental value of having more than one species available for testing would be zero.

The process of testing is much more complicated than we have depicted. Researchers do not take one batch of materials of a carefully chosen size, test them for activity against one set of diseases, and then, if search is unsuccessful, subject another batch of a carefully chosen size to the same battery of tests. Rather, in the early stages of research at least, researchers typically screen all the materials they happen to have available against all the diseases for which they are seeking treatments. Following these preliminary tests, those species that show promising activity are subjected to further tests, while less promising samples are discarded or deferred, and all the while new organisms are being acquired to subject to preliminary tests. Thus, our treatment of costs as a single, nonstochastic variable is extremely unrealistic.

We do, however, capture the essential feature of the research process. This essential feature, both analytically and empirically, is that testing occurs at a measured pace.[14] If the

---

[14] It is worth noting that we have assumed that the pace of testing is limited by an unwillingness to incur potentially redundant *variable* costs. It seems likely that an unwillingness to incur fixed costs that might afford a

expected value of the marginal species were great, pharmaceutical researchers would be in a greater hurry to test it. The very fact that most species have not been subjected to extensive chemical or genetic analysis–or even discovered–suggests that pharmaceutical researchers have, by and large, been able to meet new product needs by sampling from a relatively limited set of organisms.[15]

Another consideration we have not addressed in our model is technological change. If it may take decades to work through all species available for testing, there will almost certainly be technological progress that will affect testing strategy. It is important to think clearly about the forms such technological progress might take. Let us consider the possibilities. First, any technological progress that reduces the time and cost of testing will clearly increase the value of the marginal species. Technological progress may also, and at the same time, take other forms, however. Consider, for example, improvements in testing sensitivity. If different species contain the same or similar compounds in different concentrations, improvements in testing sensitivity will increase the probability with which a "hit" is recorded in testing any given species.[16] Improvements in synthetic chemistry may also have the effect of obviating new product development from organic sources entirely. While it is difficult to say which

---

greater rate of throughput, as well as an unwillingness to incur potentially redundant variable costs, can be identified as a factor in limiting the value of the marginal species in practice, and it does not especially matter which we model. For a model emphasizing fixed costs, see Simpson and Sedjo (1996).

[15] The search for new pharmaceutical products is not restricted solely to organic sources. In fact, we believe that one of the greatest sources of an upward bias in our estimate of values is the fact that we have ignored competition from inorganic sources.

[16] Of course, species that contain useful compounds in higher concentrations are more valuable, other things being equal, but it is often possible to increase concentrations by selective breeding or synthesize useful compounds by following the “natural blueprint.”

considerations dominate, one can reasonably argue that improvements in technology or more likely to decrease the value of the marginal species than to increase it.

In short, then, we do not offer a realistic description of the process of testing natural products for new pharmaceutical compounds. We can, however, make two arguments. First, if the process worked as we have described it, the value of the marginal species would be low. We have been justifying our second argument in this section: the simplifications and omissions or our model do not seem to induce a downward bias.[17]

IV. CONCLUSION

We have constructed a model in which pharmaceutical researchers engage in research over sequential batches of samples. We then bound the value of the "marginal species" by appeal to empirically reasonable rates of search. Even under what we argue are generous assumptions, we find low estimates of the value of the marginal species. We conclude that the economic incentives for conservation generated by biodiversity prospecting are negligible.[18]

This result has important policy implications. It does not necessarily mean that biological diversity is unimportant; biodiversity may be valuable for any of a number of other commercial,

---

[17] On this point we must note one additional consideration, however. The facts that new product demands are stochastic and that species extinctions are irreversible imply that there is an option value to delaying extinctions (see, e. g., Pindyck, 1991). In this respect, then, our estimates of value must be biased downward. We are very doubtful, however, that option value considerations are enough to induce the type of drastic revision necessary to reverse our general conclusion.

[18] Applying the commonly used model of island biogeography (MacArthur and Wilson, 1967; see also Simpson, Sedjo, and Reid, 1996 for an application in this context), we find that even the best of our best-case estimates would generate a value of about $5 per hectare in the most biologically diverse of threatened habitats. Incentives would be much lower in many other threatened regions.

ecological, esthetic, ethical, or even spiritual reasons. It may not even mean that the marginal social value of biodiversity as a source of new products is not great (since private researchers cannot appropriate the full surplus enjoyed by consumers when new products are introduced). What it does imply, however, is that advocates of biodiversity conservation should be looking to sources other than the generation of new commercial products to fund conservation.

APPENDIX

Modeling Simultaneous Discoveries

It is obvious that if a single researcher tests *m* species simultaneously, her expected payoff is given by (1). It is also straightforward to show that expression (1) is the aggregate expected payoff of *m* researchers when each successful researcher shares equally in the fixed prize, or has an equal chance of being the sole recipient of the prize. Suppose that if $k \leq m$ successful products are discovered, each successful researcher has a one-in-*k* chance of being afforded patent protection and earning a positive payoff. Now if there are *m* simultaneous sample evaluations undertaken, and each sample has a probability *p* of yielding a commercially successful product, the probability that exactly *k* successes will occur in *m* trials is given by the binomial probability density

$$\binom{m}{k} p^k (1-p)^{m-k},$$

where $\binom{m}{k} = \frac{m!}{k!(m-k)!}$, i. e., the number of ways in which *k* things can be chosen from a collection of size *m*.

Let *R* be the payoff resulting from the discovery of a successful commercial product. Let *c* be the cost of evaluating any given species for its potential as the source of a commercial product. Thus the expected payoff to any researcher from undertaking a single sample evaluation is

$$0\cdot(1-p)^{m} \;+\; pR(1-p)^{m-1} \;+\; \frac{pR}{2}\binom{m-1}{1}p(1-p)^{m-2} \;+\; \frac{pR}{3}\binom{m-1}{2}p^{2}(1-p)^{m-3} \;+\; \ldots$$
$$+\; \frac{pR}{m-1}\binom{m-1}{m-2}p^{m-2}(1-p) \;+\; \frac{pR}{m}p^{m-1} \;-\; c$$

or

$$\frac{R}{m}\left[\binom{m}{1}p(1-p)^{m-1} \;+\; \binom{m}{2}p^{2}(1-p)^{m-2} \;+\; \ldots. \;+\; \binom{m}{m-1}p^{m-1}(1-p) \;+\; p^{m}\right] \;-\; c$$

$$=\; \frac{R}{m}\left[1-(1-p)^{m}\right] \;-\; c\,. \qquad \text{(A1)}$$

If (1) describes the expected payoff to any single firm/sample evaluation, the total expected value to be received from *m* simultaneous sample evaluations is

$$R\left[1-(1-p)^{m}\right] \;-\; mc\,,$$

which is expression (1) from the text.

Differentiability and Second-Order Conditions

We have assumed that expressions are differentiable in *m*, *n*, and *p* without justifying these assumptions. In addition, we have assumed that second-order conditions for optimization are satisfied. In this section of the appendix we will prove differentiability and demonstrate the satisfaction of second-order conditions. Our general approach will be to establish results on the intervals $(0, n_1]$ and $[n_1, n_2]$ and then use expression (6) for induction on all other intervals.

Proposition 1:

$V^*(n)$ is differentiable for all *n*.

Proof: For $n \leq n_1$, we have

$$V^*(n) \;=\; R\left[1-(1-p)^n\right] - nc\,,\; n \le n_1. \tag{A2}$$

Obviously, V*(n) is differentiable with respect to n for $n \le n_1$. V*(n) will be differentiable for all positive n if we can apply expression (6) for all n. Recall that we assumed in deriving expression (6) that $\partial\mu(n)/\partial n$ exists.

Application of the implicit function theorem to (3) yields

$$\frac{\partial\mu(n)}{\partial n} \;=\; \frac{-\dfrac{\partial^2 V^*(n-m)}{\partial n\partial m}}{\dfrac{\partial^2 V^*(n-m)}{\partial m^2}} \text{ evaluated at } m = \mu(n), \text{ or}$$

$$\frac{\partial\mu(n)}{\partial n} \;=\; \frac{\delta(1-p)^{\mu}\left(p\dfrac{\partial V^*(n-\mu)}{\partial n}+\dfrac{\partial^2 V^*(n-\mu)}{\partial n^2}\right)}{-pc+\delta(1-p)^{\mu}\left(p\dfrac{\partial V^*(n-\mu)}{\partial n}+\dfrac{\partial^2 V^*(n-\mu)}{\partial n^2}\right)}. \tag{A3}$$

It is obvious, however, that $\partial\mu(n)/\partial n = 1$ on $(0, n_1]$, while, from (A3), $\dfrac{\partial\mu(n)}{\partial n} \;=\; \dfrac{\delta(1-p)^{\mu(n)}}{1+\delta(1-p)^{\mu(n)}}$ on $(n_1, n_2]$. As this quantity is necessarily less than one-half, we conclude that $\partial\mu(n)/\partial n$ does not exist at n_1. Heuristically, for all $n > n_1$, the optimal choice of μ is an interior solution; for $n < n_1$, the researcher's best choice is to adopt the corner solution of testing all remaining species simultaneously.

To find $\partial V^*(n)/\partial n$ for $n_1 < n < n_2$, (A3) justifies us in using (6), so

$$\delta(1-p)^{\mu(n)}\left[pR(1-p)^{n-\mu(n)} - c\right] \;=\; \delta pR(1-p)^n \;-\; (1-p)^{\mu(n)}c\,.$$

Now

$$\lim_{n \to n_1^+} \frac{\partial V^*(n)}{\partial n} = \delta p R(1-p)^{n_1} - \delta(1-p)^{\mu(n_1)} c = \delta(1-p)^{n_1}(pR-c)$$

$$= (1-p)^{n_1} pR - c = \lim_{n \to n_1^-} \frac{\partial V^*(n)}{\partial n},$$

where the second equality results from the definition $\mu_1 = n_1 = \mu(n_1)$ and the third from the application of (7).

We have now demonstrated that V*(n) is differentiable at least on the interval (0, n_2], and can establish that it is differentiable everywhere if we can show that the derivative $\partial\mu(n)/\partial n$ as derived in (A3) exists everywhere. Inspection of the denominator of (A3) reveals that $\partial\mu(n)/\partial n$ will exist so long as the second-order condition for profit maximization with respect to the choice of m is (strictly) satisfied. We complete the proof of our proposition, then, by proving the following lemma:

Lemma: For all $n > n_1$, (a) the second order condition is satisfied; and (b) $\partial\mu(n)/\partial n > 0$ (of course, $\partial\mu(n)/\partial n = 1$ for $0 < n \leq n_1$).

Proof: Differentiating (3) again, we have

$$\frac{\partial^2 V(n, \mu)}{\partial m^2} =$$

$$-p^2 R(1-p)^{\mu} + \delta p^2 (1-p)^{\mu} V^*(n-\mu) + 2\delta p(1-p)^{\mu} \frac{\partial V^*(n-\mu)}{\partial n} + \delta(1-p)^{\mu} \frac{\partial^2 V^*(n-\mu)}{\partial n^2}$$

or, using (3),

$$\frac{\partial^2 V(n, \mu)}{\partial m^2} = -pc + \delta(1-p)^{\mu}\left[p\frac{\partial V^*(n-\mu)}{\partial n} + \frac{\partial^2 V^*(n-\mu)}{\partial n^2}\right]. \qquad \text{(A4)}$$

The second-order condition for an optimal choice of m is that (A4) be less than zero.

Suppose that $0 < n \leq n_1$; then μ is identically n and, of course, $\partial\mu/\partial n$ is 1. $V^*(n)$ is $R[1-(1-p)^n] - nc$, $\partial V^*(n)/\partial n = pR(1-p)^n - c$, $\partial^2 V^*(n)/\partial n^2 = -p^2R(1-p)^n < 0$. Consider next the case $n_1 < n \leq n_2$. Using the results we have just noted,

$$p\frac{\partial V^*(n-\mu)}{\partial n} + \frac{\partial^2 V^*(n-\mu)}{\partial n^2} = -pc \qquad \text{(A5)}$$

for $n_1 < n \leq n_2$. Thus we see from (A4) that the second-order condition holds on the interval (n_1, n_2). From (A3) $\partial\mu(n)/\partial n > 0$ on (n_1, n_2).

We can prove the lemma and complete the proof of the proposition by establishing the induction step. Suppose that the second-order condition is satisfied when there are $n - \mu(n)$ species remaining untested. Then from (A3), $\mu(n)$ is differentiable. Thus, expression (6) from the text may be employed. We can rewrite (3) using (2) and (6) as

$$pV^*(n) + \frac{\partial V^*(n)}{\partial n} = pR - (\mu p - 1)c .$$

Differentiate totally with respect to n to find

$$-pc\frac{\partial\mu(n)}{\partial n} = p\frac{\partial V^*(n)}{\partial n} + \frac{\partial^2 V^*(n)}{\partial n^2}. \qquad \text{(A6)}$$

Thus, since we have now determined that $\partial\mu(n)/\partial n > 0$ on $[n_1, n_2]$, we can iterate again using (A6) to note that the second-order condition holds on $[n_2, n_3]$. We can use (A6) to rewrite (A3) as

$$\frac{\partial \mu(n)}{\partial n} = \frac{\delta(1-p)^{\mu}\frac{\partial \mu(n-\mu)}{\partial n}}{1 + \delta(1-p)^{\mu}\frac{\partial \mu(n-\mu)}{\partial n}}. \tag{A7}$$

Thus if μ is increasing in n on $[n_1, n_2]$, it must also be increasing in n on $[n_2, n_3]$.

We are now able to prove the general propositions that the second-order condition holds and that μ is increasing in n by induction. Suppose that μ is increasing in n on $[n_{t-1}, n_t]$ for some t. Then (A7) establishes that μ is also increasing in n on $[n_t, n_{t+1}]$. Moreover, the assumption that μ is increasing in n on $[n_{t-1}, n_t]$ is, by (A6), sufficient to establish that $p\frac{\partial V^*(n)}{\partial n} + \frac{\partial^2 V^*(n)}{\partial n^2} < 0$, and thus from (A4) that the second-order condition is satisfied on $[n_t, n_{t+1}]$.

Since these results are also sufficient to show that μ is differentiable for all $n > n_1$, we can apply expression (6) from the text and the results above to prove proposition 1.

Proposition 2: The value of the marginal species is differentiable in p for all $n > n_1$.

Proof: Differentiate expression (5) totally with respect to p:

$$\begin{aligned}\frac{\partial^2 V^*(n)}{\partial n \partial p} &= \delta(1-p)^{\mu}\left[\frac{\partial^2 V^*(n-\mu)}{\partial n \partial p} - \frac{\partial^2 V^*(n-\mu)}{\partial n^2}\frac{\partial \mu}{\partial p}\right] \\ &- \left[\delta\mu(1-p)^{\mu-1} + \delta p(1-p)^{\mu}\frac{\partial \mu}{\partial p}\right]\frac{\partial V^*(n-\mu)}{\partial n}\end{aligned} \tag{A8}$$

Differentiate the first-order condition, (3), totally with respect to p and rearrange to find

$$\frac{\partial \mu}{\partial p} = (1-p)^{\mu} x$$

$$\frac{\frac{\mu}{1-p}\left[pR - \delta p V^*(n-\mu) - \delta \frac{\partial V^*(n-\mu)}{\partial n}\right] - R + \delta p\left[\frac{\partial V^*(n-\mu)}{\partial p} + \frac{\partial^2 V^*(n-\mu)}{\partial n \partial p}\right]}{-pc + \partial(1-p)^{\mu}\left[p\frac{\partial V^*(n-\mu)}{\partial n} + \frac{\partial^2 V^*(n-\mu)}{\partial n^2}\right]}$$

Using the first-order condition, (3), again, we have

$$\cdot \frac{\partial \mu}{\partial p} = \frac{\frac{\mu c}{1-p} - R(1-p)^{\mu} + \delta(1-p)^{\mu} p\left[\frac{\partial V^*(n-\mu)}{\partial p} + \frac{\partial^2 V^*(n-\mu)}{\partial n \partial p}\right]}{-pc + \partial(1-p)^{\mu}\left[p\frac{\partial V^*(n-\mu)}{\partial n} + \frac{\partial^2 V^*(n-\mu)}{\partial n^2}\right]}. \qquad \text{(A9)}$$

We have determined in the proof of Proposition 1 that the denominator of the above expression is always negative.

Recall from above that, for $0 < n \leq n_1$, $V^*(n) = R[1- (1-p)^n] - nc$ and $\partial V^*(n)/\partial n = pR(1-p)^n - c$. Differentiating with respect to p, we have

$$\frac{\partial V^*(n)}{\partial p} = nR(1-p)^{n-1} \text{ and}$$

$$\frac{\partial^2 V^*(n)}{\partial n \partial p} = R(1-p)^n - nR(1-p)^{n-1} = [1-(n+1)p]R(1-p)^{n-1}, \; 0 < n < n_1.$$

Thus, $\partial\mu/\partial p$ exists on the interval (n_1, n_2). Iteration using (A8) then suffices to show that the value of the marginal species is differentiable in p for all n.

REFERENCES

Artuso, Anthony M. 1994. *An Economic and Policy Analysis of Biochemical Prospecting*, unpublished Ph.D. dissertation, Department of Economics, Cornell University.

Aylward, Bruce A. 1993. "A Case Study of Pharmaceutical Prospecting," Part III of Bruce A. Aylward, Jaime Echeverria, Liza Fendt, and Edward B. Barbier, *The Economic Value of Species Information and its Role in Biodiversity Conservation: Case Studies of Costa Rica's National Biodiversity Institute and Pharmaceutical Prospecting*, Report to the Swedish International Development Authority (London: London Environmental Economics Centre).

Benhabib, Jess, and Clive Bull. 1983. "Job Search: The Choice of Intensity," *Journal of Political Economy*, vol. 91, no. 5 (October), pp. 747-764.

DiMasi, Joseph A., Ronald W. Hansen, Henry G. Grabowski, and Louis Lasagna. 1991. "Cost of Innovation in the Pharmaceutical Industry," *Journal of Health Economics*, 10, pp. 107-42.

Ehrlich, Paul R., and Anne H. Ehrlich. 1981. *Extinction: The Causes and Consequences of the Disappearance of Species* (New York: Random House).

Eisner, Thomas. 1989. "Prospecting for Nature's Chemical Riches," *Issues in Science and Technology*, 6, pp. 31-34.

Farnsworth, Norman R. 1988. "Screening Plants for New Medicines," in E. O. Wilson, ed., *Biodiversity* (Washington: National Academy Press), pp. 83-97.

Joffe, Steen, and Robert Thomas. 1989. "Phytochemicals: A Renewable Global Resource," *AgBiotech News and Information*, 1, pp. 697-700.

MacArthur, R., and E. O. Wilson. 1967. *The Theory of Island Biogeography* (Princeton, N.J.: Princeton University Press).

Mendelsohn, Robert, and Michael J. Balick. 1995. "The Value of Undiscovered Pharmaceuticals in Tropical Forests," *Economic Botany*, 49 (June), pp. 223-228.

Morgan, Peter, and Richard Manning. 1985. "Optimal Search," *Econometrica*, vol. 53, no. 4 (July), pp. 923-944.

Myers, Norman. 1986. *Tackling Mass Extinction of Species: A Great Creative Challenge*, XXVIth Horace M. Albright Lecture in Conservation, University of California at Berkeley.

Myers, Norman. 1988. "Threatened Biotas: 'Hot Spots' in Tropical Forests," *The Environmentalist*, 8, pp. 187-208.

Myers, Norman. 1990. "The Biodiversity Challenge: Expanded Hot-Spots Analysis," *The Environmentalist*, 10, pp. 243-256.

Office of Technology Assessment (OTA). 1993. *Pharmaceutical R&D: Costs, Risks, and Rewards*, OTA-H-523 (Washington, D.C.: U. S. Government Printing Office).

Pearce, David, and Seema Puroshothamon. 1992. "Preserving Biological Diversity: The Economic Value of Pharmaceutical Plants," Discussion Paper 92-27, CSERGE London.

Pharmaceutical Manufacturers Association (PMA). 1982-94. "New Drug Approvals," Annual report published by PMA, Washington, D.C.

Pindyck, Robert S. 1991. "Irreversibility, Uncertainty, and Investment," *Journal of Economic Literature*, 29 (September), pp. 1110-1148.

Polasky, Steven, and Andrew Solow. 1995. "On the Value of a Collection of Species," *Journal of Environmental Economics and Management*, 29 (November), pp. 298-303.

Reid, Walter V., Sarah A. Laird, Carrie A. Meyer, Rodrigo Gamez, Ana Sittenfeld, Daniel H. Janzen, Michael A. Gollin, and Calestous Juma. 1993. "A New Lease on Life," Chapter 1 in *Biodiversity Prospecting: Using Genetic Resources for Sustainable Development* edited by Walter V. Reid, Sarah A. Laird, Carrie A. Meyer, Rodrigo Gamez, Ana Sittenfeld, Daniel H. Janzen, Michael A. Gollin, and Calestous Juma (Washington: World Resources Institute).

Roberts, Leslie. 1992. "Chemical Prospecting: Hope for Vanishing Ecosystems?" *Science* (22 May), pp. 1142-1143.

Rubin, Steven M., and Stanwood C. Fish. 1994. "Biodiversity Prospecting: Using Innovative Contractual Provisions to Foster Ethnobotanical Knowledge, Technology, and Conservation," *Colorado Journal of International Environmental Law and Policy*, 5, pp. 23-58.

Simberloff, Daniel S. 1986. "Are We on the Verge of a Mass Extinction in Tropical Rain Forests," in D. K. Elliott, ed., *Dynamics of Extinction* (Wiley: New York), pp. 165-180.

Simpson, R. David, and Roger A. Sedjo. 1996. "Investments in Biodiversity Prospecting and Incentives for Conservation," Discussion Paper 96-14, Resources for the Future, Washington, D.C.

Simpson, R. David, Roger A. Sedjo, and John W. Reid. 1996. "Valuing Biodiversity for Use in Pharmaceutical Research," *Journal of Political Economy*, vol. 104, no. 1 (February), pp. 163-184.

Stigler, George J. 1961. "The Economics of Information," *Journal of Political Economy*, vol. 69, no. 2 (June), pp. 213-225.

Wilson, Edward O. 1992. *The Diversity of Life* (Cambridge: Belknap).