Working Paper Series

WORKING PAPER NO. 819, REV.2

VALUING RESEARCH LEADS: BIOPROSPECTING AND THE CONSERVATION OF GENETIC RESOURCES

by

Gordon C. Rausser and Arthur A. Small

DEPARTMENT OF AGRICULTURE AND
RESOURCE ECONOMICS AND POLICY

BERKELEY

CALIFORNIA AGRICULTURAL EXPERIMENT STATION

University of California

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS AND POLICY
DIVISION OF AGRICULTURE AND NATURAL RESOURCES
UNIVERSITY OF CALIFORNIA AT BERKELEY

WORKING PAPER NO. 819, REV.2

VALUING RESEARCH LEADS: BIOPROSPECTING AND THE CONSERVATION OF GENETIC RESOURCES

by

Gordon C. Rausser and Arthur A. Small

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
April 1999

Valuing Research Leads: Bioprospecting and the Conservation of Genetic Resources

Gordon C. Rausser[1] Arthur A. Small[2]

University of California at Berkeley

May 7, 1998

Revised April 1, 1999

[1]Dean, College of Natural Resources, Robert Gordon Sproul Distinguished Professor, and Member, Giannini Foundation of Agricultural Economics, University of California at Berkeley.

[2]Doctoral Candidate, Department of Agricultural and Resource Economics, University of California at Berkeley. The authors acknowledge generous financial support from the US National Science Foundation–Environmental Protection Agency Partnership for Environmental Research, and from the University of California Systemwide Biotechnology Research and Education Program. The authors are grateful for comments from Peter Berck, Ignacio Chapela, Andrew Dabalen, Anthony Fisher, Peter Lupu, Astrid Scholz, Michael Ward, Brian Wright, Joshua Zivin, and an anonymous referee. The work also benefited from feedback by seminar participants at Berkeley, Columbia, Georgia Institute of Technology, Georgia State, Iowa State, Rutgers, the University of Washington, and the 1997 NSF-EPA "Valuation and Environmental Policy Workshop."

Abstract

Bioprospecting has been touted as a source of finance for biodiversity conservation. Recent work has suggested that the bioprospecting value of the "marginal unit" of genetic resources is likely to be vanishingly small, creating essentially no conservation incentive. This result is shown to flow specifically from a stylized description of the research process as one of brute-force testing, unaided by an organizing scientific framework. Scientific models channel research effort towards leads for which the expected productivity of discoveries is highest. Leads of unusual promise then command information rents, associated with their role in reducing the costs of search. When genetic materials are abundant, information rents are virtually unaffected by increases in the profitability of product discovery, and decline as technology improvements lower search costs. Numerical simulation results suggest that, under plausible conditions, the bioprospecting value of certain genetic resources could be large enough to support market-based conservation of biodiversity.

Biodiversity prospecting, the search for valuable compounds from wild organisms, has been touted as a potential source of finance for biodiversity conservation. An open question is whether, or under what conditions, revenues from bioprospecting could be large enough to offset the opportunity costs that host institutions may have to incur in order to preserve genetic resources. A recent paper by Simpson, Sedjo and Reid (1996) (SSR) argues that the returns to holding genetic resource assets are unlikely to be large enough to create significant conservation incentives. The claim is based on a model of the research process in which firms sample without replacement from a large set of research leads, incurring a fixed cost per draw. Two features of the process are key. The first is *uncertainty*: it is unknown, prior to testing, whether a given lead is good or bad, whether it will or will not generate a discovery. The second essential feature concerns the potential for *redundancy* amongst the leads. A lead that enables an innovation may not do so uniquely, just as caffeine can be found both in coffee and in tea. Assuming that a given discovery can be made only once, multiple copies of a "potential discovery" are, in this sense, superfluous. The authors pose the question: supposing that each lead carries a fixed probability of yielding a breakthrough, how much would a private firm be willing to pay to prevent the collection of leads from becoming slightly smaller? In other words, what is the value of the marginal research opportunity, in this R&D process? Formal analysis confirms what intuition suggests: if the original collection is sufficiently large, then one additional lead is likely either to be infertile (if the per-test probability of success is very low) or redundant (if the success probability is sufficiently high). Given that the number of species in the world is very large indeed, the expected research return to the "marginal species" is likely to be vanishingly small. It will, then, exert no genuine incentive towards conservation, in the context of a market for genetic resources. Extensions to cases in which discoveries vary in quality, or in which success rates covary according to an average degree of genetic distance (Polasky and Solow 1995), generate somewhat higher values, but do not alter the substance of this conclusion.

The result is intriguing. We argue in this paper, however, that it flows specifically from a stylized description of the research process as one of brute-force testing, unaided by an organizing scientific framework. Given the progress of biological and ecological science, the

realism of this assumption is suspect. While exceptions can be noted, it is a powerfully general rule that no one ever searches for anything by examining large collections of objects in random order. The essence of efficient search is the identification of clues that allow the universe of potential leads to be narrowed down. Expensive tests are then conducted, ideally, only on that handful of prospects that show special promise. In research targeting the development of innovations, the clues that identify these prospects are provided by scientific models—maps of the world that highlight areas where the productivity of research effort is likely to be highest. The ability of models to point out rich veins of ore explains exactly why applied researchers acquaint themselves with basic theory. It is substantially from this ability that the human capital of an applied research scientist derives its value. Brute-force search—the sequential testing of large numbers of leads in no particular order—is by contrast a nearly cost-maximizing approach to discovery. It is deployed, if at all, only as a back-stop technology, when all possibilities for directed search have been exhausted.

Theory, in other words, partitions collections of potential research projects into categories of greater and lesser expected quality. A useful model thereby converts leads from commodities into differentiated products. The market for research opportunities shifts from a purely competitive to a monopolistically competitive structure. This change has profound implications for the valuation of leads and, therefore, for the incentives surrounding their production or conservation. In particular, leads of unusual promise can be shown to command information rents associated with their role in reducing search costs. A private firm will be willing to pay a premium for access to a bin of promising leads in order to avoid the cost of picking through the mass of low-quality prospects, even if the latter are available at bargain prices.

The point is clarified by pursuing the metaphor of applied research as an extractive industry. Imagine that technologies are mined from a landscape marbled, unevenly, with patentable discoveries. Innovations could, in principle, be uncovered by digging without pattern or purpose, just as oil could be sought by sinking wells at randomly-chosen sites. Provided the domain of exploration is large enough, undistinguished areas will not, for this

purpose, be a scarce resource: one site is about as good as any other.

In practice, of course, the right to drill for oil can command large rents. These rents depend partly on the price of oil, and are partly a function of the prior information available about the stake. If oil has been found nearby, or if the geological structures are promising, then this intelligence raises prospectors' willingness to pay. Indeed, if the market for extraction services is competitive, then the rent accruing to the high-quality site should reflect, approximately, the amount of search and processing costs a firm avoids by focusing there, rather than on an inferior "average" site.

When genetic materials are viewed as candidate sources for new products, this image can be applied to bioprospecting. In this endeavor, leads of unusual promise are distinguished with the aid of scientific information gleaned from fields such as ecology and taxonomy. Researchers can, and do, draw on rich bases of publicly-available data describing the location and properties of plants, animals and microbes, their evolutionary history, their survival and reproductive strategies. These data, when filtered through a model that makes sense of them, can serve to tag those creatures most likely to display economically valuable characteristics. Just as a catalog helps a library patron to focus quickly on those few volumes that are most likely to contain information she desires, so can an ecological model parse the living world into categories suggesting potential use.

Through product differentiation, scientific understanding generates information rents: if a particular lead is believed to show promise as an aid to a lucrative research discovery, a rational investigator will be willing to pay an access fee. This principle is fundamental to understanding how genetic resources will be, or should be, valued in the market place. In particular, it is central to an analysis that identifies conditions under which bioprospecting creates effective market-based financial incentives for biodiversity conservation. Rents accrue to the owners of leads as they absorb part of the knowledge spillovers generated by publicly available science.

To structure this theme formally, we present a model of applied research in which leads

are differentiated by their expected quality, and then tested sequentially. A complete characterization is developed of the relationship between the costs and benefits of research, the degree of product differentiation, and the value of options on research leads. The value of a lead is shown to be driven almost entirely by its probability, relative to alternative options, of generating success, and by the costs of search. Surprisingly, the value is largely insensitive to the size of the payoff from a successful discovery. This payoff does, of course, affect the overall value of the search project to the researcher. However, so long as the jackpot is large enough to make the project a viable undertaking, additional increases have almost no effect on the rents accruing to research leads. Furthermore, a technology shock that lowers search costs actually decreases these values: since the amount of a lead's rent reflects what it contributes, in expectation, to reducing the costs of research, a drop in the price of brute-force testing shrinks the benefits of guidance.

Four sections comprise the balance of the paper. Section 1 examines evidence on the pivotal question of whether bioprospecting benefits from scientific prior information. The biological literature and the behavior of industry actors are shown to support the claim that bioprospecting takes generous advantage of a significant base of useful scientific guidance. Section 2 presents the model, analyzing the effect of product differentiation on the value of research leads. Section 3 addresses the empirical question of whether information rents associated with genetic resources might be large enough to play a significant role as a source of finance for biodiversity conservation. Numerical results indicate that, under plausible assumptions concerning the demand for discoveries and costs of testing in the pharmaceutical industry, high-quality leads could command rents large enough to tilt land use decisions toward conservation. A concluding section addresses the implications of the argument for the design of intellectual property rights and other institutions governing the market in genetic materials. Mathematical proofs, and an analysis of the model's robustness, are presented as appendices.

1 Availability of Prior Information for Bioprospecting: Evidence From Biology and Economic Behavior

A market in genetic resources will only appear if the expected benefits of conservation exceed the opportunity costs of holding these assets. While costs can be measured more-or-less directly (as a function of forgone logging profits and the like), benefits must be estimated using indirect methods. Theoretical approaches to valuing natural intellectual capital have been investigated by Brown and Goldstein (1984) and Heal (1995). The R&D option value of biodiversity has also been addressed at the microeconomic level of single species by Weisbrod (1964), and Arrow and Fisher (1974). Nonetheless, few empirical estimates have appeared in the formal economics literature.

Studies that find low values for genetic resources *in situ* share in common an assumption that the materials are, in their role as inputs to the innovation process, essentially fungible. Every "unit" of biodiversity is viewed as making an equal marginal contribution to the success of the bioprospecting enterprise; one species is about as good as any other. We claim that this assumption matters. Do bioprospectors actually treat different types of genetic materials as perfect substitutes for one another? This factual issue is most appropriately resolved by specialists in biology and in related technology industries.

In the published writings of these specialists, the economic utility of organisms appears as a topic of lively discussion, attracting contributions from ecologists, systematicists, evolutionary biologists, and ethnobotanists. Their findings occupy several journals, including such titles as *Economic Botany, Journal of Ethnopharmacology,* and *Journal of Natural Products.* One active line of inquiry concerns how the evolutionary histories and survival strategies of species relate to the kinds of useful compounds they may produce. Plants, for example, appear in general to be more chemically creative than animals. Unable to respond to predatory attacks by fleeing, plants have instead developed a wide array of complex chemical means of defense. These strategies are especially well-developed in deserts, rainforests and coral reefs—zones of "biological warfare" where biotic and environmental stresses are particularly

acute (Myers 1997). Such observations can be valuable: a fern that prospers in a region richly populated with insects may have evolved pest-control solutions that would impress the research staff at an agrochemicals firm. Other useful environmental indicators can be quite taxonomically specific. Frogs, lacking a hard defensive shell, make themselves unappetizing by secreting toxins prolifically. Abbot Laboratories reportedly has a project underway to synthesize a painkiller based on a toxin from the threatened Ecuadorian poison dart frog (Sawhill 1998). In another case, an anti-coagulant approved for human trials was developed after screening venom from only seventy species of snakes (Pollack 1992). Natural-products researchers are moving to organize known ecological and taxonomic indicators of utility into comprehensive models, in order to provide systematic guidance to the drug discovery process (Dreyfuss and Chapela 1994). Pursuing this approach, Novartis Pharmaceuticals is reportedly spending several million dollars in the attempt to identify observable factors that correlate with biochemical "creativity" of micro-organism, i.e. their fertility as sources of molecules that could form the basis for new drugs (Moeller 1996).

Clues based on general biological observations and formal models can be supplemented by indications gleaned from the historic record of product discovery. Many natural-products scientists hold that classes of organisms which have proven useful in the past are relatively likely to provide similar compounds for related uses (Phillips and Meilleur 1998). The rosy periwinkle, a flowering plant native to East Africa, has already produced two anti-cancer drugs, vincristine and vinblastine (Reid, Laird, Meyer, Gamez, Sittenfeld, Janzen, Gollin and Juma 1993). The Himalayan yew tree is now the primary source of another anti-cancer drug, taxol, that was originally sourced from its North American cousin, the Pacific yew. Indeed, some firms base their entire product discovery programs on leveraging the experience of traditional healers concerning the therapeutic properties of plants used in herbal medicine. Such historical clues need not be based only on taxonomic similarity. Certain geographic regions have particular prominence as sources of valuable research leads, functioning as ecological Silicon Valleys. Traditionally cultivated varieties of domesticated crop plants and their wild relatives, which are especially important sources of agronomically useful genes, are for example not distributed uniformly on the globe. Instead, they are concentrated in

relatively compact "centers of diversity," usually mountainous areas where species adapt locally to a fragmented set of agroclimatic conditions (Chrispeels and Sadava 1994).

The suggestion that bioprospecting is a targeted, rather than haphazard, activity is confirmed by the behavior of industry actors. If every organism was in fact an equally promising source of any given compound, then we might expect to see search projects spread more-or-less evenly across the globe. The industry's interest in the hot springs at Yellowstone National Park provides a singular counter-example. The US National Park Service now grants about twenty-five permits annually for the right to collect biological samples from this ecosystem (Pennisi 1998), which is distinctive both for its link to PCR, and as home to the most genetically diverse community of bacteria known to science (Pace 1997). This degree of localized interest is apparently inconsistent with an hypothesis that natural product potential follows a geographically uniform distribution (Polasky and Solow 1995).

If the hypothesis were nonetheless to be entertained, any observed clustering would have to be explained by differentials in the cost of collecting and evaluating samples from different regions. The location of collection site would be explained largely by variables such as distance to the firms' R&D labs, ease of access, and so on. In addition, we would expect the market for samples to be almost purely competitive, with all rents accruing to the innovating firms. These outcome do not appear to obtain in practice. Bioprospecting firms have shown a durable willingness to mount expensive collection efforts in remote but ecologically distinctive locations, including underseas. To gain admittance to these areas, they are now commonly compelled to spend a great many staff hours negotiating access contracts with host-country governments and their agents. Under these agreements, firms routinely forfeit shares in the profits arising from their product discovery efforts. The firms, one presumes, believe they are getting something in return for these resources. Their behavior cannot readily be explained by a model that assigns equal promise to every species and region.

2 A Model for Valuing Research Leads

In sum, the claim that bioprospecting is always conducted in brute-force fashion in an impoverished informational context does not appear to withstand close scrutiny. It remains to be shown whether this simplification is benign, or leads to a fundamental misspecification of the value of genetic resources in their role as research options.

The theory of valuing search options seems to have received little attention in the economics literature, particularly in the context of technology development. Search theories developed within economics (Lippman and McCall, eds 1979) and elsewhere (Stone 1975 DeGroot 1970) have focused on the identification of optimal selection and stopping rules (Weitzman 1979), and on how these rules are affected by changes in various parameters (the probability distributions of rewards, agents' appetites for risk, time preferences, Bayesian updating, etc.). Further, much of this work is based on an assumption that per-trial rewards are identically distributed, or that the number of search opportunities is unlimited (Kohn and Shavell 1974). Interactions between the size of the search space and the degree of quality differentiation amongst search opportunities have apparently not been investigated.

To address the question, we analyze a simple model of the applied research process. Research opportunities appear as lottery tickets, each characterized by a price (the cost of testing), a probability of winning, and a jackpot that is paid in the case of a lucky draw. The investigator rank tickets by quality, then draws them one at a time until she either wins the prize, or exhausts her supply. The focus is on the investigator's willingness to pay for an additional ticket of a given quality. Formally, the model generalizes that of SSR by allowing success probabilities to differ across leads. The model also echoes the representation by Evenson and Kislev (1976) of applied research as an act of repeated sampling from a known distribution, the parameters of which can be changed through basic research.

2.1 The Model

An investigator conducts a search for a compound that will make possible the development of a new product. Research is conducted by testing "leads." There are a large number N of leads available, each of which embodies some potential to yield a discovery. Data are available that describe leads in several dimensions: to each lead $n = 1, ..., N$, there is associated a vector $x_n \in X$, where X is a space of characteristics. The investigator filters these data through a model, represented by a function $P : X \to (0, 1)$, to estimate the probability that a test of a lead with a given set of characteristics will generate an enabling discovery. Both the theory and the data are available freely, as public goods. (The investigator's cost of using these resources to formulate her beliefs are considered to be small, relative to the costs of testing leads and the potential benefits of a discovery.) The collection of leads is then treated as a set of N Bernoulli trials, in which the n^{th} lead carries probability $p_n = P(x_n)$ of yielding a success, or "hit." Without loss of generality, leads are labeled in order of decreasing hit probability, so that $p_1 \geq ... \geq p_N$. The hit probabilities of different leads are assumed to be statistically independent. It is not assumed that the collection of leads contains exactly one discovery; it is also possible that the discovery is present in multiple copies, or not at all.

The investigator tests leads sequentially, at a cost c per test, where c is a positive constant. When a test is successful, a payoff R is realized. It is assumed that a discovery need be made only once; multiple hits are redundant. The sole behavioral assumption is that the investigator selects the order in which she tests leads so as to maximize the expected payoff to the project.

Given our assumptions about the incentives facing the bioprospecting firm, its behavior is characterized by the following result, a simple application in the theory of search.

Proposition 1 *An optimal search program involves testing, at each stage, a lead with maximal hit probability amongst those not yet examined. Search terminates either when a hit is reached, or when no leads remain for which testing promises a non-negative return in*

expectation.

In other words, optimal search involves checking the most promising leads first—an intuitive and well-known result (Weitzman 1979).

Proof: all proofs are included in an Appendix.

In light of the result, the probability ordering $(p_1, p_2, ..., p_N)$ determines the sequence in which leads are examined.[1] The stopping rule implies that leads for which $p_n < c/R$ are never tested under any conditions, and have no effect on search behavior or payoffs. Without violence to our results, therefore, we assume in what follows that $p_n \geq c/R$ for all n.[2]

Given optimizing behavior, a value function can be derived that gives the expected payoff of the search at each stage, conditional on results at previous stages. Let V_n denote the *ex post* expected value of continuing the search, after $n-1$ leads have been tested unsuccessfully. Applying Bellman's Principle of Optimality, this continuation value is characterized by the recursive relationship

$$V_n = p_n R + (1 - p_n)V_{n+1} - c \ , \ n = 1, ..., N \tag{1}$$

where $V_{N+1} \equiv 0$. The equation can be interpreted as follows. With probability p_n, the n^{th} test is successful, a payoff R is realized, and search terminates; with probability $1 - p_n$, the test is a failure, and search proceeds to the $n + 1^{st}$ lead. The "consolation prize" in case of failure is the opportunity to continue the search with the $n + 1^{st}$ lead, the value of which is, by definition, V_{n+1}. In either event, a cost c is incurred for the test. Solving, the expected payoff to the search at its outset is given by

$$V_1 = \sum_{n=1}^{N} a_n(p_n R - c) \tag{2}$$

[1] More exactly, the probability ordering determines the search queue up to a permutation of leads with equal hit probabilities. Such permutations have no effect on lead values, or on aggregate returns to research (Corollary 3).

[2] Alternatively, the stopping rule can be interpreted as defining endogenously the effective number of available leads. If $\bar{N}$ denotes the total number of potential leads, then $N \leq \bar{N}$ can be defined as the largest integer such that $p_N \geq c/R$.

where $a_n \equiv \prod_{i=1}^{n-1}(1-p_i)$ is the probability that the search is carried to the n^{th} stage, i.e. the probability of failure in each of the first $n-1$ tests; and $p_nR - c$ is the expected return to a test of the n^{th} lead. Here, a_np_n is the probability that search terminates with a success at the n^{th} lead. If we treat the event "the project fails" as an $N+1^{st}$ lead, then the vector $< a_1p_1, ..., a_Np_N, a_{N+1} >$ forms a probability distribution over the set $\{1, ..., N, N+1\}$, with associated payoffs $< R-c, R-2c, ..., R-Nc, -Nc >$. Specifically, this is a truncated nonhomogeneous geometric distribution, where $\sum_{i=1}^{n} a_ip_i = 1 - a_{n+1}$ gives the cumulative probability distribution. This relation can be used to express V_1 as a difference between expected benefits and costs:

$$V_1 = \sum_{n=1}^{N} a_np_nR - \sum_{n=1}^{N} a_nc = (1-a_{N+1})R - Tc \tag{3}$$

The first term denotes the *ex ante* expected benefit of the project: the probability $1-a_{N+1}$ of a successful conclusion, times the payoff R in case a hit is scored. The second term denotes the *ex ante* expected cost, expressed as the expected number $T = \sum a_n$ of trials carried out, times the per-trial cost c.

Using equation (2), an expression can be derived for the expected incremental contribution of the n^{th} lead to the overall value of the search. Define V_{-n} as the expected value of the search process, for the case in which the n^{th} lead is skipped:

$$V_{-n} = \sum_{i=1}^{n-1} a_i(p_iR - c) + a_nV_{n+1} \tag{4}$$

The incremental value of the n^{th} lead, denoted v_n, is defined as the difference between these two terms:

$$v_n \equiv V_1 - V_{-n} = a_n[p_n(R - V_{n+1}) - c] \; . \tag{5}$$

The value v_n can be viewed as the maximum a firm will be willing to pay at the start of a search project for a call option on the n^{th} lead. The option insures that the lead will be available if it should be needed, i.e. if all tests of more promising leads end in failure. The formula is interpreted as follows: With probability a_n, the first $n-1$ tests are unsuccessful and search proceeds to lead n, which is tested at cost c. With probability p_n, this test is successful, a reward R is realized, and search terminates. The effective payoff in this case

is, however, net of the continuation value V_{n+1} that would have applied if the search had instead been forced to skip the n^{th} lead. That is, since multiple discoveries are redundant, a success at the n^{th} stage destroys the value associated with the opportunity to continue searching.

Because a portfolio of leads may contain multiple copies of a given potential discovery, the expected value of the project does not equal the sum of the values of the leads. Indeed, $\sum v_n = V_1 - \sum a_n p_n V_{n+1}$; the sum of lead values equals the expected value of the project, less the expected "redundancy cost" that a discovery at one lead imposes on the others. The latter term is analogous to the social welfare cost associated with redundant R&D programs in patent races.

2.2 The Value of Research Leads: Scarcity Rents and Information Rents

If a lead might merely duplicate another one in the portfolio, how does it add value to a research project? For leads that are unusually promising, early, high-probability options contribute more than the others to the chance of an eventually successful outcome for the project. As repeated failures push investigators to pick through lower-grade "ore," it becomes increasingly unlikely that a hit will ever be scored. More importantly, the opportunity to focus initial research effort on the high quality leads increases the chance that a discovery will be made *early* in the process. In case of an early discovery, the costs of continued search are avoided. If high quality leads are removed from the menu of search options, the shift to low-quality sources implies an increase in the expected number of trials-to-discovery.

Intuition behind the argument can be strengthened by examining equation (3). Deleting a lead from the search queue raises a_{N+1}, the probability that the project ends in an expensive failure. In addition, removal of a high-quality lead can increase the expected length (hence, cost) of the search, an effect represented by an increase in T. Both effects are more pronounced for leads early in the search queue. This holds because these leads, if available,

are more likely to be tested. Removal of leads toward the back of the queue has an effect on search payoffs only in cases in which all early tests end in failure.

In sum, when search procedures are optimized to incorporate useful prior information, high-probability leads command information rents associated with their unusual contribution to the chance of success, and to the avoidance of search costs. These information rents apply in addition to any scarcity rents resulting from a limit on the total number of leads. This distinction between scarcity and information rents is formalized in the following proposition, which characterizes completely the relationship between the costs and potential benefits of search, the quality of available information, and the value of differentiated search opportunities. Here, a lead n is referred to as *marginal* if its hit probability is equal to that of the lowest-quality viable alternative, i.e. if $p_n = p_N = min\{p_i | p_i R - c \geq 0, i = 1, ..., N\}$.

Proposition 2 *Let $\{p_n\}_{n=1}^{N}$ be the sequence of hit probabilities on a collection of leads, indexed in order of decreasing probability. Let the incremental value v_n of the n^{th} lead be defined as in (5) above. Then v_n can be decomposed into components $v_n = \hat{v}_n + v_N$, where*

$$\hat{v}_n \equiv \frac{a_N}{1 - p_n} (p_n - p_N) R + \left[\sum_{i=n+1}^{N-1} \frac{a_i}{1 - p_n} (p_n - p_i) \right] c \tag{6}$$

and where $v_N = a_N (p_N R - c)$ is the value of a marginal lead. We refer to these components, respectively, as the information rent and the scarcity rent of lead n.

A lead's scarcity rent can be interpreted as the expected amount it would contribute to the value of the project if it were undistinguished from the mass of other leads and was, therefore, a perfect substitute for any other marginal lead, *ex ante*. Scarcity rents can be positive if the project is constrained by a technical bound on the number of feasible research opportunities—that is, if N is finite, and if marginal leads carry a positive expected return ($p_N R - c > 0$), so that random screening is, in expectation, profitable.

The information rent captures the degree to which a distinguishing prior increases a lead's expected incremental value. The expression for $\hat{v}_n$ can be interpreted with the aid

of a thought experiment. Suppose that a researcher who is seeking a treatment for muscle spasms learns that people of a particular location, when thus afflicted, boil the roots of a certain plant to make a tea that is locally renowned for its soothing powers. Although she had planned to visit the area eventually to conduct random screening of samples, she had previously thought it unpromising for her purpose, assigning it the lowest probability, p_N. On hearing the news, she raises her expectation that the region will provide what she seeks, increasing her prior to $p_n > p_N$. She also rearranges her search queue so that the area now occupies the prominent n^{th} position on her itinerary, where $n < N$. What effect does the change in her prior entail for her expected return to the research project?

There are two effects. First, there is an increase in the expected probability that she will find what she seeks before exhausting all her leads. The amount of the consequent rise in expected benefits is represented by the first term in the expression for $\hat{v}_n$, the probability $a_N/1-p_n$ that no other lead contains the discovery, times the increase in the expected benefit of testing this remaining lead. Second, if this n^{th} test is successful, then she will have avoided the cost of visiting at least some of the leads that now occupy positions $n+1, ..., N-1$. The second term in the expression for $\hat{v}_n$ represents the drop in her expected costs of search.

As the discussion makes clear, the magnitude of the information rent associated with a given lead depends not only (or even, primarily) on the lead's own hit probability, but also on how this value compares with those of other leads. This is because a lead's hit probability determines not only the chance of a successful test, but also its position in the search queue and, therefore, the probability that it will ever be tested at all.

2.3 Resource Abundance, Research Costs and Payoffs, and Resource Value

Under what conditions are information rents large enough to carry significant weight in land use and conservation decisions? In particular, can genetic resources have significant value, in their role as bioprospecting search opportunities, even when genetic materials are

abundant? The characterization of information and scarcity rents allows several general results to be proved that bear on these questions. It is shown that when research payoffs are large, and resources are abundant, conservation incentives are driven almost entirely by the magnitude of search costs and the quality of available information. They depend only weakly on the potential returns to the research project. In these cases, the size of the payoff will be important for a firm undertaking the research, but will have little effect on that firm's willingness to pay for access to resources. It has, therefore, little bearing on conservation decisions.

Inspection of the formulae in Proposition 2 shows that scarcity and information rents have non-negative values ($v_N \geq 0$, and $\hat{v}_n \geq 0$ for all n), and that marginal leads command zero information rents. More importantly, it confirms the claim advanced earlier that leads of unusual promise have strictly greater value than do their less promising neighbors.

Corollary 3 *For all $m, n = 1, ..., N$, if $p_n = p_m$, then $v_n = v_m$, and if $p_n > p_m$, then $v_n > v_m$. Hence the sequence $\{v_n\}$ is monotone decreasing. In particular, information rents are everywhere zero if and only if all priors are equal.*

The intuition is straightforward: since a lead's incremental value is defined as the difference in expected returns with the lead and without it, the removal of any two leads with equal probability will yield the same effect on project returns.

The comparative static effects on lead values of changes in the parameters describing search benefits and costs can also be examined. To facilitate the discussion, let $B_{n_q} \equiv \frac{a_N}{1-p_n}(p_n - p_N)R$, the first, "benefit-increasing" term in equation (6). Note that the payoff R from a success enters into the expression for v_n only through v_N and B_n and that both these terms are linear in a_N, the probability of reaching the last lead in the search queue. Since $p_n \geq c/R$ for all n, a_N is bounded above by $(1 - c/R)^{N-1}$, which becomes small as N grows large. Hence, the size of project rewards has only a limited effect on lead values, a claim formalized in the following result.

Corollary 4 *The effect on lead values of a change in research payoffs R is given by $\partial v_n/\partial R = a_{N+1}\left(\frac{p_n}{1-p_n}\right)$. Hence, $\partial v_n/\partial R$ is strictly positive, and is independent of R and c. For $p_n \leq 1/2$, it is bounded above by a_{N+1}, the probability of project failure.*

Interpreting the expression for $\partial v_n/\partial R$, an increase in project rewards raises a lead's value only in proportion to the probability that the lead contains the discovery uniquely (i.e. in proportion to p_n, the probability of success at lead n, times $a_{N+1}/1-p_n$, the probability that tests at all other leads would result in failure). In particular, when the expected probability of an eventual discovery is high ($a_{N+1} \approx 0$), lead values are largely insensitive to changes in the project payoffs. This result may seem counter-intuitive; one might expect that large increases in potential project rewards would generate substantial increases in the value of a chance to realize those rewards. To understand the result, it is important to distinguish sharply between the project's overall value V_1, and the incremental contribution to that value of any one lead. By equation (3), the overall value of the project increases linearly with R. The share of this surplus accruing to the leads is affected, however, by the potential presence of duplicate discoveries in the portfolio. In an extreme case, the space of viable options may be sufficiently large that eventual discovery becomes virtually certain. In this case, the firm assurance of realizing the discovery is unaffected by the loss of any single lead. If the project's payoff increases (e.g., if demand for a target drug rises), the associated increase in surplus accrues to the firm's research capacity, not to the leads. The precise value of a lucrative discovery has, then, little bearing on incentives to conserve the resource.

Note, however, that the "cost-reducing" term $\sum_{i=n+1}^{N-1} \frac{a_i}{1-p_n}(p_n - p_i)\,c$ in (6) remains positive even if there are many viable leads. Indeed, this term is increasing in N: as the haystack grows, information on the whereabouts of the needle becomes increasingly valuable. This observation is fundamental to the relationship between search costs and lead values.

Corollary 5 *The effect on lead values of a change in search costs c is given by $\partial v_n/\partial c = \sum_{i=n+1}^{N-1} \frac{a_i}{1-p_n}(p_n - p_i) - a_N$. This rate of change is independent of research costs c and payoffs R. Furthermore, an increase in search costs makes the n^{th} lead more valuable if and only*

if $p_n > 1/T_n$, *where* $T_n = \sum_{i=n}^{N} a_i/a_n$ *is the expected number of trials conducted after* $n-1$ *failures.*

Roughly speaking, the condition in the last sentence applies whenever a lead is "much better" than those in a large pool of alternatives. It suggests a second surprising result. Intuitively, we might expect that the value of a search opportunity to decrease with any rise in the unit cost of search effort. Such a change, after all, decreases the expected return to any trial. However, for the relatively promising leads, there is a second, counter-veiling effect. As shown in Proposition 2, a large fraction of the value of these leads is associated with their potential help in avoiding the costs of resorting to low-quality alternative sources. Conversely, as unit search costs decline, the value of this "competitive advantage" is eroded: brute-force search becomes an increasingly viable alternative to directed search. Hence, an improvement in search technology increases (weakly) the value of the lowest-quality leads, but can reduce the value of high-quality leads.[3]

As viable search opportunities become increasingly abundant, these counter-intuitive features of the value function become dominant. To show this, consider the case in which there are a relatively small number of leads known to be specially promising, and a larger number that share an "average background" probability of making a discovery at a lead sampled at random from a large pool of potential sources. Concretely, let $M \ll N$ be given, and suppose that $p_M = p_{M+1} = ... = p_N = p$, where p is a constant such that $pR - c \geq 0$; and suppose that $p_n > p$ for all $n = 1, ..., M-1$. We refer to leads $1, ..., M-1$ as "promising," while leads $M, ..., N$ are called "marginal," as above. The following proposition characterizes the value of genetic materials as marginal leads become abundant, i.e. as N grows without bound.

[3]Some readers have questioned whether combinatorial chemistry, the massively parallel engineering of drug candidates, constitutes a brute-force approach to discovery that renders bioprospecting obsolete. In practice, combinatorial chemistry rarely involves a truly random search through the set of all possible molecules; the set of configurations is simply too vast. Instead, the approach usually starts from a compound that has shown some activity against a target disease, and looks for variations that generate superior responses. Rather than supplanting bioprospecting, the technique's ability to refine the activity of naturally produced compounds arguably complements the latter's potential role in the production of pharmaceutical discoveries.

Proposition 6 *Let leads* $1, ..., M-1$ *be promising, and leads* $M, ..., N$ *be marginal, as defined above. Then for* $n = 1, ..., M-1$, *the incremental value* v_n *has a strictly positive lower bound, given by the relation*

$$v_n \geq \left[\sum_{i=n+1}^{M-1} \frac{a_i}{1-p_n} (p_n - p_i) + \frac{a_M}{1-p_n} \cdot \frac{p_n - p}{p} \right] c \quad .$$

Further, this relation becomes an equality in the limit as N *goes to infinity.*

Accordingly, a lead that is worth testing at all and is more promising than at least some of its "competitors" commands a rent that is strictly positive and bounded away from zero. Note that the bound does not depend on the degree of resource abundance N. Nor does it depend on the potential reward R, beyond a threshold level of project viability. The bound can be interpreted as the *ex ante* reduction in search costs associated with the opportunity to test the n^{th} lead before moving on to the pool of less promising sources, a pool that includes an infinite number of leads with hit probability p. Thus it depends only on search costs and prior information. (Again,we assume away the trivial case in which one or more leads contain the discovery with certainty.) The final claim of the proposition, concerning the limiting behavior as N goes to infinity, implies striking results about the value of genetic resources under conditions of abundance.

Corollary 7 *Let leads* $1, ..., M-1$ *be promising, and leads* $M, ..., N$ *be marginal, as defined above. Suppose that the payoff from a successful discovery is large enough to make random sampling of marginal leads profitable in expectation* ($R \geq c/p$). *Then as leads become abundant (in the limit as* $N \to \infty$*), the following hold:*

1. *Marginal leads have zero value.*
2. *An increase in the potential profitability of product discovery has no effect on the incremental value of any lead.*
3. *A technology improvement that lowers search costs induces a drop in the incremental value of every promising lead.*

If search opportunities are in effectively unlimited supply, scarcity rents are negligible. In this case, the value of a lead is, to a very good approximation, entirely a function of the quality of information associated with it. The composite good (material plus information) enhances the profitability of the project not so much by creating success, as by aiding the avoidance of failure.

2.4 The Effects of Basic Scientific Research

Improvements in ecological, taxonomic, and ethnobotanical knowledge can change researchers' beliefs about lead values. The next two proposition clarify the magnitude of this effect. In particular, it is shown that lead values respond to changes in scientific information in a continuous manner. The avoidance of search costs provides, therefore, a robust criterion for the valuation of research leads.

Proposition 8 *The value of a lead is a piecewise linear, continuous, increasing, and weakly convex function of its own hit probability. Furthermore, on the interval* $p_{n-1} > p_n > p_{n+1}$, *the elasticity of* v_n *with respect to* p_n *is given by*

$$\frac{\partial v_n}{\partial p_n} \cdot \frac{p_n}{v_n} = 1 + \frac{a_n c}{v_n} \quad .$$

Interpreting the elasticity formula, the twin effects of value-enhancement and cost-reduction are identified. A pair of analogous results characterizes the effect of changes in hit probabilities for other leads.

Proposition 9 *The value of a lead is a piecewise linear, decreasing function of the hit probability of every other lead. For* $m \neq n$, v_n *is continuous in* p_m *except where* $p_m = p_n$. *Furthermore, the following hold:*

(i) For $m = 1, ..., n-1$, *on the intervals* $p_{m-1} > p_m > p_{m+1}$, *the elasticity of* v_n *with respect to* p_m *is given by*

$$\frac{\partial v_n}{\partial p_m} \cdot \frac{p_m}{v_n} = -\frac{p_m}{1 - p_m} \quad .$$

(ii) For $m = n+1, ..., N$, *on the interval* $p_{m-1} > p_m > p_{m+1}$, *the elasticity of* v_n *with respect to* p_m *is given by*

$$\frac{\partial v_n}{\partial p_m} \cdot \frac{p_m}{v_n} = -\frac{p_m}{1-p_m}\left[1 + \left(a_n - p_n \sum_{i=n+1}^{m} \frac{a_i}{1-p_n}\right)\frac{c}{v_n}\right] .$$

The form of the equation in point (i) reflects the fact that, conditioned on the superiority of lead m over lead n, a further increase in p_m reduces the chance that lead n will ever be tested. (Interestingly, this effect depends only on the magnitude of p_m, and not on the position of m in the list 1,..., $n-1$.) For $m > n$, an increase in p_m reduces the perceived probability that, if the n^{th} lead provides a success, then it would have done so uniquely. It also reduces the expected cost savings associated with the opportunity to test the n^{th} lead before the m^{th}.

The claims in Propositions 8 and 9 are conditioned on constraints in the ordering of the search queue. It turns out that the qualitative results do not depend on this restriction. The effect of improved scientific information on the value of research leads can be summarized in a simple form.

Corollary 10 *An increase in the hit probability of a lead induces a more-than-proportionate increase in the value of that lead, and a decrease in the value of every other lead.*

This corollary implies in turn that the bound described in Proposition 6 applies to all probability orderings (and not just in the special case for which hit probabilities are constant for all but a few superior sites).

Corollary 11 *Let* $\{p_n\}_{n=1}^{N}$ *define a probability ordering for the bioprospecting problem, and let* p *be a constant with* $1 > p \geq p_N$. *Let* $M = M(p) = min\{m|p_m \leq p\}$. *Then for* $n = 1, ..., M-1$, *the incremental value* v_n *of lead* n *has a strictly positive lower bound given by*

$$v_n \geq \left[\sum_{i=n+1}^{M-1} \frac{a_i}{1-p_n}(p_n - p_i) + \frac{a_M}{1-p_n} \cdot \frac{p_n - p}{p}\right] c.$$

Further, in the limit as $N \to \infty$, the conclusions of Corollary 7 hold.

The inequality provides a bound on value that depends only on the first few elements $p_1, p_2, ..., p_{M-1}$ of the probability ordering, and on the search cost parameter c. In particular, it does not depend on knowledge of the entire probability ordering, nor on the exact size of the payoff R from a successful discovery. Apparently, this formula shows promise as the basis for empirical valuation studies.

3 A Numerical Illustration

Could bioprospecting information rents be large enough, in practice, to affect land use decisions? Uncertainties about relevant parameter values preclude definitive empirical estimation of these values. Insight on this qualitative question can be gained, however, by revisiting SSR's numerical calculations on the value of genetic resources as inputs to drug-discovery. The exercise has two purposes. First, it demonstrates how the framework of Section 2 could be used, in the context of suitable scientific information, as the basis for assigning economic value to genetic resources. Second, it offers support for the claim that, under a range of plausible parameter choices, such information could generate rents large enough to create significant conservation incentives.

A set of N leads is partitioned into K classes of varying quality. For $n = 1, ..., N$, let $k(n)$ denote the index of the class containing lead n. Let e_k be a measure of the quality of leads in the k^{th} class, for $k = 1, ..., K$. Hit rates are proportional to lead quality: $p_n = \bar{p} \cdot e_{k(n)}$, where $\bar{p}$ is constant. Given financial parameters c and R, and assuming an optimal program of search, the contribution v_n of the n^{th} lead to a single search project is given by equation (5). A number λ projects are carried out per year, and future costs and benefits are discounted using a constant interest rate r. The net present bioprospecting value of the n^{th} lead is then given by

$$\sum_{t=0}^{\infty} \lambda(1+r)^{-t} v_n = \frac{\lambda v_n}{r}. \tag{7}$$

To compute these values numerically, decisions must be made about how the domain of search should be characterized, how it should be parsed into individual "leads," and how the quality of leads should be measured. These questions are the subject of extensive debate in the natural-products literature. A thorough airing of the issues involved would go well beyond the scope of this paper. We proceed by considering a particular case, addressing briefly the possible sensitivity of our conclusions to these choices.

Following SSR, the search domain corresponds to a group of eighteen ecologically distinctive ecosystems listed in Table 1, described in Myers (1988) and Myers (1990) as "biodiversity hot spots." Single leads correspond to land parcels of a uniform area (1000 ha, or 10 km^2), where an investigator can collect biological samples. The quality of a parcel (or "site") as a potential source of new drugs is proxied by the degree of endemism amongst higher plant species. Specifically, for ecosystems $k = 1, ..., 18$, the quality index e_k is defined as the density of endemic higher plant species in that ecosystem, measured as the average number of species per hectare.[4]

Other parameters on the drug discovery process are based those developed by SSR, who draw on data from DiMasi, Hansen, Grabowski and Lasagna (1991) and Office of Technology Assessment (1993). The probability that a test of a site in ecosystem k will yield a discovery is $\bar{p} \cdot e_k$, where $\bar{p}$ =1.2 x 10^{-5}. The probability that a project will terminate unsuccessfully, exhausting the available leads without yielding a discovery, is $\prod_{k=1}^{18}(1 - \bar{p} \cdot e_k)^{N_k} = 63\%$. Here, N_k denotes the number of sites in ecosystem k (Table 1). To achieve a realistic yield of ten new natural-source drugs per year therefore requires that projects be launched at a rate of $\lambda = 26$ per year. Each successful discovery generates a return of $R = \$450{,}000{,}000$. Firms discount future costs and benefits at r =10% per year. In the baseline case, costs are set at c = \$485 per test, at which level the formula (7) assigns negligible value to marginal sites in the least promising ecosystem. (Alternative values of the cost parameter were also examined.) Calculations of marginal bioprospecting values per hectare in each of the eighteen ecosystems

[4] A species is "endemic" to a region if it is found nowhere else. It must be emphasized that endemism is not necessarily a realistic way to measure the quality of an area for drug-discovery purposes. Endemism is used here, rather, as a stand-in for more sophisticated indicators.

are presented in Table 1 and in Figure 1. The corresponding values generated in the SSR study are presented to facilitate comparison.[5]

Figure 1 shows the relationship between ecosystem size, hit rates, and value. Information rents can be several orders of magnitude larger than scarcity rents, and can be substantial even when scarcity rents are negligible. The values associated with the highest-quality sites—on the order of $9,000/hectare in our simulation—can be large enough to motivate conservation activities. Sensitivity analysis shows that these qualitative results are robust with respect to the selected parameter values. In particular, a ten-fold increase in the cost per test raises the net present value of the highest quality sites to approximately $11,000/hectare.

The computations reported in Table 1 are not designed to give rigorous empirical estimates of bioprospecting values for the geographic areas described. The derivation of such estimates will depend on the development of scientific and ethnobotanical bases for assigning priors to search opportunities (better P functions and data sources), over a range of potential natural products. In particular, we do not make the empirical claim that density of endemic species is a good indicator of chemical creativity or site quality. The intended take-home message is that *whatever* basis is used to differentiate leads by quality, those leads that are marked as having distinctive promise will carry distinctive value, under plausible conditions of market demand. There is no reason to believe *a priori* that the bioprospecting values of all areas are so small as to play no role in land management decisions.

4 Summary and Conclusions

A sequential-search model of biodiversity prospecting has been analyzed. When search procedures are optimized to take advantage of useful prior information, high-probability leads command information rents associated with their contribution to the chance of success

[5] In SSR, tests are carried out on individual species. Values per unit area are derived from a (constant) value per species using a widely-accepted log-linear species-area relationship. This step is not necessary in our approach because the site itself is taken as the unit of testing. Hence, values per unit area are calculated directly.

and, more importantly, to the avoidance of search costs. These rents apply in addition to any scarcity rents arising from a bound on the total number of leads available for testing. The magnitude of information rents depends on the degree to which ecological and taxonomic knowledge turns leads into "differentiated products," creating a monopolistically competitive market in research opportunities. Rents for high-quality leads can be significant even when the aggregate supply of genetic materials is large. When biological resources are abundant, an increase in the potential profitability of product discovery has virtually no effect on the value of any lead, while technology improvements that lower search costs induce (weakly) a decline in the value of every lead.. Numerical results suggest that bioprospecting information rents could, in some cases, be large enough to finance meaningful biodiversity conservation. These conclusions stand in opposition to those advanced in an earlier analysis by Simpson et al. (1996), which argued that biodiversity prospecting holds out no hope as a meaningful source of finance for biodiversity conservation. That result, it was shown, holds only in the degenerate case in which no prior information is available to sort leads into categories of differentiated quality.

Given the amount of attention that has been attached to this policy question, it is important to be clear about why the two studies draw such divergent implications. Nothing in the present analysis alters the conclusion on truly "marginal" species. Species that exhibit insignificant commercial promise in every known application will generate *de minimus* financial incentives for conservation. However, the observation that *there exist* species that will not be conserved through a market-based scheme does not imply that *all* species will be similarly condemned. The pivotal issue concerns whether every species—more generally, every unit of biodiversity—can be considered equally "marginal." This is an empirical question that cannot be resolved through *a priori* theorizing. Yet neither the biological literature nor the behavior of industry actors lends clear support to the contention that natural products prospecting proceeds on the basis of randomized search.

Viewed as inputs to the innovation process, genetic materials have the potential to become genuine resources in the context of a sufficiently rich set of complementary knowledge assets.

The effective functioning of a market in genetic resources depends on these knowledge assets just as much as, if not more than, it relies on a sound system of intellectual property rights and a robust capital market. This suggests that attempts to estimate the value of genetic resources should focus attention on how researchers form and update their beliefs. It also suggests that the institutions regulating bioprospecting, including systems of intellectual property rights, should reward the provision of helpful prior information, as well as the conservation of the base biological material.

The analysis has examined the marginal social value of research options. A significant question for further work concerns how firms' willingness to pay for research leads would be affected by competition in the race to patent biologically-based innovations. As we have seen, willingness to pay may be depressed by the potential for duplicative discovery—a given lead may not be unique as the source of an innovation. As a firm considers how much to bid, it must attend to the associated "redundancy cost" that the new lead imposes on the firm's existing portfolio of research options. The firm does not, however, internalize costs imposed on competitors. As the number of firms competing in a given patent race grows large, redundancy costs become less important; the market price of genetic resources should rise. Moreover, if firms bid against one another in the market to acquire research options, a counter-veiling effect emerges: each firm has an incentive to acquire options defensively, to keep them from the hands of competitors. When R&D firms compete both in the market for leads, and in the race to patent commercial discoveries, they will be willing to pay a premium for exclusive access to research options.

A Mathematical Appendix

Proof of Proposition 1: The proof of the first sentence, which relies on the independence of the Bernoulli trials, involves a straightforward confirmation that no alternative search sequence can improve payoffs in expectation; see Weitzman (1979). Given that multiple discoveries are strictly redundant, the stopping rule is obvious. ■

Proof of Proposition 2: Repeated application of the recursive relationship (1) yields

the closed-form expression

$$V_{n+1} = \sum_{i=n+1}^{N} \left(\frac{a_i}{a_{n+1}}\right)(p_i R - c) = \left[\sum_{i=n+1}^{N} \frac{a_i p_i}{a_{n+1}}\right] R - \left[\sum_{i=n+1}^{N} \frac{a_i}{a_{n+1}}\right] c.$$

Substituting this into (5) gives

$$v_n = a_n p_n \left[1 - \sum_{i=n+1}^{N} \frac{a_i p_i}{a_{n+1}}\right] R + a_n \left[p_n \left(\sum_{i=n+1}^{N} \frac{a_i}{a_{n+1}}\right) - 1\right] c. \tag{8}$$

Since $a_{i+1} = a_i(1 - p_i)$, the identity $a_i p_i = a_i - a_{i+1}$ can be applied to simplify the first expression in brackets:

$$a_n \left[1 - \sum_{i=n+1}^{N} \frac{(a_i - a_{i+1})}{a_{n+1}}\right] = a_n \left[1 - \frac{(a_{n+1} - a_{N+1})}{a_{n+1}}\right] = \frac{a_{N+1}}{1 - p_n}$$

where $a_{N+1} \equiv \prod_{i=1}^{N}(1 - p_i)$, the probability of project failure. Rewriting (8) yields

$$\begin{aligned} v_n &= \frac{a_{N+1}}{1 - p_n} p_n R + a_n \left[p_n \left(\sum_{i=n+1}^{N} \frac{a_i}{a_{n+1}}\right) - 1\right] c \qquad (9) \\ &= \frac{a_{N+1}}{1 - p_n}(p_n R - c) + \frac{1}{1 - p_n}\left[\left(\sum_{i=n+1}^{N} p_n a_i\right) + a_{N+1} - a_{n+1}\right] c \\ &= \frac{a_{N+1}}{1 - p_n}(p_n R - c) + \left[\sum_{i=n+1}^{N} \frac{a_i}{1 - p_n}(p_n - p_i)\right] c \end{aligned}$$

Now,

$$\frac{a_{N+1}}{1 - p_n}(p_n R - c) = \frac{a_N}{1 - p_n}(p_n - p_n p_N) R - \frac{a_N}{1 - p_n}(1 - p_N) c$$

and

$$v_N = a_N (p_N R - c) = \frac{a_N}{1 - p_n}(p_N - p_n p_N) R - \frac{a_N}{1 - p_n}(1 - p_n) c.$$

Hence

$$\begin{aligned} v_n - v_N &= \frac{a_N}{1 - p_n}(p_n - p_N) R + \left[\sum_{i=n+1}^{N-1} \frac{a_i}{1 - p_n}(p_n - p_i)\right] c \\ &\quad + \frac{a_N}{1 - p_n}[(1 - p_n) - (1 - p_N) + (p_n - p_N)] c. \end{aligned}$$

Since the last term disappears, the proposition is shown. ■

Proof of Corollary 3: Let m and n be given. Without loss of generality, assume $n \leq m$. Consider first the case $p_n = p_m \equiv p$. By Proposition 2, $v_n - v_m = \sum_{i=n+1}^{m} \frac{a_i}{1-p}(p - p_i) c$. But by

Proposition 1, $p_n = p$ and $p_m = p$ implies $p_i = p$ for all intermediate $i = m+1, m+2, ..., n-1$. Hence the first claim is shown. For the case $p_n > p_m$, Proposition 1 implies $n < m$. Hence

$$\begin{aligned} v_n - v_m &= \left(\frac{p_n - p_N}{1-p_n} - \frac{p_m - p_N}{1-p_m}\right) a_N R + \sum_{i=n+1}^{m} \frac{a_i}{1-p_n}(p_n - p_i)\, c \\ &\quad + \sum_{i=m+1}^{N-1} a_i \left(\frac{p_n - p_i}{1-p_n} - \frac{p_m - p_i}{1-p_m}\right) c. \end{aligned}$$

Since $p_n > p_m$ implies $\frac{1}{1-p_n} > \frac{1}{1-p_m}$, all three terms in the expression on the right hand side are positive, and the second claim is proven. The last two sentences are then immediate from Proposition 1 and the definition of information rents. ■

Proof of Corollary 4: Immediate from (9). ■

Proof of Corollary 5: Immediate from (9). ■

Proof of Proposition 6: It suffices to show that v_n is decreasing, as a function of N, and converges to the specified lower bound in the limit as N goes to infinity. To prove the first claim, use (9) to express v_n as a function of N:

$$v_n(N) = \frac{a_{N+1}}{1-p_n} p_n R + \left[\sum_{i=n+1}^{N} \frac{a_i}{1-p_n} p_n - a_n\right] c. \tag{10}$$

Since $a_{N+1} = (1-p)a_N$, we have

$$v_n(N) - v_n(N+1) = \frac{a_{N+1}}{1-p_n} p_n (pR - c) > 0$$

by the assumption $pR - c > 0$. Hence $v_n(N)$ is decreasing in N. To prove the convergence claim, use the relation $a_{i+1} = (1-p)a_i$ for $i > M$ to rewrite the formula in Proposition 2:

$$\begin{aligned} v_n(N) &= v_N(N) + \hat{v}_n(N) \\ &= a_M(1-p)^{N-M}(pR - c) + a_M(1-p)^{N-M}\left(\frac{p_n - p}{1-p_n}\right) R \\ &\quad + \left[\sum_{i=n+1}^{M-1} \frac{a_i}{1-p_n}(p_n - p_i) + \sum_{i=M}^{N-1} a_M (1-p)^{i-M}\left(\frac{p_n - p}{1-p_n}\right)\right] c. \end{aligned}$$

As $N \to \infty$, $(1-p)^{N-M} \to 0$, and $\sum_{i=M}^{N-1}(1-p)^{i-M} \to \frac{1}{p}$, so

$$\lim_{N\to\infty} v_n(N) = \left[\sum_{i=n+1}^{M-1} \frac{a_i}{1-p_n}(p_n - p_i) + \frac{a_M}{1-p_n} \cdot \frac{p_n - p}{p}\right] c. \tag{11}$$

Since convergence is monotonic, the expression on the right forms a lower bound for all terms in the sequence. The bound is strictly positive for $n < M$, and zero for $n \geq M$. ■

Proof of Corollary 7: Claim (i) restates equation (11) for the case $n = M, ..., N$. Claims (ii) and (iii) follow since $\lim_{N\to\infty} v_n(N)$ does not depend on R and is increasing in c. ■

Proof of Proposition 8: To carry out the proof, we expand notation to allow for variability in the ordering of the search queue. Let $I = [c/R, 1)$ be the interval of hit probabilities on which leads are viable. Then each point $\mathbf{p} \in I^N$ corresponds to an assignment of hit probabilities to an unordered collection of N viable leads. Let p_1 be variable, and let $p_{-1} =< p2, p3, \ldots, p_N >$ be a fixed vector of the other hit probabilities. Without loss of generality, label leads so that $p2 \geq \ldots \geq p_N$. Now let Z be the set of all $N!$ permutations of the set $\{1, 2, \ldots, N\}$. Let a rule $s : I \times I^{N-1} \to Z$ for ranking the N leads into a search queue be defined as follows: for leads with index $n = 2, \ldots, N$, let the position $s_n(p_1 \mid p_{-1})$ be given by

$$s_n(p_1 \mid p_{-1}) = \begin{cases} n-1 & \text{if } p_1 \leq p_n; \\ n & \text{if } p_1 > p_n. \end{cases}$$

By exhaustion, this expression determines the position $s_1(p_1 \mid p_{-1})$ of the lead with index 1. Thus leads are examined in the order $2, 3, 4, \ldots, s_1, 1, s_1 + 1, \ldots, N$. In particular, if $s_1 > 1$, the lead with index s_1 occupies the $s_1 - 1^{st}$ position in the queue, immediately preceding the lead with index 1. Since, for $n > 1$, $s_1 > s_n$ if and only if $p_1 \leq p_n$, this queuing rule satisfies the optimality condition of Proposition 1.

Now let $u_n(p_1) \equiv v_{s_n(p_1 \mid p_{-1})}$ denote the incremental value of the lead with index n, viewed as a function of p_1. By Corollary 3, this value depends only on the magnitudes of the components of $\mathbf{p}$: any other queuing rule $s\prime(\mathbf{p})$ that satisfies the optimality condition of Proposition 1 (i.e. any involving a permutation of equi-probable leads) will yield the same incremental values. Hence u_n is well-defined, for all n.

We can now prove the proposition. Rewriting (10) in terms of the new notation, the value of the lead with index 1 can be expressed as a function of its own hit probability:

$$u_1(p_1) = \prod_{i=2}^{N} (1 - p_i)\, p_1 R + \sum_{i=s_1}^{N} \prod_{j=2}^{i} (1 - p_j)\, p_1 c - \prod_{i=2}^{s_1} (1 - p_i)\, c. \tag{12}$$

By inspection, u_1 is linear in p_1 on all intervals for which s_1 is constant. Hence, u_1 is continuous on intervals of the form $p_n > p_1 > p_{n+1}$ for $n = 2, \ldots, N-1$, and on the interval $p_N > p_1 \geq c/R$. In addition, since $s_1 > s_n$ whenever $p_1 = p_n$, s_1 is constant on intervals $p_1 \in [p_n - \varepsilon, p_n]$ for ε sufficiently small. Hence u_1 is continuous from the left at the finitely-many "switching points" at which $p_1 = p_n$, for some $n > 1$. To prove continuity from the right, suppose that $p_{n-k} > p_{n-k+1} = \ldots = p_n > p_{n+1}$. Then $s_1(p_1 = p_n \mid p_{-1}) = n$, and for positive $\varepsilon < p_{n-k} - p_n$, $s_1(p_1 = p_n + \varepsilon \mid p_{-1}) = n - k$. Applying equation (12) yields

$$u_1(p_n + \varepsilon) - u_1(p_n) = \prod_{i=2}^{N} (1 - p_i)\, R\varepsilon + \sum_{i=n-k}^{N} \prod_{j=2}^{i} (1 - p_j)\, c\varepsilon + \sum_{i=n-k}^{n-1} \prod_{j=2}^{i} (1 - p_j)\, p_n c$$

$$+\left[\prod_{i=2}^{n}(1-p_i)-\prod_{i=2}^{n-k}(1-p_i)\right]c.$$

Since $p_{n-k+1}=p_{n-k+2}=\ldots=p_n$, the third and fourth terms cancel:

$$\begin{aligned}\sum_{i=n-k}^{n-1}\prod_{j=2}^{i}(1-p_j)\,p_n c &= \prod_{i=2}^{n-k}(1-p_i)\left[\sum_{j=0}^{k-1}(1-p_n)^j\right]p_n c\\ &= \prod_{i=2}^{n-k}(1-p_i)\left[1-(1-p_n)^k\right]c\\ &= \left[\prod_{i=2}^{n-k}(1-p_i)-\prod_{i=2}^{n}(1-p_i)\right]c.\end{aligned}$$

Hence $u_1(p_n+\varepsilon)\to u_1(p_n)$ as $\varepsilon\to 0$, and so u_1 is continuous for all $p_1\in I$. Equation (12) also implies that, on any open interval of arguments p_1 for which s_1 is constant, u_1 is differentiable, with

$$\frac{du_1(p_1|\,s_1)}{dp_1}=\prod_{i=2}^{N}(1-p_i)\,R+\sum_{i=s_1}^{N}\prod_{j=2}^{i}(1-p_j)\,c. \qquad (13)$$

Since the expression on the right is positive, u_1 is increasing on each such open subinterval. Continuity then implies that u_1 is increasing on the entire interval I. Further, since du_1/dp_1 is monotone decreasing in s_1, and s_1 is monotone decreasing in p_1, u_1 is weakly convex on I. Finally, (13) implies that

$$\frac{du_1(p_1|\,s_1)}{dp_1}\cdot p_1=\prod_{i=2}^{N}(1-p_i)\,p_1R+\sum_{i=s_1}^{N}\prod_{j=2}^{i-1}(1-p_j)\,p_1c=u_1+a_{s_1}c$$

where, following our established conventions, $a_{s_1}=\prod_{i=2}^{s_1}(1-p_i)$ denotes the probability that the lead with index 1 is tested. ■

Proof of Proposition 9: Let $n\in\{2,3,\ldots,N\}$ be the index for a given lead. We use the notation developed for the proof of Proposition 8 to derive a formula for $u_n(p_1)$, the value of the lead with index n, expressed as a function of p_1. For arguments $p_1>p_n$ (so that $s_n=n>s_1$), this value is given by

$$u_n(p_1|p_1>p_n)=\prod_{\substack{i=1\\ i\neq n}}^{N}(1-p_i)\,p_nR+\sum_{i=n}^{N}\prod_{\substack{j=1\\ j\neq n}}^{i}(1-p_j)\,p_nc-\prod_{i=1}^{n-1}(1-p_i)\,c, \qquad (14)$$

so u_n is linear, and hence continuous, in p_1 on the interval $1>p_1>p_n$. Further, since $\frac{du_n(p_1|p_1>p_n)}{dp_1}=-\frac{u_n}{1-p_1}$, the elasticity relation claimed in point (i) is shown. For $p_1\leq p_n$ (so that $s_n=n-1<s_1$), the value is given by

$$\begin{aligned}u_n(p_1|p_1\ \leq\ p_n)=&\prod_{\substack{i=1\\ i\neq n}}^{N}(1-p_i)\,p_nR+\left[\sum_{i=n}^{s_1}\prod_{\substack{j=2\\ j\neq n}}^{i}(1-p_j)+\sum_{i=s_1}^{N}\prod_{\substack{j=1\\ j\neq n}}^{i}(1-p_j)\right]p_nc\\ &-\prod_{i=2}^{n-1}(1-p_i)\,c.\end{aligned}$$

This function is linear, hence continuous, in p_1 on subintervals for which s_1 is constant. On each such subinterval, we have that

$$\begin{aligned}\frac{du_n(p_1|s_1 > s_n)}{dp_1} &= -\frac{1}{1-p_1}\left[\prod_{\substack{i=1\\ i\neq n}}^{N}(1-p_i)\,p_nR + \sum_{i=s_1}^{N}\prod_{\substack{j=1\\ j\neq n}}^{i}(1-p_j)\,p_nc\right] \qquad (15)\\ &= -\frac{1}{1-p_1}\left[u_n + \prod_{i=2}^{n-1}(1-p_i)\,c - \sum_{i=n}^{s_1}\prod_{\substack{j=2\\ j\neq n}}^{i}(1-p_j)\,p_nc\right]\end{aligned}$$

which expresses the elasticity relation of claim (ii) in the new notation. That u_n is continuous at points $p_1 = p_m$, for $p_m < p_n$, follows from an argument analogous to the one used in the proof of Proposition 8.

A similar argument shows that $\lim_{\varepsilon\to 0} u_n(p_n+\varepsilon) < u_n(p_n)$, so u_n is discontinuous at $p_1 = p_n$. To complete the proof, it must be shown that u_n is decreasing in p_1. Equation (14) implies that u_n is decreasing in p_1 on the interval $1 > p_1 > p_n$. For $p_1 < p_n$, equation (15) implies that u_n is decreasing on subintervals of $[c/R, p_n]$ on which s_1 is constant. But since u_n is continuous on $[c/R, p_n]$, this implies that u_n is decreasing on any interval not containing p_n. Finally, at the point of discontinuity $p_1 = p_n$, we have $\lim_{\varepsilon\to 0} u_n(p_n+\varepsilon) < u_n(p_n)$. Thus u_n is decreasing for all $p_1 \in I$. ■

Proof of Corollary 10: We showed in Proposition 8 that the elasticity of a lead's value u_1 with respect to its own hit probability p_1 is greater than unity on subintervals of I for which the search ordering is constant. The first claim of the corollary then follows from the fact that u_1 is continuous, increasing and convex in p_1. The second claim, that u_n is decreasing in p_1 for $n > 1$, was proven in Proposition 9. ■

Proof of Corollary 11: Consider a second probability ordering $\{p'_n\}_{n=1}^{N}$ such that $p'_n = p_n$ for $n < M$, and $p'_n = p$ for $n \geq M$. Let v'_n be the value of lead n with respect to this ordering. Then for $n < M$,

$$v'_n \geq \left[\sum_{i=n+1}^{M-1}\frac{a_i}{1-p_n}(p_n - p_i) + \frac{a_M}{1-p_n}\cdot\frac{p_n - p}{p}\right]c$$

by Proposition 6, and $v_n = v'_n$ by Proposition 9. To prove the limit results, note that $p_n \geq c/R$ for all n implies $a_N \leq (1-c/R)^{N-1}$. Hence $a_N \to 0$ as $N \to \infty$. Then by Proposition 2 and the above inequality, $\lim_{N\to\infty} v_n(N)$ does not depend on R, is increasing in c for $n < M(p)$, and is zero for $n = N$. ■

B Model Robustness

The model incorporates several salient features of bioprospecting, as well as applied generic research: total costs increase with project duration, innovations are nonrival goods, and

scientific knowledge gives direction to the effort overall. Nonetheless, parsimony requires that some considerations be treated in reduced form or swept aside. It is important to consider whether the implications of the analysis are likely to be robust with respect to potential generalizations.

The model treated success probabilities for different leads as statistically independent. If hit rates are instead correlated, then the investigator can update her search itinerary based on information revealed in the light of sequential test outcomes. The value of a given lead then depends both on its own fertility as a potential source of a discovery and on what a test of it reveals about the likelihood of discoveries elsewhere. Since search halts after the first success, a test only affects the course of the itinerary when it results in failure. Knowing this, the investigator can map her entire itinerary in advance, using an appropriate backwards-induction procedure to determine the optimal sequencing of leads. She can then calculate *ex ante* the probability a_n that her search will ever visit the n^{th} lead, and the value v_n of a call option on the lead, using formulae identical to those derived here. The only change involves replacing the unconditioned probability parameters $\{p_i\}_{i=1}^{N}$ with corresponding conditional probabilities. The central message—that leads have value insofar as they improve the productivity of search effort—would be unchanged, if not strengthened.

Research costs c and payoffs R were treated as constant across projects, and known in advance. In practice, bioprospecting projects often involve multi-stage testing, in which samples that show promise in an initial screening are subjected to more rigorous and expensive follow-on investigation. Samples can thus differ with respect to their ultimate cost of testing. In this case, however, the researcher will not know *ex ante* which samples will require additional expense. For the purpose of modeling the researcher's incentives we can, therefore, treat c as a constant expected cost per test, and hit rate p_n as the *ex ante* probability that the n^{th} sample will clear all the stage-hurdles associated with research and development, leading ultimately to the marketing of a new product.

Likewise, the payoff R may be uncertain at the outset of a product-development process. The payoff ultimately realized will depend on the resolution of multiple contingencies concerning consumer demand, competition, etc. It is assumed, however, that the firm can formulate beliefs about its payoffs over a rich set of contingencies and calculate R as an expected value. If the firm is risk-neutral, its approach to the project will be the same in either case, and the use of the reduced-form expression is justified. More generally, R can be viewed as the certainty equivalent of the payoff distribution under the objective function of a risk-averse firm. If R varies across different projects (e.g. if a cure for cancer is worth more than a cure for the common cold), conservation incentives can be calculated as the simple sum of the incremental values for each of the several projects.

The model incorporates no role for discounting during the life of an individual research project. Suppose instead that each test required some amount of time to carry out, and that future costs and benefits were discounted at some (possibly variable) rate. In this setting, a lead that enables rapid discovery is doubly beneficial: an early discovery not only reduces the direct expense of testing, but also increases the present value of the payoff. Incorporating a

role for in-project discounting would apparently strengthen the qualitative conclusions about the importance of information rents.

In the numerical illustration, an individual "lead" corresponds to a parcel of land of a certain size. Previous studies in the economics of bioprospecting have typically identified leads with individual species. However, in pharmaceutical natural-products especially, the objects tested are often complexes of micro-organisms, large and lumpy aggregates of living material found in soil, leaf litter, and so on. The firm may not know, nor care about, the identity of the species in a sample. The operative question is whether *something* in a biological sample generates a chemical compound with desirable properties. From this perspective, the individual species is merely an intermediate carrier of chemical creativity. Indeed, an atomistic focus on species, in abstraction from ecological and environmental relationships, can interfere with the discovery process. A plant may produce defensive chemicals, for example, only in its stems or leaves, or only when under attack from insects. The identification of leads with species presupposes an impoverished informational context. Several considerations argue for a geographic definition of "leads." Data on species distributions and other features of ecology and habitat are typically organized along geographic lines. More importantly from the policy perspective, property rights, access agreements, and investment decisions are generally location-specific.

Numerical results were based on the additional simplifying assumption that the hit rate for a given lead is constant across different classes of research projects. In practice, this pattern is unlikely to hold exactly: a mushroom that produces psychotropic chemicals may be of no use against high blood pressure. If hit rates displayed no correlation across projects, then a lead with high value in one application might have low value in others, and only modest aggregate rent. To the extent that hit rates exhibit positive correlation across applications, significant lead valuations become more likely. The degree to which organisms of potential economic use are concentrated in certain biological taxa is an empirical issue that is far from resolved. One of the few studies that attempts to address the question comprehensively within a subset of US plant species (Phillips and Meilleur 1998) finds a substantial variance across genera in the incidence of reported economic use, with certain taxa displaying uncommon fertility as product sources. As noted in the text, certain environments (Yellowstone hot springs, centers of crop diversity, areas rich in chemically creative microbes) appear to be durably promising as sources of innovations in lucrative product areas. A pattern of positive correlation could also be strengthened through path-dependent learning: as investigators examine an organism for use in one application, they may learn things about its suitability in others. While the correlation question merits further study, a reasonable null hypothesis at this point is that some biological leads are systematically superior as sources of innovations, and embody potentially significant commercial value.

References

Arrow, Kenneth J. and Anthony C. Fisher, "Environmental preservation, uncertainty, and irreversibility," *Quarterly Journal of Economics*, 1974, *88*, 312–319.

Brown, Gardner and Jon H. Goldstein, "A Model for Valuing Endangered Species," *Journal of Environmental Economics and Management*, 1984, *11*, 303–309.

Chrispeels, Maarten J. and David E. Sadava, *Plants, Genes and Agriculture*, Boston: Jones and Bartlett Publishers, 1994.

DeGroot, Morris H., *Optimal Statistical Decisions*, New York: McGraw Hill, 1970.

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

Dreyfuss, M.M. and Ignacio H. Chapela, "Potential of Fungi in the Discovery of Novel, Low-Molecular Weight Pharmaceuticals," in Vincent P. Gullo, ed., *The Discovery of Natural Products with Therapeutic Potential*, Boston: Butterworth-Heinemann, 1994.

Evenson, Robert E. and Yoav Kislev, "A Stochastic Model of Applied Research," *Journal of Political Economy*, 1976, *84* (2), 265–281.

Heal, Geoffrey M., "Markets and Biodiversity," PaineWebber Working Paper Series in Money, Economics and Finance PW/95/17, Columbia University September 1995.

Kohn, Mein G. and Steven Shavell, "The Theory of Search," *Journal of Economic Theory*, 1974, *9* (2), 93–123.

Lippman, S.A. and J.J. McCall, eds, *Studies in the Economics of Search*, Amsterdam; New York: North-Holland, 1979.

Moeller, Christine, "Biolead: A New Approach to Drug Discovery," Presentation, Seminar on Biotechnology and Genetic Resources, University of California at Berkeley, College of Natural Resources 1996.

Myers, Norman, "Threatened Biotas: 'Hot Spots' in Tropical Forests," *The Environmentalist*, 1988, *8*, 197–208.

——, "The Biodiversity Challenge: Expanded Hot-Spots Analysis," *The Environmentalist*, 1990, *10*, 243–256.

——, "Biodiversity's Genetic Library," in Gretchen C. Daily, ed., *Nature's Services*, Washington: Island Press, 1997, pp. 255–273.

Office of Technology Assessment, "Pharmaceutical R&D: Costs, Risks, and Rewards," Technical Report, United States Congress 1993.

Pace, Norman, "A Molecular View of Microbial Diversity and the Biosphere," *Science*, 1997, *276*, 734–740.

Pennisi, Elizabeth, "Lawsuit Targets Yellowstone Bug Deal," *Science*, March 13, 1998 1998, *279*, 1624.

Phillips, Oliver L. and Brien A. Meilleur, "Usefulness and Economic Potential of the Rare Plants of the United States: A Statistical Survey," *Economic Botany*, 1998, *52* (1), 57–67.

Polasky, Stephen and Andrew Solow, "On the Value of a Collection of Species," *Journal of Environmental Economics and Management*, 1995, *29*, 298–303.

Pollack, A., "Drug Industry Going Back to Nature." *The New York Times*, 3 March 1992.

Reid, Walter V., Sarah A. Laird, Carrie A. Meyer, Rodrigo Gamez, Ana Sittenfeld, Daniel H. Janzen, Michael A. Gollin, and Calestous Juma, "A New Lease on Life," in Walter V. Reid, Sarah A. Laird, Carrie A. Meyer, Rodrigo Gamez, Ana Sittenfeld, Daniel H. Janzen, Michael A. Gollin, and Calestous Juma, eds., *Biodiversity Prospecting: Using Genetic Resources for Sustainable Development*, Washington, D.C.: World Resources Institute, 1993, pp. 1–52.

Sawhill, John C., "Research and Development," *Nature Conservancy*, May/June 1998, *48* (3), 5.

Simpson, David R., Roger A. Sedjo, and John W. Reid, "Valuing Biodiversity for Use in Pharmaceutical Research," *Journal of Political Economy*, 1996, *104* (1), 163–185.

Stone, Lawrence D., *Theory of Optimal Search*, New York: Academic Press, 1975.

Weisbrod, Burton A., "Collective-Consumption Services of Individual-Consumption Goods," *Quarterly Journal of Economics*, 1964, *78*, 471–477.

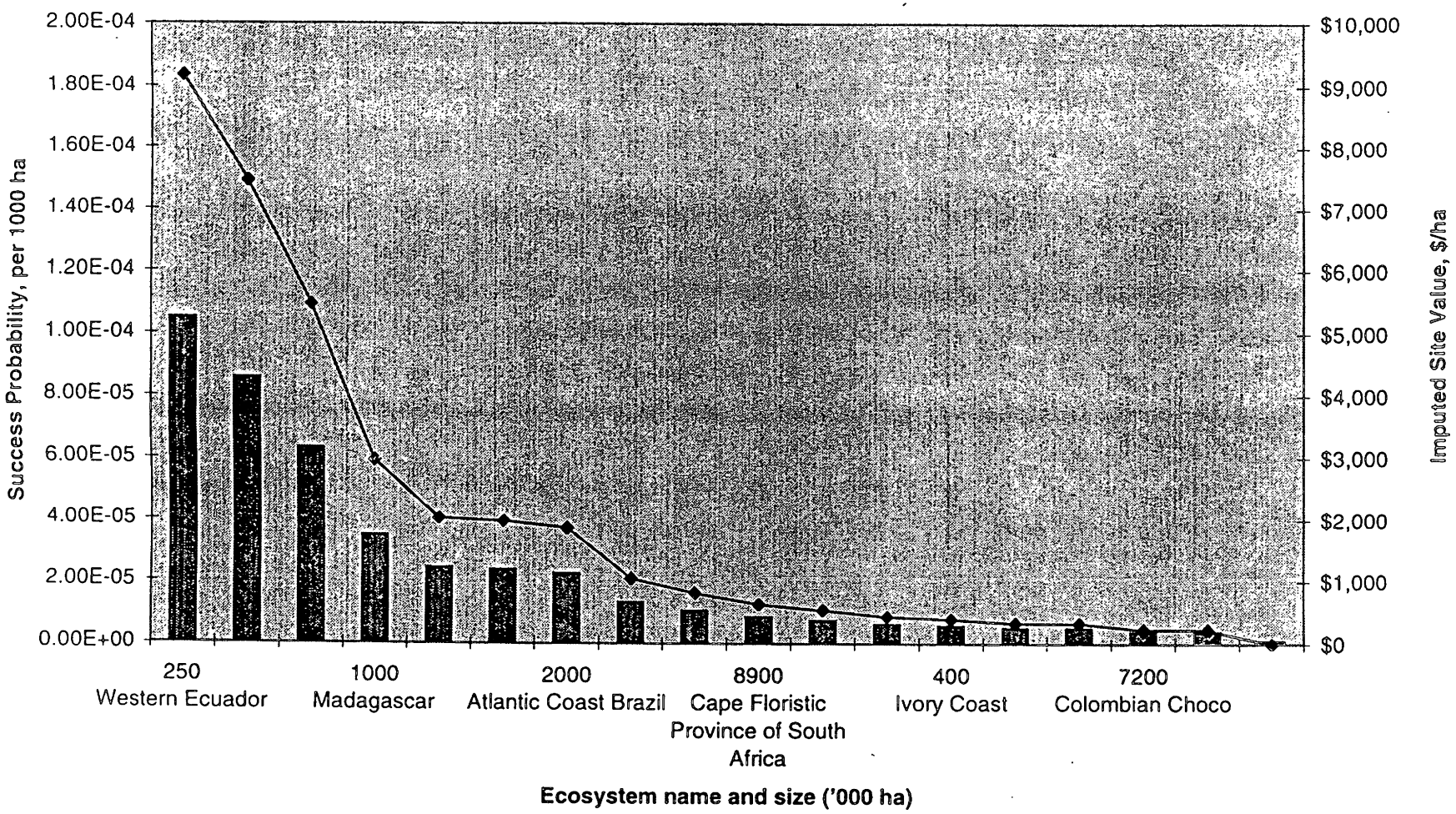
Weitzman, Martin L., "Optimal Search for the Best Alternative," *Econometrica*, 1979, *47* (3), 641–654.

Table 1: Bioprospecting Values in Several Ecosystems, as a Function of Density of Endemic Species

Biodiversity "Hot Spots"	Forest Area (1000 ha)	Density, endemic species / 1000 ha	Hit Probability (/ 1000 ha)	Incremental Value ($/hectare)	SSR Scarcity Rent ($/hectare)
Western Ecuador	250	8.75	1.05E-04	$9,177	$20.63
Southwestern Sri Lanka	70	7.14	8.57E-05	$7,463	$16.84
New Calendonia	150	5.27	6.32E-05	$5,473	$12.43
Madagascar	1000	2.91	3.49E-05	$2,961	$6.86
Western Ghats of India	800	2.03	2.44E-05	$2,026	$4.77
Philipines	800	1.98	2.38E-05	$1,973	$4.66
Atlantic Coast Brazil	2000	1.88	2.26E-05	$1,867	$4.42
Uplands of Western Amazonia	3500	1.10	1.32E-05	$1,043	$2.59
Tanzania	600	0.88	1.06E-05	$811	$2.07
Cape Floristic Province of South Africa	8900	0.71	8.52E-06	$632	$1.66
Peninsular Malaysia	2600	0.62	7.44E-06	$539	$1.47
Southwestern Australia	5470	0.52	6.24E-06	$435	$1.22
Ivory Coast	400	0.48	5.76E-06	$394	$1.14
Northern Borneo	6400	0.42	5.04E-06	$332	$0.99
Eastern Himalayas	5300	0.42	5.04E-06	$332	$0.98
Colombian Choco	7200	0.32	3.84E-06	$231	$0.75
Central Chile	4600	0.32	3.84E-06	$231	$0.74
California Floristic Province	24600	0.09	1.08E-06	$0	$0.20

Assumes 10 successes/year, revenues $450,000,000/success, cost $485/test, hit rates based on 1.2 E-05 per species, discount rate 10%.
Source: Myers (1988, 1990); Simpson, Sedjo and Reid (1996); and authors' calculations.

Figure 1: Bioprospecting Values in Several Ecosystems, as a Function of Success Probabilities



Assumes 10 successes/year, revenues $450,000,000/success, cost $483/test, hit rates based on 1.2 E-05 per species, discount rate 10%. Source: Myers (1988, 1990); Simpson, Sedjo and Reid (1996), and authors' calculations.