Valuing Plant Genetic Resources: An Economic Model of Utilization of the U.S. National Crop Germplasm Collection

Armineh Zohrabian[1], Greg Traxler[2]

Annual meeting of AAEA in
Nashville, Tennessee, August 8-11, 1999.

---

[1] Auburn University, Ph.D. candidate

[2] Auburn University, Associate professor

Armineh Zohrabian, Greg Traxler

Valuing Plant Genetic Resources: An Economic Model of Utilization of the U.S. National Crop Germplasm Collection

INTRODUCTION

Germplasm is a term used to describe seeds, plants, or plant parts useful in crop breeding, research, and conservation efforts. The terms "germplasm" and "genetic resources" are used interchangeably throughout the paper. Accessions are uniquely identified samples of seeds, plant or plant parts maintained as part of a germplasm collection. Apart from finished varieties, genetic resources are generally not exchanged in the marketplace, so there is no price assigned to them through traditional market processes. The fundamental reason for the absence of regular markets for germplasm is that it has non-rivalry characteristic similar to invention: using germplasm to create a new crop variety does not alter its performance in the original plant. The continuous extinction of plant species, advancement of biotechnology, and expansion of intellectual property rights over plant materials change the supply and demand for germplasm, as well as the ability to capture returns from using germplasm. These factors may contribute to the development of markets for germplasm.

Ex situ collections (preservation of genetic materials away from the environment of origin) exist for all important crops. There are hundreds of such collections with roughly 6 million accessions for all crops. The largest germplasm holdings are in the USA, China, and Russia. The question addressed here is whether it is economically justified to intensify the use and to expand the existing US collections. Simpson and Sedjo (1996, 1998) suggest that genetic resources may simply not be scarce, and for that reason not of much economic value, while Evenson and Lemarie (1998), and Gollin et al (1998) find large net returns from an additional search opportunity (searching germplasm collection for desired trait). Also the latest GAO study (1997) shows that US crop germplasm scientists rank germplasm acquisition, among 14 activities that gene banks undertake, as the number one priority in the event of additional funding.

Exploration of a germplasm collection for a particular trait is viewed as a search within a given distribution. An optimal strategy would be to search and collect additional accessions for traits as long as expected benefit is greater than the cost of collecting, conserving and testing it. The probability of finding a desirable trait depends on the number of accessions that are screened for the trait, and the distribution of that trait in the collection or in certain subcollections. This study will estimate the expected net return from an additional search opportunity (in regard to several pest resistance traits for soybeans) both when the additional accession is from the existing collection, and when it is newly acquired. We will use evaluation data from the USDA soybean collection, and benefit/cost data from published literature and from germplasm experts. The estimates from this study will provide gene bank managers with better information about the marginal return from a variety of research options and will provide insights into problems of allocating public funds for maintenance and further development of public gene-banks, and for setting priorities among various types of plant pre-breeding activities.

BACKGROUND

The US National Plant Germplasm System (NPGS) is primarily a federal and state-supported effort aimed at maintaining supplies of germplasm with diverse genetic traits for use in breeding and scientific research. This is a geographically dispersed network of germplasm collections administered primarily by USDA's Agricultural Research Service (ARS). Major NPGS activities are:

1. Acquisition--collecting plant germplasm from natural habitats and through exchange with other scientists or collections.
2. Preservation--storing and maintaining germplasm to ensure its diverse supply. In addition, NPGS distributes germplasm to breeders and other researchers;
3. Development and documentation of genetic information.
4. Some germplasm enhancement (the same as pre-breeding) is also undertaken by gene banks. Germplasm enhancement includes two stages: screening accessions fordesirable traits (the same as a search for traits), and incorporation of desirable traits into breeding material.

Breeding for resistance to some pests is an objective of almost all breeding programs. Despite considerable success over the years in managing them, plant diseases still cause more than nine billion dollars in annual losses in the U.S. alone. This is largely because: 1) current control strategies are not 100% effective, 2) pathogens evolve and overcome once-effective management tactics, and 3) exotic pathogens are introduced. Because only the fittest survive in nature, wild species and landraces (varieties that have resulted from farmers' selections) contain genes with resistance to survive diseases, drought, flood, and hot or cold conditions. Thus these wild species and landraces are used by plant breeders to enhance genetic resistance to pests and stress in modern cultivars. The genetically enhanced host plant resistance is sometimes the only disease management option and is often the most cost effective, environmentally benign alternative (USDA National Programs: Program Statement 1998).

STATEMENT OF PROBLEM

Two main questions are addressed in this paper are: 1) What are the expected net benefits of screening the (n+1)st vs n accession for pest resistance? I will consider two separate cases: first, the search for desirable trait is intensified through screening additional accessions from the existing collection, and, second, a new accession is added to the collection. 2) What is the optimal size of the U.S. National soybean collection in an economic model where the returns from discovering several desirable traits are maximized jointly? Due to collection and conservation cost complementarities, the optimal collection size in the search for multiple traits is greater when profits are jointly maximized than is implied by the sum of independent maximization problems for single traits (Evenson and Lemarie 1998)

Researchers simultaneously test a sample of accessions to find a desirable level of resistance. Subsequent groups may be chosen for further screening if the previous tests were not satisfactory. The screening will stop when the marginal benefits from finding and utilizing the desirable trait are equal to the marginal costs of screening. Every resistance trait has a unique distribution across accessions. The probability that a plant with a given level of resistance will be found during a screening experiment depends on the underlying distribution of the trait and the number of accessions tested. Suppose the collection has N accessions. When a new plant is added to the collection, or the number of accessions under the test is increased, the expected value of resistance changes, thus changing the expected benefits from the search. The value of testing (n +1)st accession is equal to the expectation of the improvement over the best level of resistance identified among the first n accessions tested, conditioned on the value of the resistance in the existing cultivar at the time of testing, less the cost of testing. Knowing the expected net benefits from an additional test, researchers can decide how large to make the sample of accessions for screening. The expected net benefits from testing an additional accession can be used in estimating the optimal size of the collection.

The derivation of actual probability distribution functions for several pest resistance traits will be an important contribution to the knowledge of germplasm diversity. All of the previous studies related to economic evaluation of marginal germplasm have made assumptions about the underlying distribution of the traits. I will look for the best method of deriving empirical probability distribution functions for several soybean pest resistance traits.

LITERATURE REVIEW

From November 1996 through March 1997 GAO conducted a survey of 680 crop germplasm experts from forty Crop Germplasm Committees (CGC) which provide technical advice and guidance to NPGS on germplasm activites for all 85 crops nationwide. The CGCs include NPGS curators who are responsible for maintaining and preserving the collections, as well as scientists in the area of basic research (such as biochemistry and genetics), plant pathologists, breeders, research management, etc. Response rates were above 70 percent for all but four CGCs. The acquisition of additional germplasm was a top gene bank priority for respondents. The GAO Report has summarized the rankings of the priorities to be given to 14 germplasm-related activities in the event of additional funding. The six top- priority activities are (in a decreasing level of priority): 1) acquisition, 2) evaluation, 3) characterization, 4) enhancement/pre-breeding, 5) passport, and 6) active preservation. The concerns that may contribute to the importance of increased acquisition are: the inadequate diversity of germplasm making up the existing collections, and the potential loss of germplasm that is now at risk in nature. Thus, the experts believe that there is considerable scarcity of genetic resources in certain collections.

In contrast to crop germplasm experts' belief that acquiring more germplasm must be a gene bank's top priority, Simpson and Sedjo's studies (1996, 1998) suggest that genetic resources, both for use in pharmaceutical research and in agriculture, may not be scarce, and for that reason not of much economic value. In their earlier analysis of biodiversity prospecting in pharmaceutical research (1996) they treat the testing process as a binomial outcome. The majority of natural materials tested for pharmaceutical applications simply "don't work", as opposed to working at higher or lower levels. Thus, a perfect substitutability exists among the resources that possess the trait, and a zero substitutability with the materials that do not possess it. In this setting, Simpson and Sedjo (1996) conclude that "regardless of the probability of the discovery of a commercially useful compound, if the set of organisms that may be sampled is very large, the value of the marginal species must be very small". The assumption of substitutability among genetic materials has a key implication in the assessment of their value (Gollin and Smale 1998). In agricultural research, on the other hand, the search is for a genotype that "works best" in the circumstances under which it is to be cultivated. In fact, genetic principles might lead one to suppose that many attributes of interest would be (approximately) continuously distributed (Gollin and Smale 1998). In this setting, an additional search is always expected to offer some marginal benefit unless some extreme value has already been achieved (Evenson and Kislev 1975, Evenson 1996).

In their recent work Simpson and Sedjo (1998) focus on the social, rather than the private, value of genetic diversity.

They link the value of genetic improvement to the increase in social welfare. Researchers seek to improve crop yields by finding a genotype with the highest value of some parameter θ. This random variable could represent yield per hectare, resistance to drought, pests, and/or any of a number of other attributes. Social welfare is represented as the difference between consumer surplus from the consumption of the crop and the costs of growing the crop:

$$W(\theta)={}_0\int^{q(\theta)}p(x)dx — C[\theta, q(\theta)].$$

Welfare is increasing in θ if an increase in θ induces a reduction in total cost. Under the assumption that demand for agricultural production is relatively inelastic ($\eta<1$), Simpson and Sedjo observe that the value of the marginal genotype will be smaller to the extent that relatively large realizations of θ have relatively low value to society. (Unitary elasticity will imply infinite welfare, and an elasticity greater than one will imply negative utility).

Several authors (Evenson, Lemarie, and Xiao 1997, Gollin, and Smale1998, etc) base their economic model of pre-breeding on search theory. Applied research is viewed as a search in a given distribution. Much of research work is experimentation, and often a technological development project consists of the testing of a collection of technologies (in this case, varieties of crops) to find the best one (Evenson and Kislev 1976). While Evenson, Lemarie, and Xiao (1997) develop a theoretical model which derives equations for the optimal size of search for a given trait, as well as for the marginal value of genetic material, Gollin, Smale, and Skovmand (1998) conduct an empirical study of the CIMMYT wheat collection. Using actual evaluation data from USDA GRIN database, they estimate the number of accessions for optimal search and the net benefits that result from that search. Three conclusions are made: 1) the optimal scale of a search for desirable traits is very sensitive to the size of the economic problem, as well as to the probability distribution for the trait. 2) Differences across types of genetic materials in the cost of search and in the associated time lags can lead to optimal search strategies in which some materials are systematically ignored. Until new wide-cross and molecular techniques can substantially reduce the cost and time constraints on evaluation and pre-breeding, collections of landraces and wild relatives are expected to be used seldom. 3) The results are suggestive that even in large collections, there are non-trivial benefits associated with marginal accessions.

The recent crop breeding model by Evenson and Lemarie (1998) is different from previous studies in three ways. First, plant breeding is designed to consist of two stages: a collection activity in the first stage and a selection of the best accessions within this collection in the second stage. In addition to looking for the optimal search strategy, they look for the optimal collection strategy as well. Second, the authors introduce a geographic dimension into the distribution of source materials. Assuming that the information about the geographic source region is available for each accession, the total collection can be represented as a series of subcollections. Evenson and Lemarie (1998) show that for any given trait, there is no reason to exploit more than one subcollection. Based on prior information about the distribution of a trait of interest within given subpopulation, researchers conduct independent and random draws to evaluate each accession for the level of performance in respect to that trait, and to choose the best plant for breeding. Third, the search is explicitly modeled for multiple rather than single traits. A model representing a joint search for multiple traits fits more realistically to the general case of crop improvement in a breeding program than a model of search for a single trait.

Evenson and Lemarie (1998)observe that the optimal size of collection is sensitive to the number of traits for which the collection is made, as well as to the diversity of the distribution functions among the different regions providing source materials. The optimal collection will be larger than implied by the single trait model, because conservation costs are shared among the different traits. Thus, the marginal value of accessions will most likely be higher when multiple traits are considered than in the case of simple trait model. It will also be larger if traits are clustered in “niches” rather than distributed randomly across the population.

DATA

Agronomists use scales of a discrete nature to describe resistance traits. In soybeans they generally use the scale “Resistant”, “Moderately Resistant”, “Moderately Susceptible”, and “Susceptible”, based on the extent of damage caused by the pest during agronomic tests. Based on raw data from agronomists on the performance of individual accessions during screening, I will construct a continuous scale for resistance levels, assigning 0 to the least resistant accession, and 1-- to the most resistant. It is also possible to obtain estimates of yield losses associated with different levels of resistance. Having the resistance scale and yield losses for each resistance level, I will derive several plausible “damage” functions between the fraction of losses and the resistance. I will consider cases when that function is linear, concave, or convex. This way I can generate information on the percent of crop losses corresponding to the incremental change in resistance resulting from testing an additional accession.

The expected change in the resistance level is a function of the trait distribution across accessions and the number of accessions that are tested. It is possible to obtain data on the distribution of resistance across large samples of accessions for certain pests. This will allow probability distribution functions be estimated for that trait, which will later be useful in optimizing the size of the search and the size of the collection. It is reasonable to assume that this distribution can be viewed as independent and identical when looking at groups of accessions of the same type (e.g. landraces or wild species) and of the same geographic origin. The estimates of crop losses averted as a result of reducing or eliminating a disease will be used as a proxy for economic value of crop germplasm. Estimates of areas affected by diseases, average annual yield losses, and expected lifetime of resistance will be obtained from published literature and USDA soybean pathologists. I will consult with plant breeders about the period over which the benefits are calculated which will be determined by the resistance depreciation path and the research lag. Monte Carlo simulations will give values for net benefits when different time lags, depreciation paths, and discount rates are considered.

MODELING APPROACH

I will consider two cases with regard to the distribution function of the trait: 1) researchers screen accessions for a desirable trait without prior information about its distribution, 2) prior information is available about the distribution of trait. When a new pest is identified and an experiment is designed to screen accessions for useful sources of resistance, researchers have no information about the distribution of that trait. For this case, treating the distribution of resistance as a continuous uniform distribution seems a reasonable assumption. However, for some pests, such as Soybean Cyst Nematode, throughout years sequential tests have been conducted due to the widespread character of that pest. In the subsequent experiments (subsequent groups of additional accessions to be tested) researchers can design their experiment based on the information obtained from previous search.

Problem 1. The Marginal Value of Accession.

Suppose the number of accessions tested simultaneously (as one group) is n. The resistance level of the i^{th} accession is r_i. F(r) is the distribution function and f(r) is the probability density function for the resistance level variable. Denote z as the maximum level of resistance among n accessions. The probability that z is the highest level of resistance when n accessions are tested is: $P\left(r_{(n)} \leq z\right) = P\left(r_1 \leq z, r_2 \leq z, \ldots r_n \leq z\right) = F_{(n)}(z)$ (1)

Since each test can be considered as an independent event, and $P(r_i \leq z) = F(z)$, then

$$F_{(n)}(z) = P(r_1 \leq z)P(r_2 \leq z)\ldots P(r_n \leq z) = [F(z)]^n \qquad (2)$$

Since the current cultivar has some level of resistance, say, z_0, the expected maximum increment resistance from searching n accessions will be:

$$E(\Delta z_{(n)}) = \int_0^1 nz[F(z)]^{n-1} f(z)dz - z_0 \qquad (4)$$

Given the probability distribution function for a trait (using uniform continuous, or a function derived from prior tests), through (5) we can find the expected maximum level of resistance as a function of n. Using this approach, we can illustrate how the expected value of the resistance level changes in increments when testing (n+1) vs n accessions. Having a function that relates the number of tested accessions to the expected maximum resistance level, and a function that relates levels of resistance to crop losses, I can derive the economic value of an additional accession. The total benefit stream from searching an additional accession is the sum of discounted annual savings (considering the time lags of research and the depreciation path of the resistance trait) over the average value of production.

When no prior information is available about the distribution of resistance, the distribution function is assumed to be continuous uniform in [0, 1] interval.

$$E(\Delta z_{(n)}) = n\int_0^1 z^n dz - z_0 \qquad (6)$$

$$E(\Delta z_{(n)}) = \frac{n}{n+1} - z_0 \qquad (7)$$

Problem 2. The Optimal Size of Collection.

Let the marginal value of z be ν, and the cost of agronomic screening be C^e . Then the expected profit from searching n accessions is:

$$Ep = n \cdot E(\Delta z) - n \cdot C^e \tag{8}$$

The first order condition for the optimal level of search is:

$$\frac{\partial Ep}{\partial n} = \frac{\partial Ez(n)v(n)}{\partial n} - C^e = 0 \tag{9}$$

The optimal size of search can be found by solving (9) when the distribution function and the marginal value of the trait are known. I will use computational techniques to solve this problem.

The collection (acquisition) costs should be considered while determining the optimal size of collection. In stage 1, a collection of N_1 accessions is made. Denote C^c as the cost of collection and conservation. The distribution of trait is unknown at this stage. In stage 2, the collection is screened for a desirable trait. N_2 accessions are drawn randomly from the collection without replacement and evaluated ($N_2 \leq N_1$). The expected profit from collecting, conserving and searching the accessions for a single trait is:

$$Ep^{c} = nE(\Delta z) - N_1 C^c - N_2 C^e \tag{10}$$

and profit is maximized using backward induction. The optimal collection size is determined based on the knowledge of the size of the existing collection, and the optimal search size when collection and conservation costs are not regarded (N_2^0), the latter can be found through equation (9). Imposing the constraint that $N_2 \leq N_1$, the optimal search size in the 2-stage model is: $N_2^* = \min\left(N_1, N_2^0\right)$ (11)

If $N_1 \leq N_2^0$ then $N_2^* = N_1$, and

$$Ep = nE(\Delta z) - N_1\left(C^c + C^e\right) \tag{12}$$

N_1^* can be found by maximizing (12).

If $N_1 > N_2^0$, then $N_2^* = N_2^0$, and $Ep = nE(\Delta z) - N_1 C^c - N_2^0 C^e$. This function is linear and strictly decreasing in N_1. Consequently, the profit when $N_1 > N_2^0$ is always less than the maximal profit when $N_1 \leq N_2^0$.

Finally, the unique optimal solution when maximizing on two periods is $N_1^* = N_2^*$, which can be determined optimizing equation (12).

Figure 1 in the appendix (Evenson and Lemarie 1998) provides an illustration for this optimization model. The optimal collection size when maximizing for one trait is the optimal search strategy minus “something” which will be more important if collection and conservation costs are high. Above this optimal level, the value of marginal accession is low (Evenson and Lemarie 1998).

Note that each accession is in the form of an actual plant, which contains numerous genes that may determine several useful traits. Each useful trait is indexed by j. Using the above described methods, one can determine, independently from one another, the optimal collection and selection sizes for each trait. The sign ~ is used to distinguish variables of the profits of the independent problems from the variables of joint profit maximization. Define two types of traits. A major trait is the trait for which the optimal number of accessions to collect is the largest among all the other traits, or j such that $\tilde{N}_{1j}^* = \tilde{N}_1^*$, i.e. the optimal collection size when considering the traits independently is equal to the optimal collection size estimated for the major trait. Suppose the major trait is *a*. The rest of the traits are considered minor traits, which are ranked in order of decreasing values in regard to their optimal search size: $N_{2b}^0 > N_{2c}^0 > \ldots$. The profit can then be written as the sum of profits for each trait. Notice that the profits for minor traits (traits other than *a*) do not include collection costs.

$$\begin{cases} Ep_a = n_a E\left(\Delta z_a\right) - N_1 C^c - N_{2a} C_a^e \\ Ep_j = n_j E\left(\Delta z_j\right) - N_{2j} C_j^e \end{cases} \tag{13}$$

Figure 2 (Evenson and Lemarie 1998) illustrates the profit function at the trait level when N_{2j} is chosen optimally. For the major trait, $\partial Ep_a / \partial N_1 = 0$ because $\tilde{N}_1^* = \tilde{N}_{1a}^*$. For the minor traits, we need to consider the optimal level of exploitation without collection and conservation costs, N_{2j}^0. Two cases have to be distinguished: (1) if $N_1^* > N_{2j}^0$, (e.g. trait *b* in figure 2), then the exploitation is $N_{2j}^* = N_{2j}^0$ whatever N_1^*, and consequently, $\partial Ep_j / \partial N_1 = 0$; (2) if (e.g. trait *c* in figure 2), then the optimal exploitation is $N_{2j}^* = N_1$, and $\partial Ep_j / \partial N_1 > 0$.

If we can find at least one minor trait such that $N_{2j}^0 \geq N_{1k}^*$, then the optimal collection size will be larger than the optimal collection size when maximizing profits independently for each trait. Starting from $\tilde{N}_1^*$ and increasing N_1, two changes occur. For the major trait the marginal profit becomes negative. For the minor traits which satisfy $N_{2j}^0 \geq N_1^*$, the marginal profit decreases until the level $N_1^* = N_{2j}^0$, and thereafter the marginal profit is nil (Figure 2). Because of these two changes, the marginal profit from collection decreases and becomes negative at some point, so an optimal collection size can be determined.

CONCLUDING REMARKS

The overall focus of this project is to estimate the expected net gain from an additional search opportunity for the USDA soybean germplasm collection while considering several economically important traits. It is both possible and logical to conduct simulations using data from crop breeding programs, agricultural output, probability distributions of valuable traits in the collection, and the costs of crop improvement research to derive reasonable estimates of the value of marginal accessions.

The findings of this research will provide economic justification for decisions of how much to invest in collecting landraces and wild and weedy species that are currently not in ex situ collections (many of which are presumably subject to danger of extinction). They will also contribute useful information to current debates on germplasm exchange and intellectual property rights.

LITERATURE

Evenson R. 1996. “Valuing Genetic Resources for Plant Breeding: Hedonic Trait Value and Breeding Function Methods”. Manuscript prepared for the CEIS-Tor Vergata University Symposium on the Economics of Valuation and Conservation of Genetic Resources for Agriculture, 13-15 May. Rome.

Evenson R., D. Gollin, and V. Santaniello. 1998. The Agricultural Value of Plant Genetic Resources: Introduction and Overview. In R. E. Evenson, D. Gollin, and V. Santaniello (eds.), *Agricultural Values of Plant Genetic Resources*. Wallingford, UK: CAB International.

Evenson R.. and Y. Kislev. 1976. A Stochastic Model of Applied Research. *Journal of Political Economy* 84 (2): 265-281.

Evenson R. and S. Lemarie. 1998. Optimal Collection and Search for Crop Genetic Resources. In R.E. Evenson, D. Gollin, and V. Santaniello (eds.), *Agricultural Values of Plant Genetic Resources*. Wallingford, UK: CAB International.

Evenson R., S. Lemarie and Xiao. 1997. “ Plant Breeding Models and Implications for Genetic Resources Conservation and Evaluation”. Paper presented at the international conference on Building the Basis for Economic Analysis of Genetic Resources in Crop Plants. Palo Alto, August 17-19, 1997.

Gollin D., M. Smale, and B. Skovmand. 1998. The Empirical Economics of Ex Situ Conservation: A Search Theoretic Approach for the Case of Wheat. In R. E. Evenson, D. Gollin, and V. Santaniello (eds.), *Agricultural Values of Plant Genetic Resources*. Wallingford, UK: CAB International.

Simpson R. , R. Sedjo and J. W. Reid. 1996. Valuing Biodiversity for Use in Pharmaceutical Research. *Journal of Political Economy* 104:1 February, 163-85.

Simpson R. and R. Sedjo. 1998. The Value of Genetic Resource for Use in Agricultural Improvement. In R. E. Evenson, D. Gollin, and V. Santaniello (eds.), *Agricultural Values of Plant Genetic Resources.* Wallingford, UK: CAB International.

United States General Accounting Office. 1997. “Information on the Condition of the National Plant Germplasm System”. U.S. Department of Agriculture. Report to Congressional Committees. October 1997.

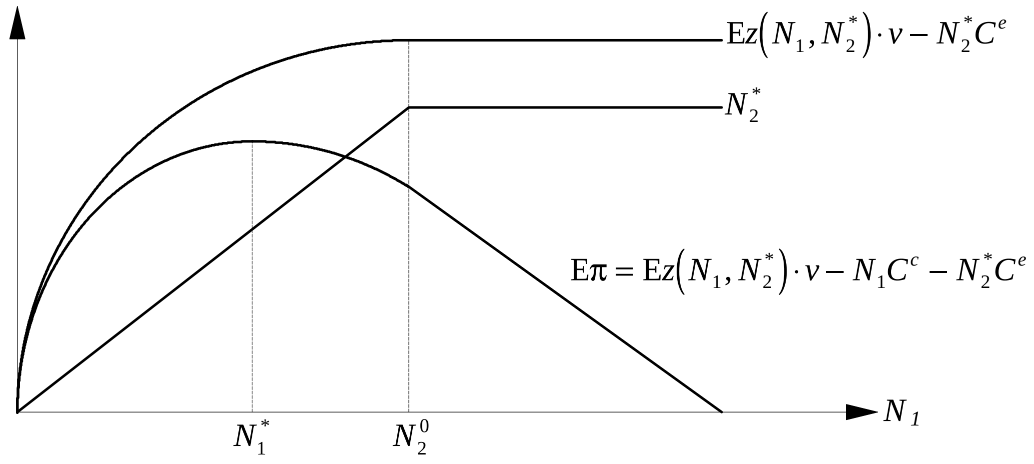

Figure 1. Stage 2 and Two-Stage Profit when N_2 is Chosen Optimally (Evenson and Lemarie, 1998)

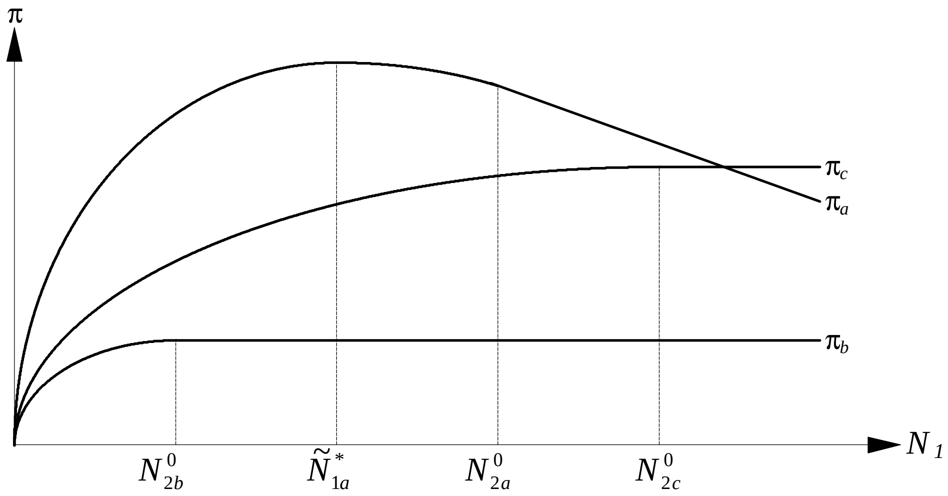

Figure 2. Trait Level Profit Function (Evenson and Lemarie, 1998)