SEARCHING FOR UNCERTAIN BENEFITS AND THE CONSERVATION OF BIOLOGICAL DIVERSITY

Stephen Polasky
Andrew Solow
James Broadus

Working Paper No. 208
February 1992

SEARCHING FOR UNCERTAIN BENEFITS AND THE CONSERVATION OF BIOLOGICAL DIVERSITY

Stephen Polasky

Department of Economics

Boston College

Andrew Solow

Marine Policy Center

Woods Hole Oceanographic Institution

James Broadus

Marine Policy Center

Woods Hole Oceanographic Institution

A common argument for protecting plant and animal species from extinction is that one or more of the preserved species may provide a benefit in the future (e.g., a cure for a disease). In this paper, we develop a simple search model that incorporates substitutability between species. Under this model, the value of a set of species depends not only on its cardinality but also on its diversity. This provides a utilitarian basis for the conservation of biological diversity.

1. INTRODUCTION

The value of protecting plant and animal species from extinction arises from at least two distinct sources. Many species have direct use value, either as productive resources (e.g., wheat) or for cultural or aesthetic reasons (e.g., giraffes). Value may also arise, however, from the relationship among species. For example, a set of species that is biologically diverse may be more valuable than a set of species that is not. In general terms, the diversity of a set refers to the joint dissimilarity of its elements. Solow, et al. (1992) considered the problem of allocating conservation resources across species to maximize measures of pure biological diversity. The problem of measuring diversity was also addressed by Weitzman (1992). The proposed measures were taken to represent a generalized preference for diversity. Under an alternative framework, the value of diversity arises from direct utilitarian considerations. For example, one justification for species conservation is that some species may yield a drug instrumental in fighting cancer or some other disease.

In this paper, we develop a simple search model to illustrate how a preference for diversity arises in such a situation. Although we will focus on the problem of finding a cure for a disease, this should be understood as a metaphor for a more general benefit. Under our model, species that are close relatives tend to be substitutes for each other (i.e., they tend to be either both cures or both not cures). As a result, a set that contains more diverse species has a higher probability of containing a cure than one containing less diverse species.

Brown and Goldstein (1984) considered the value of maintaining distinct

populations of a species to increase the probability of species survival in the face of some future threat (e.g., pest or drought). They modelled each population as an independent trial with equal probability of success. It follows that the probability of maintaining the species increases with the number of conserved populations. The independence assumption may be unrealistic, since populations that are similar genetically may exhibit similar qualities. The model described below seeks to account for this dependence.

2. THE MODEL

Suppose that it is thought that one or more of a set of species may be of future benefit. For example, one or more of these species may contain a cure for a particular disease. Each of the species is potentially subject to extinction. Limited resources are available for species conservation. One possible objective of conservation is to maximize the probability of obtaining a cure. Broadly speaking, we assume that species that are genetically similar have similar medicinal qualities.

Let $T = (s_1, \ldots, s_n)$ be the total set of species under consideration and let $D = [d_{ij}]$ be an n-by-n matrix of pairwise distances between the elements of T. We will assume that these distances are metric so that the species can be represented by points in R^q ($q \leq n - 1$). In practice, metric approximations to non-metric distances can be based on non-metric scaling (e.g., Kruskal, 1964). In our model, a species is a cure if it falls within a sphere of unknown radius R and unknown center X. Let S be a subset of T. Under this framework, the value of S is related to the probability that it contains a cure.

For a fixed value $R = r$, let $V(r; S)$ be the volume of the region within

distance r of at least one element of S and, similarly, let V(r; T) be the volume of the region within distance r of at least one element of T. If we assume that X is uniformly distributed over R^q, then the quantity:

$$p(r;\ S,\ T) = V(r;\ S)\ /\ V(r;\ T) \qquad (1)$$

is the conditional probability that S contains a cure given that T contains a cure and R = r. As long as X has an improper distribution (i.e., all finite subsets of R^q have zero measure), some scaling is necessary to compare different subsets of T. The denominator in (1) provides a scaling that is readily interpretable: p(r; S, T) measures the coverage of T by S.

As a simple example, consider Figure 1. The set T consists of species 1 through 5 lying in R^2 and the subset S consists of species 1 and 2. For $r < \min [d_{ij}/2]$, p(r; S, T) is the ratio of the number of species in S to the number of species in T (i.e., 2/5). This reflects the fact that, for small r, the species are independent. As r increases beyond $\min [d_{ij}]$, the rate of growth of V(r; S) is greater than that of V(r; T) and p(r; S, T) increases. Finally, for large r, p(r; S, T) approaches 1, reflecting the nearly total overlap of V(r; S) with V(r; T). In general, if the elements of a subset S are, on average, less dissimilar than the elements of T, then over some range of r the rate of growth of V(r; T) will be greater than that of V(r; S) and p(r; S, T) will decline. Eventually, however, V(r; S) will grow at a more rapid rate than V(r; T) and p(r; S, T) will increase.

If it is known that R = r, then the values of two subsets S_1 and S_2 can be ranked by comparing $p(r;\ S_1,\ T)$ and $p(r;\ S_2,\ T)$. In the case with unknown R, it may happen that one of the subsets dominates the other in the sense that $p(r;\ S_i,$

$T) \geq p(r;\ S_j,\ T)$ for all values of r. To compare subsets when R is unknown and neither subset dominates, it is necessary to specify a distribution for R. In that case, the subsets can be compared using:

$$p(S,\ T) = \int p(r;\ S,\ T)\ f(r)\ dr \qquad (2)$$

where f(r) is the probability density function of R. This approach is illustrated in the next section.

3. AN APPLICATION

Rodman (1991) discussed the taxonomy of 26 species of plants that produce glucosinolates (sulfur-containing compounds related to mustard oils). In this section, we apply the approach outlined above to the conservation of these species. This application is for the purposes of illustration alone and we make no claim as to its biological relevance. The species are listed in Table 1. Rodman (1991) constructed a set of distances by applying principal coordinate analysis (Gower, 1966) to a similarity matrix based on 93 measured characteristics. Figure 2 shows the locations of the species with respect to the first two principal axes.

In Figure 3(a), $p(r;\ S,\ T)$ is plotted for S_1 = (DIL, CEN, KOE) and S_2 = (PAS, CEN, KOE). The subset S_2 dominates S_1 because PAS is more distant from CEN and KOE than is DIL. In Figure 3(b), $p(r;\ S,\ T)$ is plotted for S_1 = (FLA, RES, GER) and S_2 = (FLA, RES, GER, OXA). Although S_2 dominates S_1, the curves are virtually indistinguishable, except for very small r. This reflects the high degree of overlap between GER and OXA. In addition, note that the curves in Figure 3(b)

dominate those in Figure 3(a), reflecting their greater dissimilarity.

In Figure 3(c), $p(r; S, T)$ is plotted for S_1 = (DIL, LIM, GER) and S_2 = (LIM, SAP, CEL, PEN). In this case, neither subset dominates. For small r, the larger set S_2 dominates the more dissimilar set S_1. For large r, however, the reverse is true. Since neither subset dominates the other, we compare them by calculating $p(S, T)$ for $f(r) = b \exp(-b\, r)$ for two choices of b: $b = 1.0$ and $b = 0.2$. For $b = 1.0$, $p(S_1, T) = 0.15$ and $p(S_2, T) = 0.19$, while for $b = 0.2$, these values are 0.36 and 0.33, respectively. The smaller value of b places greater weight on large values of r so that dissimilarity is weighed more heavily than cardinality.

4. IMPLICATIONS FOR DIVERSITY

The previous section shows qualitatively how a preference for diversity can arise from utilitarian considerations. Preserving species that are dissimilar tends to contribute more to the probability of finding a cure than preserving species that are similar. In this section, we explore more formally the implications of the model for the conservation of diversity.

Let $D(S)$ be a measure of the diversity of a set S. Any reasonable measure must satisfy the following three properties. First, if S' is a subset of S, then $D(S') \leq D(S)$. We will call this property monotonicity in number. Second, for a single species s_o, $D(S + s_o) = D(S)$ if and only if $d_{oi} = 0$ for some s_i in S. Weitzman (1992) referred to this property as twinning. Third, if the distance matrix for set S_1 is $[d_{ij}]$ and the distance matrix for S_2 is $[d_{ij} + e_{ij}]$, where $e_{ij} \geq 0$ for all i and j, then $D(S_1) \leq D(S_2)$. This property, which we will call monotonicity in distance, ensures that $D(S)$ does not decrease as the disimilarity

of the elements of S increases unambiguously. Clearly, p(r; S, T) satisfies these properties and, apart from its interpretation as a conditional probability, can be considered as a measure of diversity.

The three properties listed above do not define a unique measure. In ranking different subsets of species, different diversity measures tend to agree, although this is not always the case. Weitzman (1992) proposed a diversity measure by defining the distance between a point s_o and a set S as:

$$d(s_o, S) = \min_{s_i \in S} [d_{oi}] \tag{3}$$

The diversity measure is then given by:

$$W(S) = \max_{s_i \in S} [W(S - s_i) + d(s_i, S - s_i)] \tag{4}$$

with W(S) = 0 if S is composed of a single species. It is also easy to show that the measure W(S) satisfies the three properties listed above.

In Table 2, the values of W(S) are given for the examples considered in Section 3. The measures are in agreement in the first two cases. In the third, Weitzman's measure favors diversity over cardinality. The measures can, however, give contradictory results. For example, in Figure 4, p(r; S, T) is plotted for S_1 = (FLA, EUP, GER) and S_2 = (FLA, EUP, RES). In this case, S_2 dominates S_1. On the other hand, $W(S_1) = 30.02 > W(S_2) = 27.48$. If S consists of three species, W(S) is the sum of the largest and smallest pairwise distances. For this example, the largest pairwise distance is greater in S_1 than in S_2, while the smallest is the same. Unlike W(S), p(r; S,T) depends upon all three distances.

In this example, the subsets differ primarily in the middle pairwise distance, which is larger in S_2 than in S_1.

5. DISCUSSION

The key result of our model is that a preference for diversity can arise from utilitarian considerations when closely related species are regarded as partial substitutes. When this is not the case, then each species is in some sense unique and the value of a set of species depends only on its cardinality. In other words, if all species are unique, then all sets of species of equal size are equally diverse. The conclusion that the importance of diversity arises from substitutability between species does not depend on the specific model used here, but has broader generality.

In the stylized examples of Section 3, it was assumed implicitly that all species that were not conserved became extinct and all species that were conserved did not. In reality, neither of these assumptions is true. The approach can be easily modified to account for general extinction probabilities (e.g., Solow, et al., 1992). In that case, a conservation program operates by decreasing extinction probabilities for selected species (e.g., by preserving habitat).

To go beyond measuring the diversity of a set of species and to assign a value to its conservation, it is necessary to attach a value to diversity. In the context of finding a cure for a disease, the value of a set of species is simply the product of the probability of finding a cure among the set and the value of finding a cure. This would involve converting p(S, T) into an unconditional probability that S contains a cure. Despite this and other

potential difficulties, this problem seems to be more straightforward than attaching value in the absence of a utilitarian objective - for example, in conserving diversity for its own sake.

REFERENCES

Brown, G.M. and Goldstein, J.H., "A Model for Valuing Endangered Species," Journal of Environmental Economics and Management, 1984, 11, 303-309.

Gower, J.C., "Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis," Biometrika, 1966, 53, 325-338.

Kruskal, J.B., "Nonmetric Multidimensional Scaling: A Numerical Method," Psychometrika, 1964, 29, 115-129.

Rodman, J.E., "A Taxonomic Analysis of Glucosinolate-Producing Plants, Part 1: Phenetics," Systematic Botany, 1991, 16, 598-618.

Solow, A., Polasky, S., and Broadus, J., "On the Measurement of Biological Diversity," Journal of Environmental Economics and Management, 1992, in press.

Weitzman, M., "On Diversity," Quarterly Journal of Economics, in press.

TABLE 1

Taxa of glucosinolate-producing plants and putative relatives.

Taxon	Code	Taxon	Code
Akaniaceae	AKA	Tovariaceae	TOV
Bataceae	BAT	Tropaeolaceae	TRO
Brassicaceae	BRA	Balsaminacea	BAL
Bretschneideraceae	BRE	Celastraceae	CEL
Capparaceae	CAP	Centrospermae	CEN
Cariacaceae	CAR	Dilleniaceae	DIL
Drypetes	DRY	Euphorbiaceae	EUP
Gyrostemonaceae	GYR	Flacouticeae	FLA
Limnanthaceae	LIM	Geraniaceae	GER
Moringaceae	MOR	Koeberliniaceae	KOE
Pentadiplandraceae	PEN	Oxalidaceae	OXA
Resedaceae	RES	Passifloraceae	PAS
Salvadoraceae	SAL	Sapindaceae	SAP

TABLE 2

Values of W(S) for subsets considered in Section 3.

S	W(S)
(DIL, CEN, KOE)	17.94
(PAS, CEN, KOE)	22.40
(FLA, RES, GER)	38.82
(FLA, RES, GER, OXA)	39.56
(DIL, LIM, GER)	23.40
(LIM, SAP, CEL, PEN)	13.64

FIGURE CAPTIONS

Figure 1. In this example, T consists of species 1-5 with locations (0,2), (3,2), (0,0), (1,0), (2,0). The subset S consists of species 1 and 2. For r = 0.9, V(r; T) is the area enclosed in the circles and V(r; S) is the hatched area. The function p(r; S,T) is plotted below.

Figure 2. Locations of the species in Table 1 in terms of the first two principal axes (Rodman, 1986).

Figure 3. $p(r; S_1, T)$ (solid) and $p(r; S_2, T)$ (dotted) for:

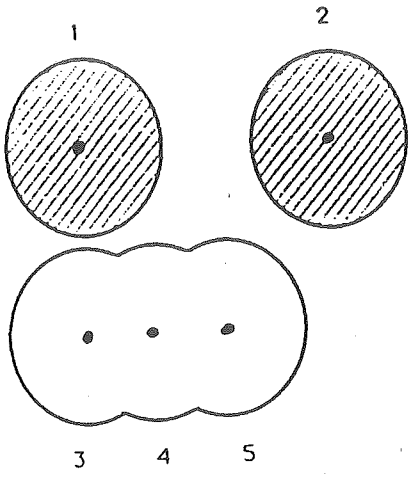
(a) S_1 = (DIL, CEN, KOE), S_2 = (PAS, CEN, KOE)

(b) S_1 = (FLA, RES, GER), S_2 = (FLA, RES, GER, OXA)

(c) S_1 = (DIL, LIM, GER), S_2 = (LIM, SAP, CEL, PEN).

Figure 4. $p(r; S_1, T)$ (solid) and $p(r; S_2, T)$ for S_1 = (FLA, EUP, GER) and S_2 = (FLA, EUP, RES).

Figure 1.



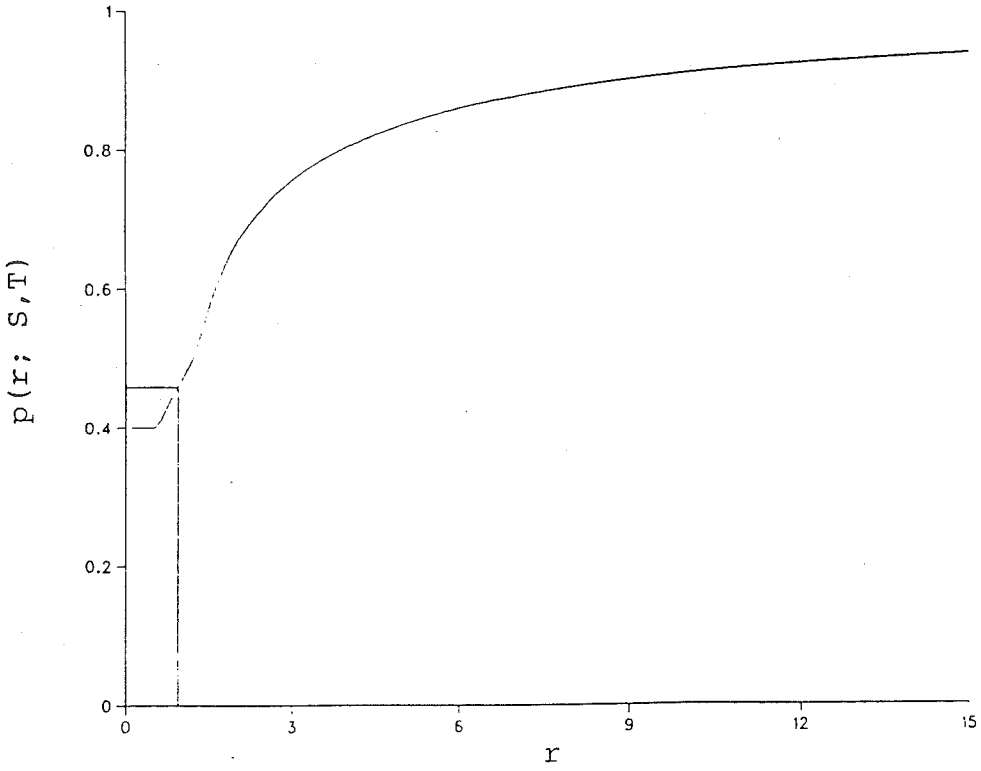


Figure 2.

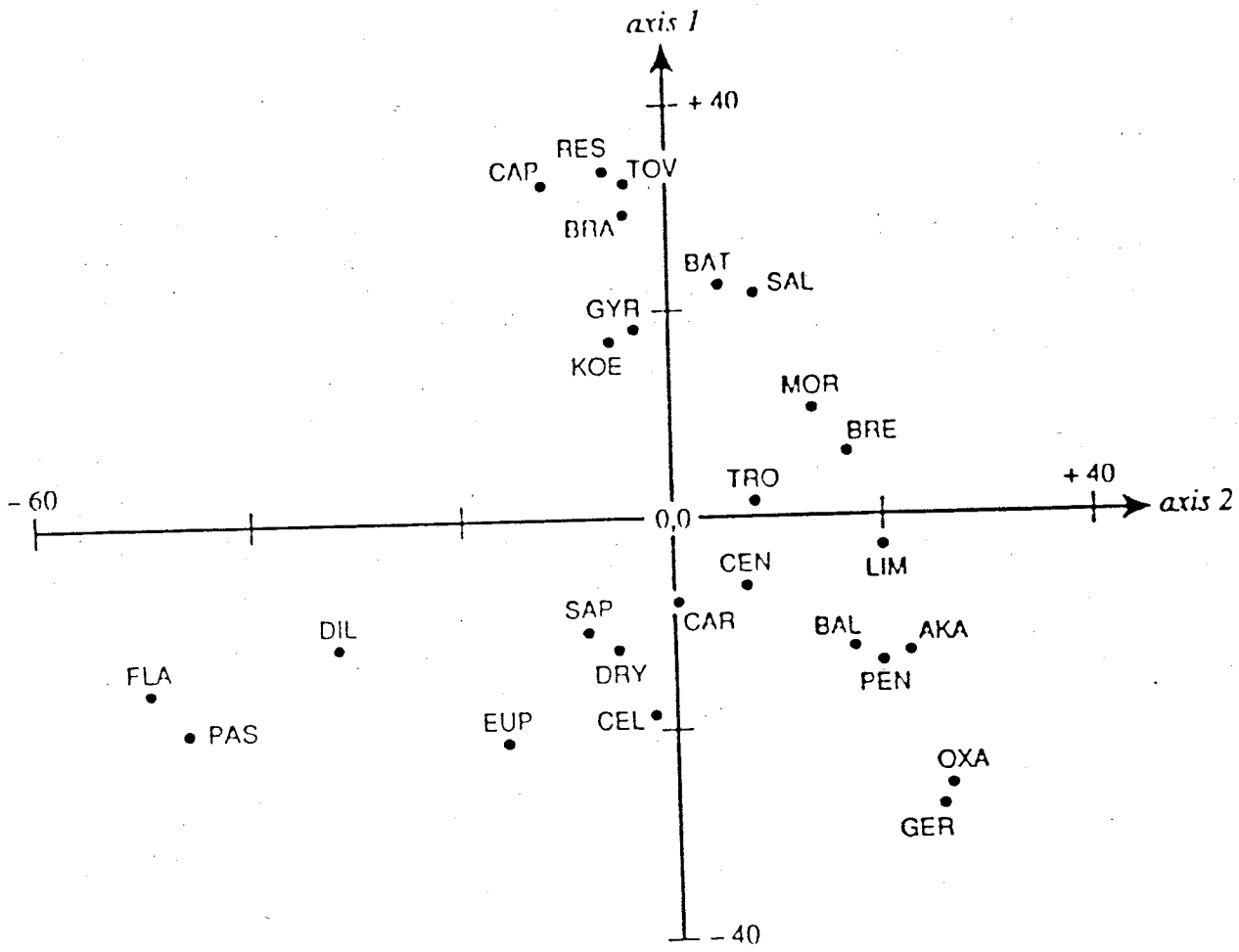

axis 1
+ 40
RES
CAP
TOV
BOA
BAT
SAL
GYR
KOE
MOR
BRE
TRO
+ 40
axis 2
- 60
0,0
CEN
LIM
SAP
CAR
DIL
BAL
AKA
FLA
DRY
PEN
EUP
CEL
PAS
OXA
GER
- 40

Figure 3a.

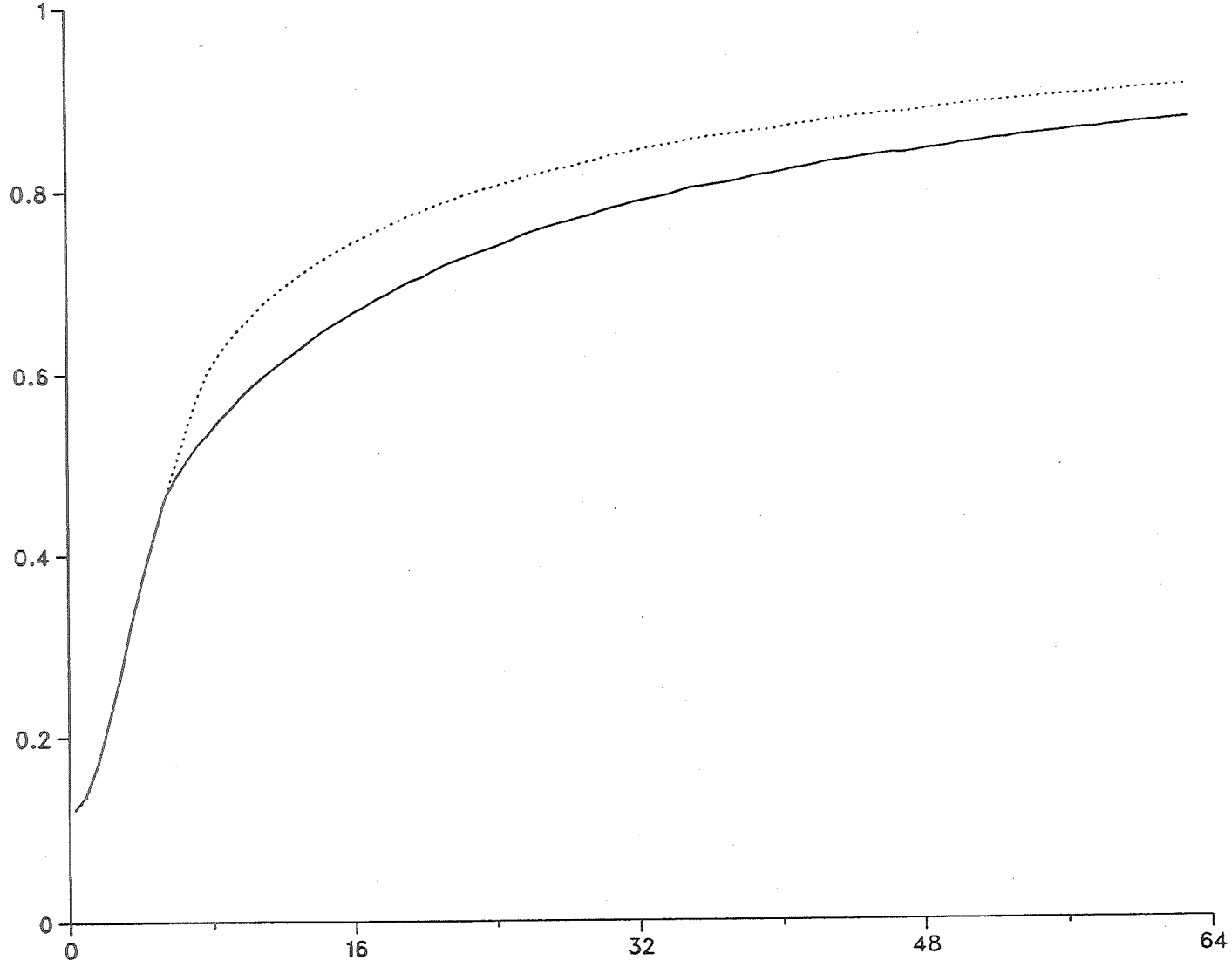


Figure 3b.

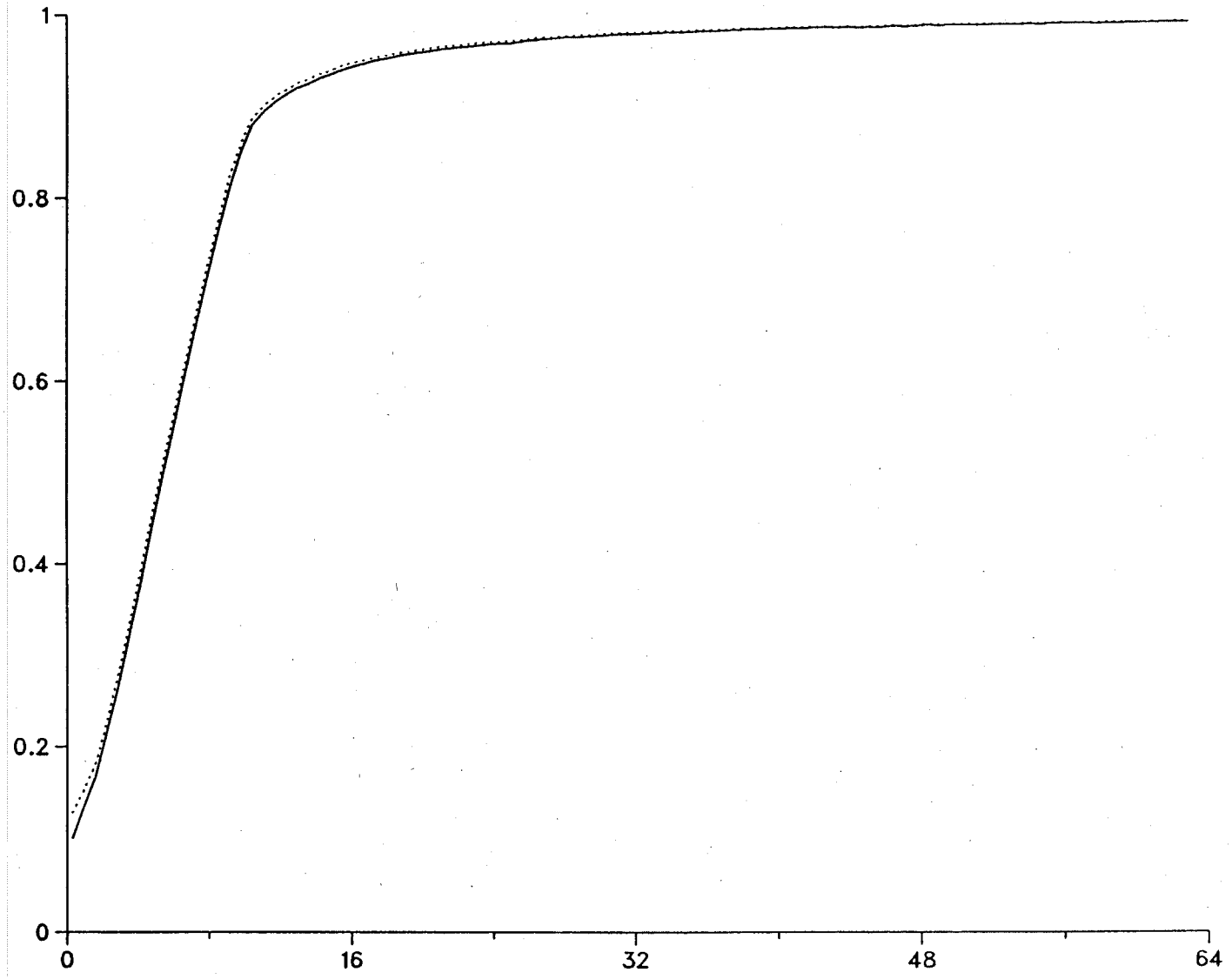


Figure 3c.

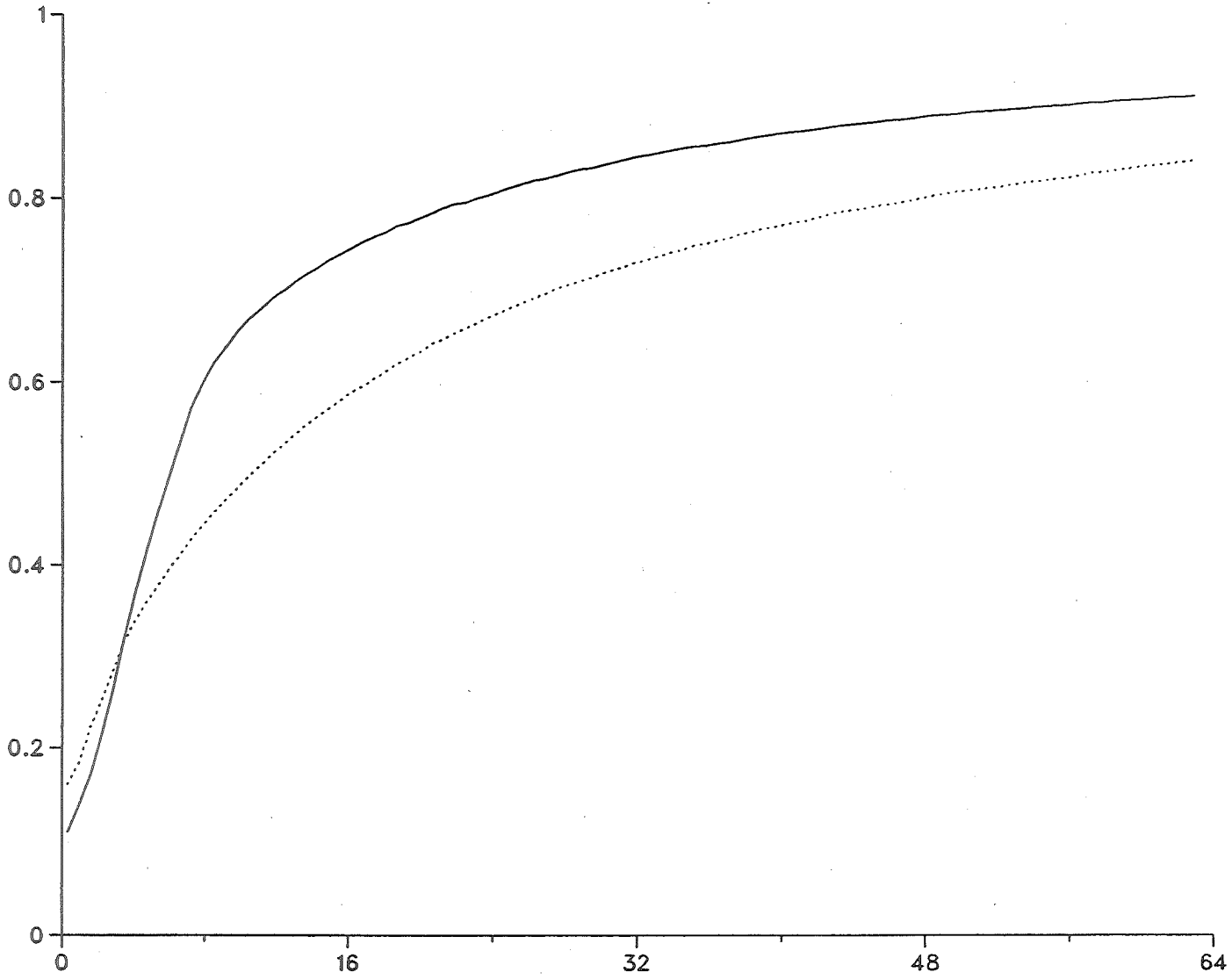


Figure 4.