CAE Working Paper #02-16

On Literacy Rankings

by

Tapan Mitra

November 2002

.

On Literacy Rankings*

Tapan Mitra†

Abstract

This paper is concerned with the issue of characterizing the situations in which all the literacy indices, consistent with a set of reasonable axioms, would provide the same ranking of societies. It is shown that a theory, analogous to that developed for the Lorenz order in the study of income inequality, can be obtained in the study of literacy, by extending the standard mathematical theory relating gauge functions to convex functions, and the theory of majorization.

Journal of Economic Literature Classification Numbers: C60, D60, I20, I31.

Keywords: Literacy, Externality, Gauge Functions, Convex Functions, Majorization.

*I am grateful to Paula Valenti for her comments and suggestions on an earlier version of this paper.

†Department of Economics, 448 Uris Hall, Cornell University, Ithaca, NY 14853; Email: tm19@cornell.edu

1 Introduction

The importance of literacy in the process of economic development of a society is well recognized. The standard measure of literacy of a society is the percentage of literates in the adult population, called the *literacy rate*. This measure has been the subject of careful scrutiny in recent years. Basu and Foster (1998) have argued that literates in a household provide a positive externality to the illiterates in that household. Thus, an illiterate person in a household which has some literate person, is *more literate* than an illiterate person in a household which has no literate person. Consequently, the percentage of literates in a household does not truly capture the literacy of the household. Some account needs to be taken of those illiterates who benefit from the presence of literates in the same household.

New indices of literacy have been proposed, taking this aspect into account, by Basu and Foster (1998), and Subramanian (2001), among others. They differ in the specification of the nature of the externality. Recently, Valenti (2002) has completely characterized the class of *externality functions* which are consistent with a set of reasonable axioms on a literacy index. In comparing the literacy of two societies by literacy measures (which obey such a set of axioms), one might get conflicting results, depending on the externality function that has been specified. That this is not just a theoretical issue but one of practical significance is demonstrated in Valenti (2002) by looking at some literacy data of South Africa.

This paper is primarily concerned with the issue of characterizing the situations in which *all* the literacy indices, consistent with a set of reasonable axioms, would provide the *same* ranking of societies. That is, I am interested in characterizing those comparison situations, in which the literacy rankings are *independent* of the particular choice of the externality function. It will be shown that a theory, analogous to that developed for the Lorenz order in the study of income inequality, can be obtained in the study of literacy.

To indicate a flavor of the kind of comparisons that might be settled unequivocally, we present a simple example. Let us denote a *household* by a vector (r, n) of two integers, the number of literates (r), and the total number of persons in the household (n).[1] A *society* is a collection of such households.

[1] One would typically count only the *adult* members of the household, and I take this for granted henceforth, without further mention of the word "adult".

We compare two societies, A and B, which are defined as follows:

$$A = \{(r_1, n_1), (r_2, n_2), (r_3, n_3)\} = \{(3, 10), (4, 10), (6, 10)\} \tag{1.1}$$

$$B = \{(r'_1, n'_1), (r'_2, n'_2), (r'_3, n'_3)\} = \{(1, 10), (5, 10), (7, 10)\} \tag{1.2}$$

Notice that there is no *obvious* sense in which society A can be called more literate than society B, or the other way round. It is not true, for example, that every household in society A is more literate (in the sense of a higher literacy rate, or equivalently in the sense of having a higher number of literates, since each household in each society has exactly the same number of persons) than every household in society B, or the other way round. It is not even true that such a dominance holds in the sense that, after ordering the households in each society from the least literate to the most literate (they are already presented in (1.1) and (1.2) so ordered), household k of society A is more literate than household k of society B, for $k = 1, 2, 3$, or the other way round. Nevertheless, we will show that every literacy index, consistent with the set of axioms mentioned above, would rank society A at least as high as society B. This follows from the simple observation that the vector of literacy rates of society A, (0.3, 0.4, 0.6), *super- majorizes* the vector of literacy rates of society B, (0.1, 0.5, 0.7); that is, $0.3 > 0.1$, $0.3 + 0.4 > 0.1 + 0.5$, and $0.3 + 0.4 + 0.6 = 0.1 + 0.5 + 0.7$.

The theory, which allows us to make such comparisons, is somewhat more complicated than the example might suggest. The complications arise from two sources. First, the total number of persons in the two societies need not be the same. Second, even if they were, the number of persons in each household (in either society) need not be the same. But, these aspects can be readily accommodated (as we shall see below) by defining the vectors to be compared for each society appropriately.

The axioms on literacy indices are formally presented in Section 3, along with a discussion which compares them to similar axioms presented in the literature. The basic theory of unequivocal literacy rankings is presented in Theorem 1 of Section 4. What is crucial for the theory is the result that the literacy axioms restrict the class of acceptable literacy indices to those for which literacy of a household is a "concave" and "monotone increasing" function of its literacy rate.[2]

[2] The quotes are to remind us that the domain of the function is not an interval, unlike the usual setting for concave and monotone functions on subsets of the real line.

We view this result as an extension of the standard mathematical theory relating *gauge functions* to *convex functions*. Equipped with this result, one can develop a theory of super-majorization along the lines of the well-known Tomic-Weyl theorem. The standard theory cannot be directly applied, because the number of literates or illiterates in a household are integers, and so variables like literacy rates take on rational values only. Thus the natural domain of measures of literacy of societies is not a convex set. The relevant material, covering this technical part of the paper, is presented in Section 2, and the proofs of the results of Section 2 are presented in Section 6.

Section 5 presents a characterization of the class of externality functions consistent with the set of axioms proposed in Section 3. An alternative characterization, due to Valenti (2002), is also discussed. Her characterization result is compared to ours through two propositions. These are based on two technical lemmas, the proofs of which are again presented in Section 6.

It would appear from our analysis that the crucial variable of interest in literacy rankings, even after taking the externality aspect of literacy into account, is the literacy rate, although now at the micro-level of the household.

2 Preliminaries

2.1 Super-Additive and Concave Functions

In the standard mathematical theory of convex functions, it is known that *gauge functions* defined on convex cones are also convex functions (see, for example, Rosenbaum (1950), Green (1954), and Roberts and Varberg (1973)). Thus, for example, if F is a gauge function on $\mathbb{R}_+^2$ (that is, F is sub-additive and homogeneous of degree one), then F is convex on $\mathbb{R}_+^2$. This means that the function $f(z) = F(z, 1)$ is a convex function of z, for $z \geq 0$. A similar theory can be developed for functions, defined on a subset of $\mathbb{M}^2$, where $\mathbb{M}$ is the set of non-negative integers.[3]

Let $\mathbb{N}$ denote, as usual, the set of natural numbers $\{1, 2, 3, ...\}$, and let $\mathbb{M}$ denote the set $\{0, 1, 2, 3, ...\}$. We define the set $\mathbb{X} = \{(x, y) : x \in \mathbb{M}, y \in \mathbb{N}$, and $x \leq y\}$. Let $\mathbb{Q}$ denote, as usual, the set of rational numbers. We denote

[3] Properties of such functions, discussed below, are called by familiar names such as "homogeneity of degree one", "super-additivity", even though the domains are restricted. This is deliberately done to show the similarity of the theory presented here to the one usually developed for domains not so restricted.

by $\mathbb{Q}_+$ the set of non-negative rational numbers $\{z = p/q$, where $p \in \mathbb{M}$, and $q \in \mathbb{N}\}$, and by $\mathbb{Y}$ the set $\mathbb{Q} \cap [0,1]$; the set $\mathbb{Q} \cap (0,1)$ is denoted by Y.

Let F be a function from $\mathbb{X}$ to $\mathbb{R}$. Consider the following properties, which may be satisfied by such a function:

Homogeneity of Degree One(H): If $(x,y) \in \mathbb{X}$, and $t \in \mathbb{N}$, then $F(tx,ty) = tF(x,y)$.
Super Additivity(SA): If $(x,y) \in \mathbb{X}$, and $(x',y') \in \mathbb{X}$, then $F(x+x', y+y') \geq F(x,y) + F(x',y')$.
X-Monotonicity(XM): If $(x,y) \in \mathbb{X}$, and $(x',y) \in \mathbb{X}$, with $x' > x$, then $F(x',y) > F(x,y)$.
Origin and Scale(OS): $F(1,1) = 1, F(0,1) = 0$.

Given any $F : \mathbb{X} \to \mathbb{R}$, satisfying **H**, we can define $f : \mathbb{Y} \to \mathbb{R}$ as follows:

$$f(z) = F(p,q)/q \quad where \quad z = (p/q), with \quad (p,q) \in \mathbb{X} \tag{2.1}$$

Note that the function, f, is well-defined by (2.1). For, if $z = (p/q)$, with

$(p,q) \in \mathbb{X}$, and if z is also equal to (p'/q'), with $(p',q') \in \mathbb{X}$, then we have $F(p',q')/q' = qF(p',q')/qq' = F(qp',qq')/qq'$ (by **H**) $= F(pq',qq')/qq' = q'F(p,q)/qq'$ (by **H**) $= F(p,q)/q$.

Suppose $F : \mathbb{X} \to \mathbb{R}$ satisfies properties **H, SA, XM** and **OS**. We can show that the function, $f : \mathbb{Y} \to \mathbb{R}$ must satisfy the following properties[4]:

Concavity(C): If $z, z' \in \mathbb{Y}$, and $t \in Y$, then $f(tz + (1-t)z') \geq tf(z) + (1-t)f(z')$.
Monotonicity(M): If $z, z' \in \mathbb{Y}$, and $z' > z$, then $f(z') > f(z)$.
End Point Condition(E): $f(0) = 0$ and $f(1) = 1$.

Conversely, given any $f : \mathbb{Y} \to \mathbb{R}$, we can associate with it a function, $F : \mathbb{X} \to \mathbb{R}$, defined as follows:

$$F(p,q) = qf(p/q) \quad where \quad (p,q) \in \mathbb{X} \tag{2.2}$$

[4]Here, as indicated in an earlier footnote, the names "concavity" and "monotonicity" are deliberately used, even though the domain of f is not an interval of the real line, unlike the usual setting for concave and monotone functions of a real variable.

If $f : \mathbb{Y} \to \mathbb{R}$ satisfies the properties **C, M** and **E**, then we can show that

the associated function, $F : \mathbb{X} \to \mathbb{R}$, defined by (2.2), satisfies properties **H, SA, XM** and **OS**.

We summarize the above discussion, for ready reference, in the following proposition.

Proposition 1 *(i) Suppose $F : \mathbb{X} \to \mathbb{R}$ satisfies properties **H, SA, XM** and **OS**. Then, the function, $f : \mathbb{Y} \to \mathbb{R}$, defined by (2.1), satisfies properties **C, M** and **E**. (ii) Suppose $f : \mathbb{Y} \to \mathbb{R}$ satisfies properties **C, M** and **E**. Then, the function, $F : \mathbb{X} \to \mathbb{R}$, defined by (2.2), satisfies properties **H, SA, XM** and **OS**.*

2.2 Majorization Theory

In the standard theory of convex functions, defined on convex subsets of the real line, the class of functions which are convex and monotone is of special significance, because a very useful *majorization* theory can be developed for it. The main result of this theory is known as the Tomic-Weyl theorem; see, for example, Mitrinovic-Vasic (1970, p.165) for this result, as well as some of the important variations of it. Something similar can be achieved for functions, $f : \mathbb{Y} \to \mathbb{R}$, which satisfy properties **C** and **M**. We summarize this theory by stating three lemmas, followed by a Proposition.

The three lemmas are analogous to the basic results for concave functions defined on an interval of the real line; see, for example, Nikaido (1968, Theorem 3.15, p.47) for purpose of comparison. The first lemma relates to the comparison of slopes of chords.

Lemma 1 *Let $f : \mathbb{Y} \to \mathbb{R}$ be a function satisfying property **C**. Then:*

(i) If $a, b, c \in \mathbb{Y}$, and $a < b \leq c$, then $[f(b) - f(a)]/[b - a] \geq [f(c) - f(a)]/[c - a]$.
(ii) If $a, b, c \in \mathbb{Y}$, and $a \leq b < c$, then $[f(c) - f(a)]/[c - a] \geq [f(c) - f(b)]/[c - b]$.
(iii) If $a, b, c \in \mathbb{Y}$, and $a < b < c$, then
$[f(b) - f(a)]/[b - a] \geq [f(c) - f(a)]/[c - a] \geq [f(c) - f(b)]/[c - b]$.

The second lemma is analogous to the result for concave functions (defined on an interval of the real line) that the right-hand derivative is well-defined in the interior of the interval and is non-increasing.

Lemma 2 *Let $f : \mathbb{Y} \to \mathbb{R}$ be a function satisfying property* ***C****. Then, the function, $g : Y \to \mathbb{R}$ is well-defined by:*

$$g(y) \equiv \lim_{\substack{\varepsilon \downarrow 0 \\ y+\varepsilon \in Y}} \{[f(y+\varepsilon) - f(y)]/\varepsilon\} \tag{2.3}$$

Further, g is monotone non-increasing on Y.

The third lemma is analogous to the result for concave functions (defined on an interval of the real line) comparing the slope of a chord with the (right-hand)derivative at an end-point of the chord.

Lemma 3 *Let $f : \mathbb{Y} \to \mathbb{R}$ be a function satisfying property* ***C****. Then for $x \in \mathbb{Y}$, $y \in Y$, we have:*

$$f(x) - f(y) \leq g(y)(x - y) \tag{2.4}$$

The above lemmas, together with Abel's inequality, can be used to establish the result (in our context) analogous to the Tomic-Weyl theorem of majorization theory, which we now proceed to state.

Proposition 2 *Let z and z' be vectors in $\mathbb{Y}^n$, such that $z_1 \leq z_2 \leq \ldots \leq z_n$, and $z'_1 \leq z'_2 \leq \ldots \leq z'_n$. Let $f : \mathbb{Y} \to \mathbb{R}$ be a function satisfying properties* ***C*** *and* ***M****. Suppose, for each integer $k \in \{1, 2, ..., n\}$, we have:*

$$\sum_{i=1}^{k} z'_i \leq \sum_{i=1}^{k} z_i \tag{2.5}$$

Then, we must have:

$$\sum_{i=1}^{n} f(z'_i) \leq \sum_{i=1}^{n} f(z_i) \tag{2.6}$$

Remark:

In Proposition 2, if there is a strict inequality in (2.5) when $k = n$, then we can infer that (2.6) must also hold with a strict inequality, by using property **M** of the function f.

3 Literacy Indices

3.1 Decomposable Literacy Indices

A *household* is a pair $(r, n) \in \mathbb{X}$. Here n is to be interpreted as the total number of individuals in the household, and r is to be interpreted as the number of *literate* individuals in the household. Thus, $s = r - n$ is the number of *illiterate* individuals in the household.

A *society* is a (non-empty) collection of households; a society consisting of $m \in \mathbb{N}$ households is denoted by the set $\{(r_1, n_1), ..., (r_m, n_m)\}$.

A literacy index is a function from the set of all possible societies to the reals. To formalize this, we define the set of all possible societies to be:

$$\mathbb{U} = \bigcup_{k=1}^{\infty} \mathbb{X}^k \tag{3.1}$$

and we define a *literacy index* as a function, $L : \mathbb{U} \to \mathbb{R}$. We will confine our attention exclusively to *decomposable literacy indices*; that is, those which satisfy:

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) = \sum_{i=1}^{n} (n_i/n) L(\{(r_i, n_i)\}) \tag{3.2}$$

where $n = (n_1 + \cdots + n_m)$.

Given our restriction to decomposable literacy indices, it is clear that any axiom system on literacy indices of a society can be expressed as an axiom system on the literacy indices of the single-household society. This allows us to focus on the micro-level, and see what reasonable restrictions one might wish to impose on the literacy index of a household. We shall impose four such axioms, and provide some justification for each one.

Remarks:

(i) The restriction of the exercise to decomposable literacy indices is a serious one. There is little in the way of a theoretical justification for this restriction. This is especially true in a context in which externalities of literates on illiterates within a household is being emphasized, for the restriction rules out any inter-household externality of literates on illiterates. One would think that such externalities are prevalent, even when they are not formalized in the institution of a school. Of course, formally, one can think of a "household" more broadly as the unit in which the externalities are prevalent. But, this creates problems in defining precisely the boundaries of "households", and therfore in using the theory on standard household data. From the practical point of view, one might argue that decomposable indices are the only indices which stand a chance of being used by policy makers.

(ii) The depiction of a household as a pair (r, n) hides a lot of relevant information about the household. For example, there is no special significance attached to whether the father (or the mother) in the household is literate. A whole range of policy issues which are tied to such aspects of the household cannot, therefore, be addressed in our framework. However, the representation of a household as $(1, 3)$ in my notation (for example), conveys all the relevant information that is conveyed by its alternative representation, used in Basu and Foster(1998), as $(0, 0, 1)$ [or equivalently as $(0, 1, 0)$ or $(1, 0, 0)$, using their Anonymity axiom] in which each 0 represents an illiterate and each 1 represents a literate person in the household.

3.2 Axioms on Decomposable Literacy Indices

Consider a single household, with a single literate member (and no illiterate member). If we compare this with a single household, with a single illiterate member (and no literate member), it should be obvious that any reasonable literacy index would pronounce the first one more literate than the second. If we are to assign higher numbers for higher literacy, then any reasonable literacy index would assign a higher number to the first household than to the second. Our first axiom treats these two households as "benchmarks" relative to which other households are evaluated, by assigning the number 1 to the first household, and 0 to the second household.

Axiom N (Normalization Axiom):
$L(\{(1, 1)\}) = 1; L(\{(0, 1)\}) = 0.$

Consider, next, a comparison of one household, (r, n), with another household, (r', n), where the number of literates in the second household (r') exceeds the number of literates in the first household (r) , while the total number of individuals in both households is the same. It should be obvious that any reasonable literacy index should assign a higher number to the second household relative to the first. This is the content of the monotonicity axiom.

Axiom RM (Monotonicity Axiom):
If $(r, n) \in \mathbb{X}$, *and* $(r', n) \in \mathbb{X}$, *and* $r' > r$, *then* $L(\{(r', n)\}) > L(\{(r, n)\})$.

We now come to an axiom, which might be viewed as an extension of the idea that there is a positive externality of literates on the illiterates in a household. Consider a society $A = \{(1, 1),\ (0, 1)\}$, with two households: a household, consisting of a literate person, and no illiterate person, and another household, consisting of an illiterate person, and no literate person. Contrast this with a society $B = \{(1, 2)\}$, consisting of a single household, obtained by merging the two households of society A into one. The presence of positive externality of literates on illiterates in the same household means precisely that a literacy index should assign at least as high a number (possibly higher) to society B relative to society A.

Notice that the literate in the first household cannot have a positive externality on the illiterate in the second household in society A. The decomposability of the literacy index rules out an inter-household externality. But, when the two individuals are part of the same household, as in society B, then the illiterate can gain from the literate.

Thus, we might view the presence of a positive intra-household externality as saying that $L(\{1+0, 1+1\}) \geq L(\{(1, 1), (0, 1)\})$. Extending this idea, we could say that if society C consists of two households, and is described by $\{(r_1, n_1), (r_2, n_2)\}$, and society $D = \{(r_1 + r_2), (n_1 + n_2)\}$ is obtained by merging the two households of society C into one, then we should have $L(D) \geq L(C)$; that is, $L(\{(r_1+r_2), (n_1+n_2)\}) \geq L(\{(r_1, n_1), (r_2, n_2)\})$. After all, the single household society of D can always function like two households living under one roof; that is, without any interaction between the literates of the first household (r_1) and the illiterates of the second household ($n_2 - r_2$), and without any interaction between the literates of the second household (r_2) and the illiterates of the first household ($n_1 - r_1$). But, in general, there is the *possibility* of these positive interactions in society D, which are absent in society C. That is, society D is capable of doing everything that society C is capable of doing, in terms of positive effects of its literates on its illiterates, and possibly more. We formalize these ideas in the positive externality axiom.

Axiom PE(Positive Externality Axiom):
If $(r_1, n_1) \in \mathbb{X}$, *and* $(r_2, n_2) \in \mathbb{X}$, *then* $L(\{(r_1 + r_2), (n_1 + n_2)\}) \geq L(\{(r_1, n_1), (r_2, n_2)\})$.

Our final axiom pertains to scale invariance. Consider society A, consisting of a single household, with one literate, and no illiterate person. Contrast this with society B, consisting again of a single household, with two literates, and no illiterate person. We agreed to assign a literacy index of 1 to society A (by **Axiom N**). It would seem reasonable to assign the same literacy index to society B. Similarly, society C, consisting of a single household, with one illiterate, and no literate person, gets a literacy index of 0 by Axiom N. And, it again seems reasonable to assign the same index of 0 to society D, consisting of a single household, with two illiterates, and no literate person. These are rather clear-cut cases of invariance of the literacy index to the scale of the household in question (in a single-household society). The next axiom postulates that such scale invariance holds for all single household societies.

Axiom SI (Scale Invariance Axiom):
If $(r, n) \in \mathbb{X}$ *, and* $k \in \mathbb{N}$ *, then* $L(\{(kr, kn)\}) = L(\{(r, n)\})$.

Remarks:
(i) The usual statement of the normalization axiom (in Basu and Foster(1998) and Valenti(2002)) is stronger than our **Axiom N**, asserting that if $n \in \mathbb{N}$, then $L(\{(n, n)\}) = 1$, and $L(\{(0, n)\}) = 0$. That is, it combines our **Axiom N** with a scale invariance property for some particular single household societies.

(ii) The Monotonicity Axiom is basically the same as the one used in Basu and Foster(1998) and Valenti(2002).

(iii) **Axiom PE** and **Axiom SI** are combined into one axiom, called the Equality Axiom, in Valenti (2002). I have found it more useful to separate the two. This makes **Axiom PE** acceptable, if one does believe in the positive externality of literates on illiterates in the same household. And, it makes **Axiom SI** acceptable, if one believes that a literacy index should be a *relative* (not an absolute) index. The interpretation given in Valenti (2002) for the Equality Axiom is somewhat different from that provided in the discussion above. Valenti (2002) would also assert a strict inequality in **Axiom PE**, when (r_1, n_1) is not proportional to (r_2, n_2). This stronger requirement rules out some literacy indices that I wish to accomodate. This is discussed in greater detail in Section 5.

(iv) The externality axiom in Basu and Foster(1998) implies **Axiom PE**, but is stronger. The first part of their externality axiom also implies **Axiom SI**.

4 Literacy Ranking of Societies

Recall that we defined the set of all possible societies to be:

$$\mathbb{U} = \bigcup_{k=1}^{\infty} \mathbb{X}^k$$

and a literacy index as a function from $\mathbb{U}$ to the reals. We now show that, given a decomposable literacy index, $L : \mathbb{U} \to \mathbb{R}$ which satisfies Axioms **N,RM,PE** and **SI**, there is a function $f : \mathbb{Y} \to \mathbb{R}$ satisfying properties **C, M** and **E**, such that for any society $\{(r_1, n_1), ..., (r_m, n_m)\} \in \mathbb{U}$, we have:

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) = \sum_{i=1}^{n} (n_i/n) f(r_i/n_i) \tag{4.1}$$

This will allow us to use our majorization result, developed in Proposition 2 of Section 2, on the literacy index, L.

To this end, given a decomposable literacy index, $L : \mathbb{U} \to \mathbb{R}$, let us associate with it a *literacy measure*, M, defined on $\mathbb{U}$ by:

$$M(\{(r_1, n_1), ..., (r_m, n_m)\}) = L(\{(r_1, n_1), ..., (r_m, n_m)\})(n_1 + \cdots + n_m) \tag{4.2}$$

Thus, M is also a real-valued function on $\mathbb{U}$. Since L is decomposable (that is, it satisfies (3.2)), M must be *additively separable*; that is, it must satisfy:

$$M(\{(r_1, n_1), ..., (r_m, n_m)\}) = \sum_{i=1}^{m} M(\{(r_i, n_i)\}) \tag{4.3}$$

We now proceed to infer properties on the literacy measure, M, associated with a decomposable literacy index, L, which satisfies Axioms **N,RM,PE** and **SI**. Using Axiom **N**, it follows that:

$$M(\{(1, 1)\}) = 1, \; M(\{(0, 1)\}) = 0 \tag{4.4}$$

Using Axiom **RM**, it follows that:

$$If \; (r, n) \in \mathbb{X}, \; and \; (r', n) \in \mathbb{X}, \; then \; M(\{(r', n)\}) > M(\{(r, n)\}) \tag{4.5}$$

Using Axiom **PE**, and the additive separability of M, we can infer that:

$$\begin{aligned} If\ (r_1, n_1), (r_2, n_2) &\in \mathbb{X}, then \\ M(\{(r_1 + r_2, n_1 + n_2)\}) &\geq M(\{(r_1, n_1)\}) + M(\{r_2, n_2)\}) \end{aligned} \tag{4.6}$$

Finally, using Axiom **SI**, we can deduce that:

$$If\ (r, n) \in \mathbb{X}, and\ k \in \mathbb{N}, then\ M(\{(kr, kn)\}) = kM(\{(r, n)\}) \tag{4.7}$$

Let us now define a function, $F : \mathbb{X} \to \mathbb{R}$ by:

$$F(r, n) = M(\{(r, n)\}) \tag{4.8}$$

Notice that the domain of F is $\mathbb{X}$, while the domain of M is the set of all societies, $\mathbb{U}$; F is defined by restricting the domain of M to single-household societies. It now follows readily from (4.4),(4.5), (4.6) and (4.7) that F, defined by (4.8), satisfies properties **H, SA, XM** and **OS** of Section 2.1.

Thus, as indicated in Section 2.1, there is a well-defined function, $f : \mathbb{Y} \to \mathbb{R}$, which satisfies:

$$f(z) = F(p, q)/q\ where\ z = p/q,\ and\ (p, q) \in \mathbb{X} \tag{4.9}$$

Further, f satisfies properties **C, M** and **E** of Section 2.1 by Proposition 1. Clearly, (4.8) and (4.9) imply that $[M(\{(r, n)\})/n]$ is a function, f, of the single variable (r/n), and that this function, f, satisfies properties **C, M** and **E**. But, by (4.2), $[M(\{(r, n)\})/n]$ is $L(\{(r, n)\})$, and so any decomposable literacy index, L, satisfying the axioms **N,RM,PE** and **SI** can be written as:

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) = \sum_{i=1}^{n} (n_i/n) f(r_i/n_i)$$

where $n = (n_1 + \cdots + n_m)$, and $f : \mathbb{Y} \to \mathbb{R}$ satisfies properties **C, M** and **E**.

We can now state and prove the principal result of this paper on literacy rankings.

Theorem 1 *Let $L : \mathbb{U} \to \mathbb{R}$ be any decomposable literacy index satisfying* ***Axioms N, RM, PE*** *and* ***SI****. Consider two societies, $A = \{(r_1, n_1), ..., (r_m, n_m)\}$, and $B = \{(r'_1, n'_1), ..., (r'_k, n'_k)\}$. Let $n = (n_1 + \cdots + n_m)$, and $n' = (n'_1 + \cdots +$*

n'_k).*Define two vectors, x and y, in $\mathbb{Y}^{nn'}$ as follows:*

$$x = (\underbrace{(r_1/n_1), \ldots, (r_1/n_1)}_{n'n_1 \text{ times}}, \underbrace{(r_2/n_2), \ldots, (r_2/n_2)}_{n'n_2 \text{ times}}, \ldots, \underbrace{(r_m/n_m), \ldots, (r_m/n_m)}_{n'n_m \text{ times}})$$
$$y = (\underbrace{(r'_1/n'_1), \ldots, (r'_1/n'_1)}_{nn'_1 \text{ times}}, \underbrace{(r'_2/n'_2), \ldots, (r'_2/n'_2)}_{nn'_2 \text{ times}}, \ldots, \underbrace{(r'_k/n'_k), \ldots, (r'_k/n'_k)}_{nn'_m \text{ times}}) \tag{4.10}$$

Denote by $\hat{x}$ the increasing rearrangement of x, and by $\hat{y}$ the increasing rearrangement of y. Assume that, for each $K \in \{1, \ldots, nn'\}$, the following inequalities hold:

$$\sum_{i=1}^{K} \hat{x}_i \geq \sum_{i=1}^{K} \hat{y}_i \tag{4.11}$$

Then, we have:

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) \geq L(\{(r'_1, n'_1), ..., (r'_k, n'_k)\}) \tag{4.12}$$

That is, society A is at least as literate as society B, according to the literacy index, L.

Proof. Given (4.11), we can apply Proposition 2 of Section 2 to obtain:

$$\sum_{i=1}^{nn'} f(\hat{x}_i) \geq \sum_{i=1}^{nn'} f(\hat{y}_i) \tag{4.13}$$

whenever $f : \mathbb{Y} \to \mathbb{R}$ satisfies properties **C** and **M**. This can be rewritten as:

$$\sum_{j=1}^{m} (n'n_j) f(r_j/n_j) \geq \sum_{j=1}^{k} (nn'_j) f(r'_j/n'_j) \tag{4.14}$$

Dividing (4.14) by nn', we obtain:

$$\sum_{j=1}^{m} (n_j/n) f(r_j/n_j) \geq \sum_{j=1}^{k} (n'_j/n') f(r'_j/n'_j) \tag{4.15}$$

Given the literacy index, $L : \mathbb{U} \to \mathbb{R}$, we know that there is $f : \mathbb{Y} \to \mathbb{R}$ satisfying conditions **C, M** and **E**, such that (4.1) holds. Thus, using (4.1) and (4.15), we have:

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) \geq L(\{(r'_1, n'_1), ..., (r'_k, n'_k)\}) \tag{4.16}$$

This means, of course, that society A is at least as literate as society B, according to the literacy index, L. ■

Remark:

In the example of Section 1, the vectors x and y , associated with societies A and B respectively, are:

$$\hat{x} = x = (\underbrace{(0.3, \ldots, 0.3)}_{300\ times}, \underbrace{(0.4, \ldots, 0.4)}_{300\ times}, \underbrace{(0.6, \ldots, 0.6)}_{300\ times})$$

$$\hat{y} = y = (\underbrace{(0.1, \ldots, 0.1)}_{300\ times}, \underbrace{(0.5, \ldots, 0.5)}_{300\ times}, \underbrace{(0.7, \ldots, 0.7)}_{300\ times})$$

Then, $\hat{x}$ super-majorizes $\hat{y}$ if and only if the vector $(0.3, 0.4, 0.6)$ super-majorizes the vector $(0.1, 0.5, 0.7)$, which it does, as checked in Section 1. Thus, $L(A) \geq L(B)$, as claimed in Section 1.

5 Externality Functions

5.1 A Characterization

Given a decomposable literacy index, $L : \mathbb{U} \to \mathbb{R}$, satisfying the axioms **N,RM,PE** and **SI**, we can associate an *externality function* with it in the following way. Note that the literacy measure, $M : \mathbb{U} \to \mathbb{R}$, associated with L, satisfies (4.4)-(4.7). Consider any $(r, n) \in \mathbb{X}$. If $r = n$, then clearly $L(\{(r, n)\}) = 1 = (r/n)$. And, if $r = 0$, then $L(\{(r, n)\}) = 0 = (r/n)$. Finally, if $0 < r < n$, then $(n - r) \in \mathbb{N}$, and $r \in \mathbb{N}$, so we have $M(\{(r, n)\}) \geq M(\{(r, r)\}) + M(\{(0, n - r)\})$ [by (4.6)] $= rM(\{(1, 1)\}) + (n - r)M(\{(0, 1)\})$ [by (4.7)] $= r$ [by (4.4)], and $L(\{(r, n)\}) \geq (r/n)$. Thus, we have:

$$L(\{(r, n)\}) \geq (r/n) \ for\ all\ (r, n) \in \mathbb{X} \tag{5.1}$$

As noted in Section 4 above, associated with L is a function, $f : \mathbb{Y} \to \mathbb{R}$, such that $L(\{(r, n)\}) = f(r/n)$ for all $(r, n) \in \mathbb{X}$. We define the *externality function* associated with L as:

$$e(r/n) = f(r/n) - (r/n) \ for\ all\ (r, n) \in \mathbb{X} \tag{5.2}$$

Clearly, using (5.1), e is a function from $\mathbb{Y}$ to $\mathbb{R}_+$. Since f satisfies property **C**, so does e. And, since $f(0) = 0$, while $f(1) = 1$, we have $e(0) = e(1) = 0$.

Note that e is *not* a monotone increasing function on $\mathbb{Y}$, but since f is a monotone increasing function on $\mathbb{Y}$, so is the function $e(z) + z$.

These three properties in fact characterize externality functions consistent with decomposable literacy indices, satisfying axioms **N,RM,PE** and **SI**. That is, if e is a function from $\mathbb{Y}$ to $\mathbb{R}_+$, which satisfies:

(i) $e(0) = e(1) = 0$,
(ii) $e(z) + z$ satisfies property **M** on $\mathbb{Y}$,
(iii) e satisfies property **C** on $\mathbb{Y}$,

then by defining:

$$L(\{(r, n)\}) = e(r/n) + (r/n) \ for\ all\ (r, n) \in \mathbb{X} \tag{5.3}$$

and :

$$L(\{(r_1, n_1), \ldots, (r_m, n_m)\}) = \sum_{i=1}^{m} (n_i/n)L(\{(r_i, n_i)\}) \tag{5.4}$$

for all $\{(r_1, n_1), \ldots, (r_m, n_m)\} \in \mathbb{U}$, with $n = (n_1 + \cdots + n_m)$, one can check that L is a decomposable literacy index, which satisfies axioms **N,RM,PE** and **SI**. Of these, axioms **N,RM** and **SI** are easy to check. To verify axiom **PE**, let (r_1, n_1) and (r_2, n_2) belong to $\mathbb{X}$, and denote $(n_1 + n_2)$ by n. Then,

$$\begin{aligned}
L(\{(r_1 + r_2), (n_1 + n_2)\}) &= e((r_1 + r_2)/(n_1 + n_2)) + ((r_1 + r_2)/(n_1 + n_2)) \\
&= e((n_1/n)(r_1/n_1) + (n_2/n)(r_2/n_2)) \\
&+ ((n_1/n)(r_1/n_1) + (n_2/n)(r_2/n_2)) \\
&\geq (n_1/n)e(r_1/n_1) + (n_2/n)e(r_2/n_2) \\
&+ ((n_1/n)(r_1/n_1) + (n_2/n)(r_2/n_2)) \\
&= (n_1/n)L(\{(r_1, n_1)\}) + (n_2/n)L(\{(r_2, n_2)\}) \\
&= L(\{(r_1, n_1), (r_2, n_2)\})
\end{aligned}$$

where the single inequality follows from property **C** of e, and the last two equalities use (5.3) and (5.4).

Remarks:

(i) The traditional index of literacy, the literacy rate, arises by defining the externality function to be: $e(z) = 0$ for all $z \in \mathbb{Y}$. Clearly, this e satifies properties (i),(ii) and (iii) described above.

(ii) The literacy index used by Basu and Foster (1998) arises by defining the externality function to be: $e(z) = \alpha(1-z)$ for $z \neq 0$, and $e(0) = 0$, where α is a number in $(0, 1)$. It is easy to verify that this e satisfies properties (i),(ii) and (iii) described above.

(iii) The literacy index proposed by Subramanian(2001) arises by defining the externality function to be: $e(z) = z(1-z)$. It can be checked that this e satifies properties (i),(ii) and (iii) described above.

5.2 An Alternative Characterization

In this subsection, we compare our characterization result with that proposed by Valenti (2002). This alternative characterization of the class of externality functions consistent with a decomposable literacy index, satisfying axioms **N,RM,PE** and **SI** can be developed as follows.

Lemma 4 *Suppose h is a function from $\mathbb{Q}_+$ to $[0, 1)$, such that: (a) $h(0) = 0$, (b) h satisfies property* $\boldsymbol{M}$ *on $\mathbb{Q}_+$, (c) h satisfies property* $\boldsymbol{C}$ *on $\mathbb{Q}_+$. If $e : \mathbb{Y} \to \mathbb{R}_+$ is a function defined as follows:*

$$e(z) = \begin{cases} h(z/(1-z))(1-z) & for\ z \neq 1 \\ 0 & for\ z = 1 \end{cases} \tag{5.5}$$

then: (i) $e(0) = e(1) = 0$, (ii) $e(z) + z$ satisfies property $\boldsymbol{M}$ *on $\mathbb{Y}$, (iii) e satisfies property* $\boldsymbol{C}$ *on $\mathbb{Y}$.*

Given Lemma 4, and using the characterization of the previous subsection, the following Proposition is immediate.

Proposition 3 *Any decomposable literacy index, L, defined by:*

$$L(\{(r,n)\}) = \begin{cases} (r/n) + h(r/(n-r))[(n-r)/n] & for\ r \neq n \\ (r/n) & for\ r = n \end{cases} \tag{5.6}$$

where $h : \mathbb{Q}_+ \to [0, 1)$ satisfies (a),(b) and (c) of Lemma 4 above, must satisfy axioms ***N,RM,PE*** *and* ***SI****.*

The interpretation of the functions e and h are as follows. If (r, n) is a household, then the total externality generated (by the literates on the illiterates) is measured by $e(r/n)$. If there are some illiterates in the household, so that $s \equiv n - r \neq 0$, then $h(r/s) \equiv h(r/(n-r))$ measures the externality generated as a proportion of the fraction who are illiterate in the household[5]. Thus, if the household is $(2,3)$, then the fraction of the illiterates in the household is $(1/3)$, and $h(r/s) = h(2)$ measures $[e(2/3)/(1/3)]$. In the terminology of Valenti (2002), the function h itself would be called the externality function, rather than the function, e.

Valenti (2002) is concerned with the class of externality functions (in her terminology), h, which satisfy properties (a) and (b) of Lemma 4, and a property stronger than property (c), which is analogous to "strict concavity":

$$(c+)\ If\ z, z' \in \mathbb{Q}_+,\ with\ z \neq z',\ and\ t \in Y,\ then\ h(tz+(1-t)z') > th(z)+(1-t)h(z')$$

By Proposition 3, that class of functions will generate literacy indices, defined by (5.6), which also satisfy axioms **N,RM,PE** and **SI**.

The converse of Proposition 3 is not true. As mentioned in the previous subsection, the traditional index of literacy, the literacy rate, arises by defining the externality function to be: $e(z) = 0$ for all $z \in \mathbb{Y}$. Since this e function satisfies properties (i), (ii) and (iii) of the previous subsection, the corresponding literacy index defined by (5.3) and (5.4) will satisfy axioms **N,RM,PE** and **SI**. However, the corresponding h function, defined implicitly by (5.6), must clearly satisfy $h(z) = 0$ for all $z \in \mathbb{Q}_+$. This h does not satisfy property (b), and it also does not satisfy property (c+).

Similarly, the literacy index used by Basu and Foster (1998) arises by defining the externality function to be: $e(z) = \alpha(1-z)$ for $z \neq 0$, and $e(0) = 0$, where α is a number in $(0,1)$.Since this e function satisfies properties (i), (ii) and (iii) of the previous subsection, the corresponding literacy index defined by (5.3) and (5.4) will satisfy axioms **N,RM,PE** and **SI**. The corresponding h function, defined implicitly by (5.6), must clearly satisfy $h(z) = \alpha$ for all $z \in \mathbb{Q}_+$. Thus, this h violates property (b) and property (c+).

Suppose L is a decomposable literacy index satisfying axioms **N,RM,PE** and **SI**, and h is the function associated with it, defined implicitly by (5.6). What properties must h satisfy ? Our two examples above show that h need not satisfy property (b). What *can* be shown is that h must satisfy properties

[5] Note that $h(r/s)$ *does not* measure the externality per illiterate. For the household $(2,3)$, the externality per illiterate is $e(2/3)$ itself, since there is only one illiterate, while $h(r/s) = h(2) = e(2/3)/(1/3)$.

(a), (c), and a weaker form of (b):

$$(b-)\ If\ z, z' \in \mathbb{Q}_+, and\ z' > z, then\ h(z') \geq h(z)$$

We can develop this result by first establishing the following lemma.

Lemma 5 *Suppose $e : \mathbb{Y} \to \mathbb{R}_+$ is a function such that (i) $e(0) = e(1) = 0$, (ii) $e(z) + z$ satisfies property **M** on $\mathbb{Y}$, (iii) e satisfies property **C** on $\mathbb{Y}$. Then, $h : \mathbb{Q}_+ \to \mathbb{R}_+$, defined by:*

$$h(w) = (1 + w)e(w/(1 + w))$$

*satisfies: (a) $h(0) = 0$, (b-) If $z, z' \in \mathbb{Q}_+$,and $z' > z$, then $h(z') \geq h(z)$,(c) h satisfies property **C** on $\mathbb{Q}_+$. Further, h maps from $\mathbb{Q}_+$ to $[0, 1)$.*

Given a literacy index L satisfying axioms **N,RM,PE** and **SI**, by the characterization of the previous subsection, there is a function e from $\mathbb{Y}$ to $\mathbb{R}_+$, satisfying properties (i),(ii) and (iii) stated there, and:

$$L(\{(r, n)\}) = (r/n) + e(r/n) \tag{5.7}$$

The corresponding function, h, from $\mathbb{Q}_+$ to $\mathbb{R}$, implicitly defined by (5.6), must then satisfy:

$$h(w) = (1 + w)e(w/(1 + w)) \tag{5.8}$$

Lemma 5 then immediately leads to the following result.

Proposition 4 *Let L be a literacy index satisfying axioms **N,RM,PE** and **SI**. Then a function, h, from $\mathbb{Q}_+$ to $\mathbb{R}$, implicitly defined by (5.6), must be a map from $\mathbb{Q}_+$ to $[0, 1)$, with properties (a), (b-) and (c) of Lemma 5.*

Remarks:

(i) Note that the functions, h, associated with the literacy rate and the Basu-Foster literacy index (through (5.6)), discussed above, satisfy properties (a), (b-) and (c), as they clearly should, since they satisfy the axioms **N,RM,PE** and **SI**.

(ii) Valenti (2002) provides a complete characterization of externality functions (in her terminology), h, consistent with her set of axioms on literacy indices, in terms of properties (a), (b) and (c+). As is to be expected, given the observations already made in this sub-section above, her set of axioms is stronger than mine, leading to a more restricted class of literacy indices. In particular, the literacy rate and the Basu-Foster literacy index cannot be accommodated in her class of literacy indices as special cases. Specifically, her axiom on equality is stronger than axioms **PE** and **SI** combined (as I have already noted in Section 3).

(iii) Valenti's characterization is in terms of functions (which I have denoted by h) defined on the ratio (r/s),the ratio of literates to *illiterates* in a household, rather than on (r/n), the literacy rate of the household. This makes it difficult to see the literacy ranking result of Section 4 directly from her characterization.

6 Proofs

In this section, the proofs of the mathematical results of Section 2 are presented.

Proof. (of **Proposition 1**): (i) Suppose $F : \mathbb{X} \to \mathbb{R}$ satisfies properties **H, SA, XM** and **OS**. Define the function, $f : \mathbb{Y} \to \mathbb{R}$ by (2.1). Then, we have $f(0) = F(0,1) = 0$ and $f(1) = F(1,1) = 1$ by **OS**, so property **E** is verified.

To verify property **M**, let $z, z' \in \mathbb{Y}$, with $z' > z$. Then there exist $(p, q) \in \mathbb{X}$, and $(p', q') \in \mathbb{X}$, with $z' = (p'/q')$, and $z = (p/q)$. Then, we have $qq' \in \mathbb{N}$, $pq' \in \mathbb{M}$, $p'q \in \mathbb{M}$ and $pq' < p'q \leq q'q$. Thus, we can write $f(z') = f(p'/q') = f(p'q/q'q) = F(p'q, q'q)/q'q > F(pq', qq')/qq' = f(pq'/qq') = f(p/q) = f(z)$, the inequality following from property **XM** of F.

To verify property **C,** let $z, z' \in \mathbb{Y}, \textit{and } t \in Y$. Then there exist $(p, q) \in \mathbb{X}$, $(p', q') \in \mathbb{X}$, and $(s, r) \in \mathbb{X}$, with $z = (p/q)$, $z' = (p'/q')$, and $t = (s/r)$. Then, we have:

$$
\begin{aligned}
f(tz+(1-t)z') &= f(\frac{spq'}{rqq'}+\frac{(r-s)p'q}{rq'q}) \\
&= f(\frac{spq'+(r-s)p'q}{rqq'}) \\
&= F(spq'+(r-s)p'q, rqq')/rqq' \\
&= F(spq'+(r-s)p'q, sqq'+(r-s)qq')/rqq' \\
&\geq \frac{F(spq', sqq')+F((r-s)p'q,(r-s)qq')}{rqq'} \\
&= \frac{sq'F(p,q)}{rqq'}+\frac{(r-s)qF(p',q')}{rqq'} \\
&= (s/r)\frac{F(p,q)}{q}+(1-(s/r))\frac{F(p',q')}{q'} \\
&= (s/r)f(p/q)+(1-(s/r))f(p'/q') \\
&= tf(z)+(1-t)f(z')
\end{aligned}
$$

where property **SA** of F is used to obtain the inequality, and property **H** of F is used in the very next line in the above computations.

(ii) To establish the converse, given any $f:\mathbb{Y}\to\mathbb{R}$, satisfying properties **C, M,** and **E**, we associate with it a function, $F:\mathbb{X}\to\mathbb{R}$, defined by (2.2). Then, $F(1,1)=f(1)=1$, and $F(0,1)=f(0)=0$, verifying property **OS** of F.

To verify property **H** of F, let $(x,y)\in\mathbb{X}$, and $t\in\mathbb{N}$. Then, $(tx,ty)\in\mathbb{X}$, and $F(tx,ty)=tyf(tx/ty)=tyf(x/y)=tF(x,y)$.

To verify property **XM** of F, let $(x,y)\in\mathbb{X}$, and $(x',y)\in\mathbb{X}$, with $x'>x$. Then, by property **M** of f, we have $F(x',y)=yf(x'/y)>yf(x/y)=F(x,y)$.

Finally, to verify property **SA** of F, let $(x,y)\in\mathbb{X}$, and $(x',y')\in\mathbb{X}$. Then,

$$
\begin{aligned}
F(x+x',y+y') &= (y+y')f(\frac{x+x'}{y+y'}) \\
&= (y+y')f((x/y)\frac{y}{y+y'}+(x'/y')\frac{y'}{y+y'}) \\
&\geq yf(x/y)+y'f(x'/y') \\
&= F(x,y)+F(x',y')
\end{aligned}
$$

where property **C** of f was used to obtain the inequality in the above computations. ■

Proof. (of **Lemma 1**): We will only prove (i). The proof of (ii) is similar; (iii) follows directly from (i) and (ii). We are given that $a, b, c \in \mathbb{Y}$ and $a < b \leq c$. Then, there is a real number $\lambda \in (0, 1]$, such that $b = (1-\lambda)a+\lambda c$. Then, $\lambda(c - a) = (b - a)$, so that:

$$\lambda = \frac{b-a}{c-a} \tag{6.1}$$

Thus, λ must be a rational number in $(0, 1]$. Using property **C** of the function f, we then have: $f(b) \geq (1 - \lambda)f(a) + \lambda f(c)$, which can be re-written as:

$$\lambda[f(c) - f(a)] \leq [f(b) - f(a)] \tag{6.2}$$

Using (6.1) in (6.2), and dividing through by $(b - a) > 0$, we get:

$$\frac{[f(c) - f(a)]}{(c-a)} \leq \frac{[f(b) - f(a)]}{(b-a)} \tag{6.3}$$

This establishes (i). ■

Proof. (of **Lemma 2**): For $y \in Y$, we can pick $\delta \in Y$, such that $y - \delta > 0$, and $y + \delta < 1$. For all $\varepsilon \in Y$, with $\varepsilon < \delta$, we have by Lemma 1,

$$\frac{f(y) - f(y-\delta)}{\delta} \geq \frac{f(y+\varepsilon) - f(y)}{\varepsilon} \tag{6.4}$$

Also, by Lemma 1, the right-hand side expression in (6.4) is non-decreasing as ε decreases to zero. Since it is bounded above by the left-hand side expression in (6.4), it must converge to a limit. Thus, $g : Y \to \mathbb{R}$, is well-defined by:

$$g(y) \equiv \lim_{\substack{\varepsilon \downarrow 0 \\ y+\varepsilon \in Y}} \{[f(y+\varepsilon) - f(y)]/\varepsilon\} \tag{6.5}$$

Let $y, z \in Y$ with $y < z$. We can pick $\delta \in Y$, such that $0 < y + \delta < z < z + \delta < 1$. Then, for all $n \in \mathbb{N}$, we have $0 < y + (\delta/n) < z < z + (\delta/n) < 1$, and so by Lemma 1,

$$\frac{f(y+(\delta/n)) - f(y)}{(\delta/n)} \geq \frac{f(z+(\delta/n)) - f(z)}{(\delta/n)} \tag{6.6}$$

Letting $n \to \infty$ in (6.6), and using (6.5), we have $g(y) \geq g(z)$. ■

Proof. (of **Lemma 3**): The result is trivial for $x = y$. So, we consider two cases (i) $x > y$; (ii) $x < y$. In case (i), we can choose $\delta \in Y$, such that $x > y + \delta$. Then, for all $n \in \mathbb{N}$, $x > y + (\delta/n)$. Consequently, for all $n \in \mathbb{N}$, we have:

$$\frac{f(x) - f(y)}{(x - y)} \leq \frac{f(y + (\delta/n)) - f(y)}{(\delta/n)} \tag{6.7}$$

by Lemma 1. Letting $n \to \infty$ in (6.7), and using Lemma 2, we have:

$$\frac{f(x) - f(y)}{(x - y)} \leq g(y) \tag{6.8}$$

Multiplying through in (6.8) by $(x - y) > 0$, we get the desired result.

In case (ii), we can choose $\delta \in Y$, such that $x < y - \delta$. Then, for all $n \in \mathbb{N}$, $x < y - (\delta/n)$. Therefore, for all $n \in \mathbb{N}$, we obtain:

$$\frac{f(y) - f(x)}{(y - x)} \geq \frac{f(y) - f(y - (\delta/n))}{(\delta/n)} \geq \frac{f(y + (\delta/n)) - f(y)}{(\delta/n)} \tag{6.9}$$

by Lemma 1. Letting $n \to \infty$ in (6.9), and using Lemma 2, we get:

$$\frac{f(y) - f(x)}{(y - x)} \geq g(y) \tag{6.10}$$

Multiplying through in (6.10) by $(y - x) > 0$, we obtain:

$$f(y) - f(x) \geq g(y)(y - x) \tag{6.11}$$

Transposing terms in (6.11) yields the desired result. ■

Proof. (of **Proposition 2**): We first show the result under the assumption that $z_i \in Y$ for all $i \in \{1, \ldots, n\}$. Then, we show that the general case (without this assumption) follows easily.

Since $z_i \in Y$ and $z_i' \in \mathbb{Y}$, we can use Lemma 3 to write, for each $i \in \{1, \ldots, n\}$,

$$f(z_i') - f(z_i) \leq g(z_i)(z_i' - z_i) \tag{6.12}$$

Summing over $i \in \{1, \ldots, n\}$, we get:

$$\sum_{i=1}^{n} [f(z_i') - f(z_i)] \leq \sum_{i=1}^{n} g(z_i)(z_i' - z_i) \tag{6.13}$$

Using Lemma 2, we have:

$$g(z_1) \geq g(z_2) \geq \cdots \geq g(z_n) \tag{6.14}$$

Using the fact that f satisfies property **M** in Lemma 2, we also have:

$$g(z_i) \geq 0 \; for \; i \in \{1, \ldots, n\} \tag{6.15}$$

Thus, we can use Abel's inequality (Mitrinovic-Vasic (1970, p.32)) to write:

$$\sum_{i=1}^{n} g(z_i)(z_i' - z_i) \leq g(z_1)[\max_{1 \leq k \leq n} \sum_{i=1}^{k}(z_i' - z_i)] \tag{6.16}$$

By (2.5), we have:

$$[\max_{1 \leq k \leq n} \sum_{i=1}^{k}(z_i' - z_i)] \leq 0 \tag{6.17}$$

Thus, using (6.15) and (6.17) in (6.16), we have:

$$\sum_{i=1}^{n} g(z_i)(z_i' - z_i) \leq 0 \tag{6.18}$$

Using (6.18) in (6.13), we obtain (2.6), the desired result.

Turning now to the more general situation, we see that the following cases can arise: (i) $z_i \in Y$ for $i \in \{1, \ldots, n\}$; (ii) $z_i = 0$ for $i \in \{1, \ldots, n\}$; (iii) $z_i = 1$ for $i \in \{1, \ldots, n\}$; (iv) there is $1 \leq p < n$ such that $z_i = 0$ for $i \in \{1, \ldots, p\}$ and $z_i = 1$ for $i \in \{p + 1, \ldots, n\}$; (v) there is $1 \leq p < n$ such that $z_i = 0$ for $i \in \{1, \ldots, p\}$ and $z_i \in Y$ for $i \in \{p + 1, \ldots, n\}; (vi)$ there is $1 \leq q < n$ such that $z_i \in Y$ for $i \in \{1, \ldots, q\}$ and $z_i = 1$ for $i \in \{p + 1, \ldots, n\}; (vii)$ there exist $1 \leq p < q < n$ such that $z_i = 0$ for $i \in \{1, \ldots, p\}$, $z_i \in Y$ for $i \in \{p + 1, \ldots, q\}$and $z_i = 1$ for $i \in \{q + 1, \ldots, n\}$.

We have already established the result in case (i). In case (ii), (2.5) implies that $z_i' = 0$ for $i \in \{1, \ldots, n\}$, so (2.6) follows trivially. In case (iii), $z_i' \leq z_i$ for $i \in \{1, \ldots, n\}$, so (2.6) follows from property **M** of f. In case (iv), we have $z_i' = 0$ for $i \in \{1, \ldots, p\}$ by (2.5), while $z_i' \leq z_i$ for $i \in \{p+1, \ldots, n\}$, so (2.6) follows again from property **M** of f.

In case (v), we have $z_i' = 0$ for $i \in \{1, \ldots, p\}$ by (2.5), so we can define $j = i - p$ for $i = p + 1, \ldots, n$, and $m = n - p$. Then, by (2.5), we have for $k = 1, \ldots, m$,

$$\sum_{j=1}^{k} z_j' \leq \sum_{j=1}^{k} z_j$$

so that by the analysis of case (i), we obtain:

$$\sum_{j=1}^{m} f(z'_j) \leq \sum_{j=1}^{m} f(z_j)$$

and this yields (2.6), since $z'_i = z_i = 0$ for $i = 1, \ldots, p$.

In case (vi), using the analysis of case (i), we have:

$$\sum_{i=1}^{q} f(z'_i) \leq \sum_{i=1}^{q} f(z_i)$$

and this yields (2.6), by using property **M** of f, since $z'_i \leq z_i = 1$ for $i = q+1, \ldots, n$.

In case (vii), we have $z'_i = 0$ for $i \in \{1, \ldots, p\}$ by (2.5), so we can define $j = i - p$ for $i = p+1, \ldots, q$, and $m = q - p$. Then, by (2.5), we have for $k = 1, \ldots, m$,

$$\sum_{j=1}^{k} z'_j \leq \sum_{j=1}^{k} z_j$$

so that by the analysis of case (i), we have:

$$\sum_{j=1}^{m} f(z'_j) \leq \sum_{j=1}^{m} f(z_j)$$

This yields (2.6), by (a) using property **M** of f, and $z'_i \leq z_i = 1$ for $i = q+1, \ldots, n$, while noting that (b) $z'_i = z_i = 0$ for $i = 1, \ldots, p$. ∎

Proof. (of **Lemma 4**): Property (i) of e being clear from (5.5), we proceed to prove property (ii). Let $z, z' \in \mathbb{Y}$, with $z' > z$. There are two cases to consider: (I) $z' = 1$, (II) $z' < 1$. In case (I), $e(z') + z' = 1$, from (5.5), while $e(z) + z = h(z/(1-z))(1-z) + z < (1-z) + z$ (since $z < 1$, and $h(z/(1-z)) \in [0, 1)$)$= 1$.

In case (II), we have:

$$\begin{aligned} e(z') + z' &= h(z'/(1-z'))(1-z') + z' \\ &> h(z/(1-z))(1-z') + z' \\ &= h(z/(1-z))(1-z) + z + h(z/(1-z))(z - z') + (z' - z) \\ &= e(z) + z + (z' - z)[1 - h(z/(1-z))] \\ &> e(z) + z \end{aligned}$$

the first inequality following from property **M** of h and the fact that $z'/(1-z') > z/(1-z)$ [using $1 > z' > z$], and the second inequality following from the fact that $h(z/(1-z)) \in [0,1)$ and $z' > z$.

We turn now to property (iii) of the function, e. Let $z, z' \in \mathbb{Y}$, and let $t \in Y$. We will first prove the property, assuming that z and z' are both less than 1. Then, we will show that the property is also valid without this assumption. When $z, z' < 1$, let us denote $[1-(tz+(1-t)z')]$ by Δ. Then, we have:

$$\Delta = t(1-z) + (1-t)(1-z') \tag{6.19}$$

Then, we can write:

$$\begin{aligned}
e(tz+(1-t)z') &= h(\frac{(tz+(1-t)z')}{\Delta})\Delta \\
&= h(\frac{tz(1-z)}{\Delta(1-z)} + \frac{(1-t)z'(1-z')}{\Delta(1-z')})\Delta \\
&= h(\frac{t(1-z)}{\Delta}\frac{z}{1-z} + \frac{(1-t)(1-z')}{\Delta}\frac{z'}{1-z'})\Delta \\
&\geq [\frac{t(1-z)}{\Delta}h(\frac{z}{1-z}) + \frac{(1-t)(1-z')}{\Delta}h(\frac{z'}{1-z'})]\Delta \\
&= te(z) + (1-t)e(z')
\end{aligned}$$

the inequality following from property **C** of the function, h, after using (6.19).

In the general case, where z, z' are not necessarily less than 1, we proceed as follows. Define, for all $n \in \mathbb{N}$, $z(n) = z$ if $z < 1$, and $z(n) = [n/(1+n)]z$, when $z = 1$; similarly, define for all $n \in \mathbb{N}$, $z'(n) = z'$ if $z' < 1$, and $z'(n) = [n/(1+n)]z'$, when $z' = 1$. Then, $z(n), z'(n)$ are both less than 1 for all $n \in \mathbb{N}$, with $z(n) \leq z$ and $z'(n) \leq z'$ for all $n \in \mathbb{N}$. Thus, we can write:

$$\begin{aligned}
e(tz+(1-t)z') + [tz+(1-t)z'] &\geq e(tz(n)+(1-t)z'(n)) \\
&\quad +[tz(n)+(1-t)z'(n)] \\
&\geq te(z(n)) + (1-t)e(z'(n)) \\
&\quad +[tz(n)+(1-t)z'(n)] \qquad (6.20)
\end{aligned}$$

the first inequality following from property (ii) of e (already established above), and the second inequality from the above analysis of property (iii) of e, since $z(n), z'(n)$ are both less than 1. The inequality (6.20) yields:

$$\begin{aligned} e(tz+(1-t)z') &\geq te(z(n))+(1-t)e(z'(n)) \\ &\quad +t(z(n)-z)+(1-t)(z'(n)-z') \\ &\geq te(z)+(1-t)e(z') \\ &\quad +t(z(n)-z)+(1-t)(z'(n)-z') \qquad (6.21) \end{aligned}$$

the second inequality following from the fact that $e(1)=0$, while e maps to $\mathbb{R}_+$. Now, letting $n \to \infty$ in (6.21), we obtain the desired result. ■

Proof. (of **Lemma 5**): Property (a) of the function h is clear from (5.8) and property (i) of the function, e. We proceed to verify property (c) of h. Let $w, w' \in \mathbb{Q}_+$, and let $t \in Y$. Define $w'' = tw + (1-t)w'$, and note that $w'' \in \mathbb{Q}_+$, and:

$$(1+w'') = t(1+w) + (1-t)(1+w') \qquad (6.22)$$

Clearly, we have $[w/(1+w)]$, $[w'/(1+w')]$, $[w''/(1+w'')]$ in $\mathbb{Y}$. Using (5.8), we can write:

$$\begin{aligned} h(w'') &= e(w''/(1+w''))(1+w'') \\ &= e\Big(\frac{tw(1+w)}{(1+w'')(1+w)} + \frac{(1-t)w'(1+w')}{(1+w'')(1+w')}\Big)(1+w'') \\ &= e\Big(\frac{t(1+w)}{(1+w'')}\frac{w}{(1+w)} + \frac{(1-t)(1+w')}{(1+w'')}\frac{w'}{(1+w')}\Big)(1+w'') \\ &\geq \Big[\frac{t(1+w)}{(1+w'')}e\Big(\frac{w}{(1+w)}\Big) + \frac{(1-t)(1+w')}{(1+w'')}e\Big(\frac{w'}{(1+w')}\Big)\Big](1+w'') \\ &= th(w) + (1-t)h(w') \end{aligned}$$

the inequality following from property (iii) of e, after using (6.22).

We now turn to property (b-) of the function, h. Let $w, w' \in \mathbb{Q}_+$, with $w' > w$. We claim that $h(w') \geq h(w)$. Suppose, on the contrary, that $h(w') < h(w)$. Define $\delta = [h(w) - h(w')]/(w' - w)$; then $\delta > 0$. Using property (c) of h (which has already been established), and using the proof of Lemma 1, we have for all $n \in \mathbb{N}$,

$$-\delta = \frac{h(w') - h(w)}{(w' - w)} \geq \frac{h(w'+n) - h(w)}{(w'+n-w)}$$

This yields the inequality:

$$-h(w) \leq h(w' + n) - h(w) \leq (w' + n - w)(-\delta) < -n\delta \tag{6.23}$$

using the facts that $h(w'+n) \geq 0$ and $(w'-w) > 0$. But, (6.23) clearly leads to a contradiction for large n. This establishes our claim, and hence property (b-) of h.

To verify that h maps from $\mathbb{Q}_+$ to $[0,1)$, suppose on the contrary there is some $w \in \mathbb{Q}_+$, with $h(w) \geq 1$. Define $z = w/(1+w)$; then, we have $z \in [0,1)$, and $(1-z) = 1/(1+w)$. Using (5.8), we then have $e(z) = (1-z)h(w) \geq (1-z)$, so that $e(z) + z \geq 1$. But, by property **M** of e, we must have $e(z) + z < e(1)+1 = 1$, since $e(1) = 0$ by property (i) of e. This contradiction establishes the result. ■

References

[1] Basu, K. and J.E. Foster (1998), "On Measuring Literacy", *Economic Journal*, 108, 1733-1749.

[2] Green, J.W. (1954), "Recent Applications of Convex Functions", *American Mathematical Monthly*, 61(7), 449-454.

[3] Mitrinovic, D.S. and P.M.Vasic (1970), *Analytic Inequalities*, Springer-Verlag, Heidelberg.

[4] Nikaido, H. (1968), *Convex Structures and Economic Theory*, Academic Press, New York.

[5] Roberts, A.W. and D.E. Varberg (1973), *Convex Functions*, Academic Press, New York.

[6] Rosenbaum, R.A. (1950), "Sub-Additive Functions", *Duke Mathematical Journal*, 17, 227-247.

[7] Subramanian, S. (2001), "Equity, Efficiency, and the Measurement of Literacy: Some Extensions of the Basu-Foster Framework", mimeo, Madras Institute of Development Studies.

[8] Valenti, P.M. (2002), "Should We Be Concerned About the Distribution of Literacy Across Households? An Axiomatic Characterization", CAE Working Paper 02-15, Cornell University.