17
On Literacy Rankings*

Tapan Mitra

The importance of literacy in the process of economic development of a society is well recognised. 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 *effectively more literate* than an illiterate person in a household which has no literate person. Consequently, some account needs to be taken of this externality in capturing the *effective* literacy of the household.

Empirical work shows that such an intra-household externality of literacy is present, and can be quite large. In an early essay in this area, Green et al. (1985) studied the role of 'shared literacy' in the adoption of modern farm practices in Guatemala. More recent work include Gibson (2001), who studied the effects of adult literacy on children's nutritional status in Papua New Guinea, and Basu et al. (2002), who studied the effects of education on individual earnings in the non-farm sector in Bangladesh.

This essay is primarily concerned with the observation of Basu and Foster (1998) that, taking into account the fact that literate household members generate a positive externality for illiterate members, 'a more even distribution of literacy across households leads to greater effective literacy'. The observation would be of greater interest if its validity (*a*) does not depend on the choice of a specific externality function, but applies to an entire class of externality functions, consistent with literacy indices satisfying a set of reasonable axioms, and (*b*) the distributions

* This essay is dedicated, with affection and admiration, to Satish Jain. For comments on an earlier version of this essay, I am grateful to Paola Valenti. For encouragement in preparing the current revised version, I am greatly indebted to S. Subramanian.

of literacy that are being compared encompass a broad class and are not confined merely to extreme polar cases.

With this objective in mind, we present an axiomatic study of the class of decomposable literacy indices in the second section of the essay. Four axioms are imposed on the literacy indices, reflecting monotonicity with respect to literates, scale invariance (to capture the idea that the index is a relative measure, not an absolute one), a standard normalisation rule and, crucially, a positive externality axiom (reflecting the fact that literates generate an externality within a household for its illiterate members). Theorem 1 in the second section presents our main representation result for decomposable literacy indices. It shows that any decomposable literacy index of a society (satisfying the set of four axioms) can be written as the weighted sum of a *concave* and *increasing* function of the *literacy rates* of the individual households of that society, and that this form of representation in fact characterises decomposable literacy indices, satisfying the four axioms. 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 majorisation 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 essay, is presented in the Appendix. It leads to Theorem 2 in the second section, which can be considered to be the principal theoretical application of Theorem 1. It asserts that, given two societies A and B, if the distribution of the literacy rates of the households in society A is more equitable than the distribution of the literacy rates of the households in society B, in the sense of the standard Lorenz quasi-order, then society A has a higher effective literacy index compared to society B.[1]

Theorem 2 meets the primary objective of the essay, described earlier. The distributions of literacy that can be compared are as numerous

[1] Expressions like 'higher' and 'more equitable' are used in the weak sense. So, a more precise formulation of the result would be that if society B has a strictly less equitable distribution of literacy rates than society A in the sense of the standard Lorenz quasi-order, then society B cannot have a strictly higher literacy index compared to society A.

as those that are Lorenz comparable. And, the result does not depend on a choice of a specific externality function, but on the positive externality axiom which asserts merely the presence of an intra-household externality of literacy. However, Theorems 1 and 2 naturally lead to the explicit characterisation of externality functions implicit in our set of four axioms on decomposable literacy indices; this is presented in Theorem 3 in the third section 3. It would appear from our results in the second and third sections 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.

In the fourth section, we compare our characterisation result in Theorem 3, with those proposed by Dutta (2004) and Valenti (2002).[2] The framework used in these essays is comparable to but somewhat different from ours, and we describe it explicitly in the section 'The Alternative Framework'. The axiom system used by Valenti (2002) implies that all four of our axioms hold, and is slightly stronger. Thus, our main conclusion with respect to the distribution of literacy (Theorem 2) holds under her axiom system. Further, using our characterisation of externality functions (Theorem 3), we can present (in Theorem 4) the main contribution in her essay: a characterisation of externality functions implicit in her axiom system.

The axiom system used by Dutta (2004) is weaker than the one in Valenti (2002) and, in particular, the 'weak externality' axiom used by him asserts the presence of an intra-household externality of literacy only in more restricted circumstances. As a result, the characterisation of externality functions implicit in his axiom system (presented in Theorem 5) imposes less structure on the externality functions. In particular, our principal result on the distribution of literates (Theorem 2) does not hold under his axiom system, and we demonstrate this with a simple example. In the example, a society with an unequal distribution of literates among households is seen to have a *strictly higher* effective

[2] An earlier version (Mitra 2002) of the current essay was circulated as a CAE Working paper in 2002. In the decade that has followed, there has been considerable research in this area. In terms of axiomatic analysis, the contributions by Valenti (2002) and Dutta (2004) are noteworthy. The purpose of the fourth section is to provide the reader with a better perspective of alternative axiomatic approaches that have been pursued in studying literacy indices with intra-household externality to literacy.

literacy index than a society with a completely equal distribution of literates, and this literacy index is consistent with the axiom set used by Dutta (2004).

Literacy Indices

Notation

Let $\mathbb{N}$ denote the set of natural numbers $\{1, 2, 3, ...\}$, and let $\mathbb{M}$ denote the set $\{0, 1, 2, 3, ...\}$. We define the set $\mathbb{X} = \{(x, y) : x \in \mathbb{M}, y \in \mathbb{N}, \text{ and } x \leq y\}$. Let $\mathbb{Q}$ denote the set of rational numbers, and $\mathbb{R}$ the set of real numbers. We denote by $\mathbb{Q}_+$ the set of non-negative rational numbers $\{z = p/q, \text{ where } p \in \mathbb{M}, \text{ and } q \in \mathbb{N}\}$, and by $\mathbb{Y}$ the set $\mathbb{Q} \cap [0, 1]$; the set $\mathbb{Q} \cap (0, 1)$ is denoted by Y.

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 formalise this, we define the set of all possible societies to be:

$$\mathbb{U} = \bigcup_{k=1}^{\infty} \mathbb{X}^k \tag{17.1}$$

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:

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) = \sum_{i=1}^{m} (n_i/n)L(\{(r_i, n_i)\}), \tag{17.2}$$

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.

Remark 1.

1. *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 emphasised, 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 formalised 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, therefore, 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.*
2. *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.*

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 (Normalisation 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):

$$\text{If } (r,n) \in \mathbb{X}, \text{ and } (r',n) \in \mathbb{X}, \text{ and } r' > r, \text{ then } L(\{(r',n)\}) > L(\{(r,n)\}).$$

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 formalise these ideas in the positive externality axiom.

Axiom PE (Positive Externality Axiom):

$$\text{If } (r_1, n_1) \in \mathbb{X}, \text{ and } (r_2, n_2) \in \mathbb{X}, \text{ then } L(\{(r_1 + r_2), (n_1 + n_2)\}) \geq L(\{(r_1, n_1), (r_2, n_2)\}).$$

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):

$$\text{If } (r, n) \in \mathbb{X}, \text{ and } k \in \mathbb{N}, \text{ then } L(\{(kr, kn)\}) = L(\{(r, n)\}).$$

Remark 2.

1. *The usual statement of the normalisation axiom (in Basu and Foster [1998], Dutta [2004], 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.*
2. *The monotonicity axiom (**Axiom RM**) is basically the same as the one used in Basu and Foster (1998) and Valenti (2002).*

3. *Valenti (2002) uses an 'equality axiom' instead of* ***Axiom PE*** *and* ***Axiom SI****. Her equality axiom implies* ***PE*** *and* ***SI****, but is stronger than axioms* ***PE*** *and* ***SI*** *combined. I have found it more useful to separate the two axioms. This makes* ***Axiom PE*** *acceptable, if one believes 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. A detailed description of Valenti's axiom system can be found in the fourth section.*
4. *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****.*

Representation of Decomposable Literacy Indices

In this subsection, we discuss the main representation result of the class of decomposable literacy indices, satisfying the set of four axioms introduced in the previous subsection.

To formulate the result precisely, let us define the following properties that a function $\phi : \mathbb{Q}_+ \to \mathbb{R}$ might satisfy.[3]

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

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 will establish the following representation result.

Theorem 1. *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*

[3] Here, the names 'concavity' and 'monotonicity' are deliberately used, even though the domain of ϕ is not an interval of the real line, unlike the usual setting for concave and monotone functions of a real variable.

$\{(r_1, n_1), ..., (r_m, n_m)\} \in \mathbb{U}$, *we have:*

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) = \sum_{i=1}^{m} (n_i/n) f(r_i/n_i) \qquad (17.3)$$

Conversely, any literacy index $L : \mathbb{U} \to \mathbb{R}$, *defined by (17.3), where* $f : \mathbb{Y} \to \mathbb{R}$ *satisfies properties* ***C, M*** *and* ***E,*** *must satisfy Axioms* ***N, RM, PE*** *and* ***SI.***

Proof: The second statement in Theorem 1 is, of course, trivial to verify. So, we proceed to demonstrate the first statement. 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:

$$M(\{(r_1, n_1), ..., (r_m, n_m)\}) = L(\{(r_1, n_1), ..., (r_m, n_m)\})(n_1 + \cdots + n_m). \qquad (17.4)$$

Thus, M is also a real-valued function on $\mathbb{U}$. Since L is decomposable (that is, it satisfies (17.2)), M must be *additively separable*; that is, it must satisfy:

$$M(\{(r_1, n_1), ..., (r_m, n_m)\}) = \sum_{i=1}^{m} M(\{(r_i, n_i)\}). \qquad (17.5)$$

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:

$$M(\{(1, 1)\}) = 1, \; M(\{(0, 1)\}) = 0. \qquad (17.6)$$

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

$$\text{if } (r, n) \in \mathbb{X}, \text{ and } (r', n) \in \mathbb{X}, \text{ then } M(\{(r', n)\}) > M(\{(r, n)\}). \qquad (17.7)$$

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

$$\text{if } (r_1, n_1), (r_2, n_2) \in \mathbb{X}, \text{ then}$$
$$M(\{(r_1 + r_2, n_1 + n_2)\}) \geq M(\{(r_1, n_1)\}) + M(\{(r_2, n_2)\}). \qquad (17.8)$$

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

$$\text{if } (r, n) \in \mathbb{X}, \text{ and } k \in \mathbb{N}, \text{ then } M(\{(kr, kn)\}) = kM(\{(r, n)\}). \qquad (17.9)$$

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

$$F(r, n) = M(\{(r, n)\}). \qquad (17.10)$$

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.

Given F, one can define a function $f : \mathbb{Y} \to \mathbb{R}$ by:

$$f(z) = F(p, q)/q \text{ where } z = p/q, \text{ and } (p, q) \in \mathbb{X}, \tag{17.11}$$

as formally demonstrated in the Appendix. Further, it follows from (17.6),(17.7), (17.8) and (17.9) that f satisfies properties **C, M** and **E,** by Proposition 1 in the Appendix.

Clearly, (17.10) and (17.11) 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 (17.4), $[M(\{(r, n)\})/n]$ is $L(\{(r, n)\})$, and so any decomposable literacy index, L, satisfying the axioms **N, RM, PE** and **SI** can be written as:

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

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

Distribution of Literacy

Our representation theorem shows that any decomposable literacy index of a society (satisfying a set of reasonable axioms) can be written as the weighted sum of a *concave* and *increasing* function of the *literacy rates* of the individual households of that society. This representation has the advantage that it can provide a formal demonstration of one of the principal themes of the recent literature on literacy that 'effective literacy is enhanced by a more, rather than less, equitable distribution of literates across households.' (Subramanian 2008: 839). A particularly appealing aspect of this result (stated formally in Theorem 2) is that its validity does not depend on the choice of a specific literacy index of a society, but applies to the *entire class* of literacy indices of society, satisfying the set of four axioms described in subsection on 'Axioms on Decomposable Literacy Indices'. Specifically, the result states that, given two societies A and B, if the distribution of the literacy rates of the households in society A is more equitable than the distribution of the literacy rates of the households in society B, in the sense of the standard Lorenz quasi-order, then society A has a higher literacy index compared to society B.

Theorem 2. *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:*

$$x = (\underbrace{(r_1/n_1), \ldots, (r_1/n_1)}_{n'n_1 \text{ times}}, \underbrace{(r_2/n_2), \ldots, (r_2/n_2)}_{n'n_2 \text{ times}}, \ldots, \underbrace{(r_m/n_m), \ldots, (r_m/n_m)}_{n'n_m \text{ times}})$$
$$y = (\underbrace{(r'_1/n'_1), \ldots, (r'_1/n'_1)}_{nn'_1 \text{ times}}, \underbrace{(r'_2/n'_2), \ldots, (r'_2/n'_2)}_{nn'_2 \text{ times}}, \ldots, \underbrace{(r'_k/n'_k), \ldots, (r'_k/n'_k)}_{nn'_k \text{ times}}). \quad (17.12)$$

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:

$$\sum_{i=1}^{K} \hat{x}_i \geq \sum_{i=1}^{K} \hat{y}_i \quad (17.13)$$

Then, we have:

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) \geq L(\{(r'_1, n'_1), ..., (r'_k, n'_k)\}), \quad (17.14)$$

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

Proof: Given (17.13), we can apply Proposition 2 of the Appendix to obtain:

$$\sum_{i=1}^{nn'} f(\hat{x}_i) \geq \sum_{i=1}^{nn'} f(\hat{y}_i) \quad (17.15)$$

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

$$\sum_{j=1}^{m} (n'n_j) f(r_j/n_j) \geq \sum_{j=1}^{k} (nn'_j) f(r'_j/n'_j). \quad (17.16)$$

Dividing (17.16) by nn', we obtain:

$$\sum_{j=1}^{m} (n_j/n) f(r_j/n_j) \geq \sum_{j=1}^{k} (n'_j/n') f(r'_j/n'_j). \quad (17.17)$$

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 (17.3) holds. Thus, using (17.3) and (17.17), we have:

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

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

Remark 3. *To illustrate Theorem 2, consider two societies A and B, defined as follows:*

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

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

The vectors x and y (and $\hat{x}$ and $\hat{y}$) associated with societies A and B respectively are as follows:

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

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

Then, $\hat{x}, \hat{y}$ satisfy (17.13) if and only if the vector (0.3, 0.4, 0.6) Lorenz-dominates the vector $(0.1, 0.5, 0.7)$, which it does, since $0.3 > 0.1$, $0.3 + 0.4 > 0.1 + 0.5$ and $0.3 + 0.4 + 0.6 = 0.1 + 0.5 + 0.7$. Thus, $L(A) \geq L(B)$ for *every* decomposable literacy index L satisfying **Axioms N, RM, PE** and **SI.**

Externality Functions

We have defined the class of decomposable literacy indices $L : \mathbb{U} \to \mathbb{R}$ by:

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) = \sum_{i=1}^{m} (n_i/n)L(\{(r_i, n_i)\}). \tag{17.18}$$

Note that for any household $i \in \{1, \ldots, m\}$, we can write:

$$
\begin{aligned}
(n_i/n)L(\{(r_i, n_i)\}) &= (n_i/n)[(r_i/n_i) + L(\{(r_i, n_i)\}) - (r_i/n_i)] \\
&= \frac{n_i[(r_i/n_i) + L(\{(r_i, n_i)\}) - (r_i/n_i)]}{n} \\
&= \frac{[r_i + n_i L(\{(r_i, n_i)\}) - r_i]}{n} \\
&\equiv \frac{[r_i + E(r_i, n_i)]}{n},
\end{aligned} \tag{17.19}
$$

where we have defined the function $E : \mathbb{X} \to \mathbb{R}$ by:

$$E(r, n) \equiv nL(\{(r, n)\}) - r \text{ for all } (r, n) \in \mathbb{X}. \tag{17.20}$$

Thus, any decomposable literacy index $L : \mathbb{U} \to \mathbb{R}$ can be written as:

$$L(\{(r_1, n_1), \ldots, (r_m, n_m)\}) = \frac{\sum_{i=1}^{m}[r_i + E(r_i, n_i)]}{n}, \tag{17.21}$$

where $n = (n_1 + \cdots + n_m)$, and $E : \mathbb{X} \to \mathbb{R}$ is defined by (17.20).

Conversely, if E is a map from $\mathbb{X}$ to $\mathbb{R}$, and $L : \mathbb{U} \to \mathbb{R}$ is defined by (17.21), then clearly E satisfies (17.20), and:

$$
\begin{aligned}
L(\{(r_1, n_1), \ldots, (r_m, n_m)\}) &= \frac{\sum_{i=1}^{m}[r_i + E(r_i, n_i)]}{n} \\
&= \frac{\sum_{i=1}^{m}[r_i + n_i L(\{(r_i, n_i)\}) - r_i]}{n} \\
&= \sum_{i=1}^{m}(n_i/n)L(\{(r_i, n_i)\}),
\end{aligned} \tag{17.22}
$$

so that L is a decomposable literacy index.

Note that the expression (17.21) is valid for any decomposable literacy index $L : \mathbb{U} \to \mathbb{R}$, independent of the four axioms imposed on L discussed earlier. The form (17.21) of decomposable literacy indices is useful for the discussion of externality in measuring literacy, which follows.[4]

Crucial to the representation result stated in Theorem 1, and its principal application (stated in Theorem 2), is the recognition of the fact that the literates in a household bestow on the illiterates of that household

[4] If E (defined by [17.20]) happens to be a map from $\mathbb{X}$ to $\mathbb{R}_+$ (that is, it is non-negative on its domain), then it can be interpreted as an intra-household *externality measure*.

an *externality*, so that the effective literacy rate of a household is typically different from the literacy rate measured in the standard way as the percentage of literates in a household. This key observation of Basu and Foster (1998) has been formalised in different ways in the literature on literacy. In our formulation, it appears in the form of the positive externality axiom. Because of the extremely convenient form of the representation of decomposable literacy indices (satisfying **Axioms N, RM, PE** and **SI**) obtained in Theorem 1, the externality implicit in **Axiom PE** can be given an explicit form. This section is devoted to studying the form of the *class of externality functions*, associated with the class of decomposable literacy indices satisfying **Axioms N, RM, PE** and **SI**.

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 (17.6)–(17.9). 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 [17.8]) $= rM(\{(1, 1)\}) + (n - r)M(\{(0, 1)\})$ (by [17.6]) $= r$ (by [17.9]), and $L(\{(r, n)\}) \geq (r/n)$. Thus, we have:

$$L(\{(r, n)\}) \geq (r/n) \text{ for all } (r, n) \in \mathbb{X} \qquad (17.23)$$

and so by (17.20),

$$E(r, n) \geq 0 \text{ for all } (r, n) \in \mathbb{X}.$$

As noted earlier, 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 can, therefore, define an *externality function* $e : \mathbb{Y} \to \mathbb{R}$ associated with L as:

$$e(r/n) \equiv f(r/n) - (r/n) = L(\{(r, n)\}) - (r/n) \text{ for all } (r, n) \in \mathbb{X} \quad (17.24)$$

and so by (17.20),

$$E(r, n) = ne(r/n) \text{ for all } (r, n) \in \mathbb{X} \qquad (17.25)$$

Clearly, using (17.23) and (17.24), e is a function from $\mathbb{Y}$ to $\mathbb{R}_+$, and using (17.25), E is a function from $\mathbb{X}$ to $\mathbb{R}_+$. We refer to E as an *externality measure*.

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 $e(z) + z$. To summarise, e is a function from $\mathbb{Y}$

to $\mathbb{R}_+$, which satisfies the following:

$$\left.\begin{array}{l} 1.\ e(0) = e(1) = 0 \\ 2.\ e(z) + z \text{ satisfies property } \mathbf{M} \text{ on } \mathbb{Y} \\ 3.\ e \text{ satisfies property } \mathbf{C} \text{ on } \mathbb{Y} \end{array}\right\} \tag{17.26}$$

These three properties, in fact, *characterise* 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 (17.26), and $E : \mathbb{X} \to \mathbb{R}$ is of the form given in (17.25), then $L : \mathbb{U} \to \mathbb{R}$ defined by (17.21), where $n = (n_1 + \cdots + n_m)$, is a decomposable literacy index which satisfies axioms **N, RM, PE** and **SI.**

To verify this claim, note that since L is defined by (17.21), $E : \mathbb{X} \to \mathbb{R}$ must satisfy (17.20), so that

$$L(\{(r, n)\}) = [E(r, n)/n] + (r/n) = e(r/n) + (r/n) \text{ for all } (r, n) \in \mathbb{X}, \tag{17.27}$$

the second equality in (17.27) following from (17.25). Further, since L is defined by (17.21), we have:

$$L(\{(r_1, n_1), \ldots, (r_m, n_m)\}) = \sum_{i=1}^{m} (n_i/n)L(\{(r_i, n_i)\}) \tag{17.28}$$

for all $\{(r_1, n_1), \ldots, (r_m, n_m)\} \in \mathbb{U}$, with $n = (n_1 + \cdots + n_m)$, as verified in (17.22), so that L is a decomposable literacy index. Now, using (17.27), axiom **N** follows from (17.26) (1), axiom **RM** follows from (17.26)(2). Axiom **SI** follows directly from (17.27). 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 first equality uses (17.27), the single inequality follows from (17.26)(3), and the last two equalities use (17.27) and (17.28) respectively.

We can summarise the findings in the following characterisation of decomposable literacy indices in terms of externality functions.

Theorem 3. *1. If $L : \mathbb{U} \to \mathbb{R}$ is any decomposable literacy index, then it can be written as:*

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) = \frac{\sum_{i=1}^{m}[r_i + E(r_i, n_i)]}{n}, \tag{17.29}$$

where $(n_1 + \cdots + n_m) = n$ and $E : \mathbb{X} \to \mathbb{R}$ is defined by:

$$E(r, n) \equiv nL(\{(r, n)\}) - r \text{ for all } (r, n) \in \mathbb{X}. \tag{17.30}$$

Conversely, if E is a map from $\mathbb{X}$ to $\mathbb{R}$, and $L : \mathbb{U} \to \mathbb{R}$ is defined by (17.29), then E satisfies (17.30) and L is a decomposable literacy index.

*2. If $L : \mathbb{U} \to \mathbb{R}$ is a decomposable literacy index, which satisfies Axioms **N, RM, PE** and **SI**, then L can be written as (17.29), where $(n_1 + \cdots + n_m) = n$ and $E : \mathbb{X} \to \mathbb{R}$ is defined by (17.30), and there is a function $e : \mathbb{Y} \to \mathbb{R}_+$ satisfying (17.26), such that $E : \mathbb{X} \to \mathbb{R}$ takes the form:*

$$E(r, n) = ne(r/n) \text{ for all } (r, n) \in \mathbb{X} \tag{17.31}$$

*Conversely, if e is a function from $\mathbb{Y}$ to $\mathbb{R}_+$, which satisfies (17.26), and $E : \mathbb{X} \to \mathbb{R}$ is of the form given in (17.31), then $L : \mathbb{U} \to \mathbb{R}$ defined by (17.29), where $n = (n_1 + \cdots + n_m)$, is a decomposable literacy index which satisfies axioms **N, RM, PE** and **SI**.*

Remark 4.

1. *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 satisfies (17.26).*
2. *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 (17.26).*
3. *The literacy index proposed by Subramanian (2004: 456) arises by defining the externality function to be: $e(z) = z(1 - z)$. It can be checked that this e satisfies (17.26).*

Alternative Characterisations

Alternative axiomatic characterisations of literacy indices have been proposed in the literature, with the aim of generalising the Basu–Foster analysis to appropriate *classes* of externality functions. In this section, we compare our characterisation results, provided in Theorems 1 and 3, with those proposed by Valenti (2002) and Dutta (2004). The basic framework is somewhat different from ours, and the axiom systems used also differ from ours in some respects. It turns out that the characterisation result of Valenti (2002) is essentially similar to the result stated in Theorem 3. In contrast, the result of Dutta (2004) is different from ours in a crucial respect, so that the result on the distribution of literates (contained in our Theorem 2) does not follow from his characterisation. The connection between our analysis (of the second and third sections) and these alternative approaches might provide the reader with a better perspective of axiomatic characterisations of literacy indices.

The Alternative Framework

The framework used by both Valenti (2002) and Dutta (2004) may be presented as follows. A household is described by the number of literates (r) and the number of illiterates (s).[5] This description might appear to be equivalent to the one presented in the second section. However, a difference arises when considering a *change* in literates only in a household. Thus, an increase of literates in a household depicted as (r, n), with n remaining constant, reflects a situation in which some illiterates in the household become literates. On the other hand, an increase of literates in a household depicted as (r, s), with s remaining constant, reflects a situation in which new members are added to the household (so that the total number of persons in the household $(r + s)$ goes up) and all of the newly added members are literates.

Let us define a set:

$$\mathbb{Z} = \{(r, s) \in \mathbb{M} \times \mathbb{M} : (r + s) \in \mathbb{N}\}$$

In this alternative framework, a *household* is a pair $(r, s) \in \mathbb{Z}$. Here r is to be interpreted as the number of *literate* individuals in the household, and s is to be interpreted as the number of *illiterate* individuals in the household. Thus, $n \equiv r + s$ is the total number of 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 $\{(r_1, s_1), ..., (r_m, s_m)\}$.

[5] Actually, Valenti (2002) employs the more extensive form representation of the household used by Basu and Foster (1998), but her analysis can be adequately presented as described here.

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

$$\mathbb{V} = \bigcup_{k=1}^{\infty} \mathbb{Z}^k \tag{17.32}$$

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

$$\Lambda(\{(r_1, s_1), ..., (r_m, s_m)\}) = \sum_{i=1}^{m} (n_i/n)\Lambda(\{(r_i, s_i)\}), \tag{17.33}$$

where $n_i = r_i + s_i$ for all $i \in \{1, ..., m\}$ and $n = (n_1 + \cdots + n_m)$. Both Valenti (2002) and Dutta (2004) impose the restriction (17.33) in their analysis, by including a 'decomposability axiom' in their set of axioms; I prefer not to state it as an axiom, but as a restriction on the scope of the analysis.

To relate this framework to the one we have used in the second section, we can define $L : \mathbb{U} \to \mathbb{R}$ in terms of Λ by:

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) = \sum_{i=1}^{m} (n_i/n)\Lambda(\{(r_i, n_i - r_i)\}), \tag{17.34}$$

where $n = (n_1 + \cdots + n_m)$. Then, clearly, we have:

$$L(\{(r, n)\}) = \Lambda(\{(r, n - r)\}) \text{ for all } (r, n) \in \mathbb{X} \tag{17.35}$$

and so (using (17.34) and (17.35)) $L : \mathbb{U} \to \mathbb{R}$ is decomposable:

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) = \sum_{i=1}^{m} (n_i/n)L(\{(r_i, n_i)\}) \tag{17.36}$$

in exactly the sense described in (17.2).

Both Valenti (2002) and Dutta (2004) include a ‘normalisation axiom’ in their axiom systems;[6] this can be written as:

$$\Lambda(\{(1,0)\}) = 1;\ \Lambda(\{(0,1)\}) = 0. \tag{17.37}$$

This translates, using (17.35), to:

$$L(\{(1,1)\}) = 1;\ L(\{(0,1)\}) = 0, \tag{17.38}$$

which is precisely our **Axiom N**, described in section 2.3.

Beyond this point, the approaches of Valenti (2002) and Dutta (2004) differ from each other, and we highlight the differences in the next two subsections.

Monotonicity and Equality

Valenti (2002) uses a ‘monotonicity axiom’ which can be stated as follows:

$$\Lambda(\{(r+1, s-1)\}) > \ \Lambda(\{(r,s)\}) \ \text{ for all } (r,s) \in \mathbb{Z} \text{ with } s \in \mathbb{N}, \tag{17.39}$$

that is, the household on the left side of (17.39) is obtained from the household on the right side of (17.39) by a ‘simple increment’, involving an illiterate becoming literate. Notice that the total number of persons in the household remains unchanged in such a comparison, and we see that (17.39) translates, using (17.35), to:

$$L(\{(r+1, n)\}) > \ L(\{(r,n)\}) \ \text{ for all } (r,n) \in \mathbb{X}, \text{ with } r < n \tag{17.40}$$

This yields our **Axiom RM**, described earlier, by induction.

[6] Actually, Valenti (2002) and Dutta (2004) state a stronger normalisation axiom, which assigns a literacy index of 1 to all households which have only literates, and a literacy index of 0 to all households which have only illiterates. This is superfluous in Valenti’s axiom system, because of (17.33) and (17.41) (1). It is superfluous in Dutta’s axiom system because of (17.33) and (17.63).

Crucial to Valenti's axiom system is the 'equality axiom', which can be stated as follows.[7] If $(r_1, s_1) \in \mathbb{Z}$, and $(r_2, s_2) \in \mathbb{Z}$, then:

$$\left.\begin{array}{l} 1.\ \Lambda(\{(r_1{+}r_2), (s_1{+}s_2)\}){=}\Lambda(\{(r_1, s_1), (r_2, s_2)\}) \text{ when } (r_1, s_1){=}(r_2, s_2) \\ 2.\ \Lambda(\{(r_1{+}r_2), (s_1{+}s_2)\}) > \Lambda(\{(r_1, s_1), (r_2, s_2)\}) \text{ when } (r_1, s_1) \neq (r_2, s_2) \end{array}\right\} \tag{17.41}$$

Note that (17.41)(1) translates, using (17.33) and (17.34), to the following condition. If $(r_1, n_1) \in \mathbb{X}$ and $(r_2, n_2) \in \mathbb{X}$, then:

$$L(\{(r_1 + r_2), (n_1 + n_2)\}) = L(\{(r_1, n_1), (r_2, n_2)\}) \text{ when } (r_1, n_1) = (r_2, n_2). \tag{17.42}$$

Using (17.36) and (17.42), and using induction, it follows that if $(r, n) \in \mathbb{X}$ and $k \in \mathbb{N}$, then:

$$L(\{(kr, kn)\}) = L(\{(r, n)\}), \tag{17.43}$$

which is our **Axiom SI**, described earlier.

Also, (17.41) implies, using (17.33) and (17.34), that the following condition holds. If $(r_1, n_1) \in \mathbb{X}$ and $(r_2, n_2) \in \mathbb{X}$, then:

$$L(\{(r_1 + r_2), (n_1 + n_2)\}) \geq L(\{(r_1, n_1), (r_2, n_2)\}), \tag{17.44}$$

which is our **Axiom PE**, described earlier. Thus, Valenti's axiom system on $\Lambda : \mathbb{V} \to \mathbb{R}$ implies that $L : \mathbb{U} \to \mathbb{R}$, defined by (17.34), is a decomposable literacy index, which satisfies our axioms **N**, **RM**, **PE** and **SI**.

It follows that, under Valenti's axiom system, the literacy index L (associated with Λ) satisfies (17.3) of Theorem 1 and, therefore, its principal application on the distribution of literacy rates (Theorem 2) holds.

Valenti's principal contribution (contained in her Theorem, p. 14) is a characterisation of literacy indices $\Lambda : \mathbb{V} \to \mathbb{R}$ satisfying the decomposability, normalisation, monotonicity and equality axioms (written as (17.33), (17.37), (17.39) and (17.41)) in terms of *externality functions* $\alpha : \mathbb{Q}_+ \to [0, 1)$ satisfying concavity, monotonicity and an end-point condition. [See remark (2) following Theorem 4 for a more precise statement].

[7] The wording is somewhat different in Valenti (2002). However, note that a single household always satisfies 'perfect literacy equality', according to her terminology. And, the society appearing on the right side of (17.41) satisfies 'perfect literacy equality', when $(r_1, s_1) = (r_2, s_2)$, while it violates 'perfect literacy equality', when $(r_1, s_1) \neq (r_2, s_2)$.

We show how a characterisation result similar to Valenti's can be obtained, by using our characterisation result in Theorem 3. This will clarify the connection between the externality function $e : \mathbb{Y} \to \mathbb{R}_+$ appearing in Theorem 3 and the externality function $\alpha : \mathbb{Q}_+ \to [0, 1)$ which appears in her characterisation result. In accordance with our own axiom system used in the second section, we work with a slightly weaker axiom set than Valenti's. Specifically, we replace her equality axiom (17.41), with the following weaker version.[8] If $(r_1, s_1) \in \mathbb{Z}$, and $(r_2, s_2) \in \mathbb{Z}$, then:

$$\left.\begin{array}{l} 1.\ \Lambda(\{(r_1{+}r_2), (s_1{+}s_2)\}){=}\Lambda(\{(r_1, s_1), (r_2, s_2)\}) \text{ when } (r_1, s_1){=}(r_2, s_2) \\ 2.\ \Lambda(\{(r_1{+}r_2), (s_1{+}s_2)\}) \geq \Lambda(\{(r_1, s_1), (r_2, s_2)\}) \text{ when } (r_1, s_1) \neq (r_2, s_2) \end{array}\right\} \tag{17.45}$$

Our preference for this weaker version arises from the fact that the characterisation result on the class of literacy indices can thereby accommodate the standard literacy rate and the Basu–Foster literacy index as special cases. (For further elaboration on this point see remark (3) following Theorem 4).

Theorem 4. [Valenti]

1. *Suppose $\Lambda : \mathbb{V} \to \mathbb{R}$ is a literacy index satisfying the axioms (17.33), (17.34), (17.39) and (17.45). Then, $\Lambda : \mathbb{V} \to \mathbb{R}$ can be written as*

$$\Lambda(\{(r_1, s_1), ..., (r_m, s_m)\}) = \frac{\sum_{i=1}^{m}[r_i + \eta(r_i, s_i)]}{n}, \tag{17.46}$$

where $n_i = r_i + s_i$ for all $i \in \{1, ..., m\}$ and $(n_1 + \cdots + n_m) = n$ and $\eta : \mathbb{Z} \to \mathbb{R}$ satisfies:

$$\eta(r, s) = (r + s)\Lambda(\{(r, s)\}) - r \text{ for all } (r, s) \in \mathbb{Z}. \tag{17.47}$$

Further, there is a function $\alpha : \mathbb{Q}_+ \to [0, 1)$, satisfying (a) $\alpha(0) = 0$, (b) If $z, z' \in \mathbb{Q}_+$, and $z' > z$, then $\alpha(z') \geq \alpha(z)$, (c) α satisfies

[8] Note that the weaker axiom set on $\Lambda : \mathbb{V} \to \mathbb{R}$, given by (17.33), (17.37), (17.39) and (17.45), is enough to ensure that $L : \mathbb{U} \to \mathbb{R}$, defined by (17.34), is a decomposable literacy index, which satisfies our axioms **N**, **RM**, **PE** and **SI**. Thus, under this weaker axiom system, the literacy index L (associated with Λ) satisfies (17.3) of Theorem 1 and, therefore, the result on the distribution of literacy rates (Theorem 2) holds.

property **C** *on* $\mathbb{Q}_+$, *such that* $\eta : \mathbb{Z} \to \mathbb{R}$ *takes the form:*

$$\eta(r, s) = \begin{cases} s\alpha(r/s) \text{ for all } (r, s) \in \mathbb{Z}, \text{ with } s \in \mathbb{N} \\ 0 \qquad\quad \text{for all } (r, s) \in \mathbb{Z}, \text{ with } s = 0 \end{cases} \tag{17.48}$$

2. *Conversely, if there is a function* $\alpha : \mathbb{Q}_+ \to (0, 1)$, *satisfying (a)* $\alpha(0) = 0$; *(b) if* $z, z' \in \mathbb{Q}_+$, *and* $z' > z$, *then* $\alpha(z') \geq \alpha(z)$; *(c)* α *satisfies property* **C** *on* $\mathbb{Q}_+$, *and* $\eta : \mathbb{Z} \to \mathbb{R}$ *is of the form given in (17.48), then* $\Lambda : \mathbb{V} \to \mathbb{R}$ *defined by (17.46), where* $n_i = r_i + s_i$ *for all* $i \in \{1, ..., m\}$ *and* $n = (n_1 + \cdots + n_m)$, *is a literacy index which satisfies axioms (17.33), (17.37), (17.39) and (17.45).*

Proof: 1. Given $\Lambda : \mathbb{V} \to \mathbb{R}$ satisfying the axioms (17.33), (17.37), (17.39) and (17.45), we define $L : \mathbb{U} \to \mathbb{R}$ by (17.34) where $n = (n_1 + \cdots + n_m)$. Then, as verified earlier, $L : \mathbb{U} \to \mathbb{R}$ is a decomposable literacy index satisfying axioms **N**, **RM**, **PE** and **SI**. Applying Theorem 3, L can be written as:

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) = \frac{\sum_{i=1}^{m} [r_i + E(r_i, n_i)]}{n}, \tag{17.49}$$

where $(n_1 + \cdots + n_m) = n$ and $E : \mathbb{X} \to \mathbb{R}$ is defined by:

$$E(r, n) \equiv nL(\{(r, n)\}) - r \text{ for all } (r, n) \in \mathbb{X}. \tag{17.50}$$

Further, there is a function $e : \mathbb{Y} \to \mathbb{R}_+$ satisfying (17.27), such that $E : \mathbb{X} \to \mathbb{R}$ takes the form:

$$E(r, n) = ne(r/n) \text{ for all } (r, n) \in \mathbb{X}. \tag{17.51}$$

Using (17.33), (17.34) and (17.49), $\Lambda : \mathbb{V} \to \mathbb{R}$ can be written as:

$$\Lambda(\{(r_1, s_1), ..., (r_m, s_m)\}) = \frac{\sum_{i=1}^{m} [r_i + \eta(r_i, s_i)]}{n},$$

where $n_i = r_i + s_i$ for all $i \in \{1, ..., m\}$ and $(n_1 + \cdots + n_m) = n$ and $\eta : \mathbb{Z} \to \mathbb{R}$ is defined by:

$$\eta(r, s) = E(r, r + s) \text{ for all } (r, s) \in \mathbb{Z}, \tag{17.52}$$

Then, by (17.34), (17.40) and (17.52),

$$\eta(r, s) = (r + s)L(\{(r, r + s)\}) - r = (r + s)\Lambda(\{(r, s)\}) - r \text{ for all } (r, s) \in \mathbb{Z},$$

which is (17.47).

Now, define $\alpha : \mathbb{Q}_+ \to \mathbb{R}_+$ by:

$$\alpha(w) = (1 + w)e(w/(1 + w)) \text{ for all } w \in \mathbb{Q}_+$$

Then, since $e : \mathbb{Y} \to \mathbb{R}_+$ satisfies (17.30), Proposition 3 of the Appendix can be used to show that $\alpha : \mathbb{Q}_+ \to \mathbb{R}_+$ maps from $\mathbb{Q}_+$ to $[0, 1)$, and that (a) $\alpha(0) = 0$; (b) If $w, w' \in \mathbb{Q}_+$, and $w' > w$, then $\alpha(w') \geq \alpha(w)$; (c) α satisfies property **C** on $\mathbb{Q}_+$.

Further, using (17.51) and (17.52),

$$\eta(r, s) = E(r, r + s) = (r + s)e(r/(r + s)) \text{ for all } (r, s) \in \mathbb{Z}. \qquad (17.53)$$

Thus, for all $(r, s) \in \mathbb{Z}$ with $s = 0$, we have $\eta(r, s) = 0$. And, for all $(r, s) \in \mathbb{Z}$ with $s \in \mathbb{N}$, we have:

$$\eta(r, s)/s = [(r + s)/s]e(r/(r + s)) = \alpha(r/s) \text{ for all } (r, s) \in \mathbb{Z}$$

which establishes (17.48).

2. Since $\alpha : \mathbb{Q}_+ \to (0, 1)$, satisfies (a) $\alpha(0) = 0$; (b) If $z, z' \in \mathbb{Q}_+$, and $z' > z$, then $\alpha(z') \geq \alpha(z)$; (c) α satisfies property **C** on $\mathbb{Q}_+$, Proposition 4 in the Appendix can be applied to show that the function $e : \mathbb{Y} \to \mathbb{R}_+$ defined by:

$$e(z) = \begin{cases} \alpha(z/(1 - z))(1 - z) & \text{for } z \neq 1 \\ 0 & \text{for } z = 1 \end{cases} \qquad (17.54)$$

has the following properties: (a) $e(0) = e(1) = 0$, (b) $e(z) + z$ satisfies property **M** on $\mathbb{Y}$, (c) e satisfies property **C** on $\mathbb{Y}$. Thus, by (2) of Theorem 3, if $E : \mathbb{X} \to \mathbb{R}$ is defined by:

$$E(r, n) = ne(r/n) \text{ for all } (r, n) \in \mathbb{X} \qquad (17.55)$$

and $L : \mathbb{U} \to \mathbb{R}$ is defined as:

$$L(\{(r_1, n_1), ..., (r_m, n_m)\}) = \frac{\sum_{i=1}^{m} [r_i + E(r_i, n_i)]}{n}, \qquad (17.56)$$

where $(n_1 + \cdots + n_m) = n$, then $L : \mathbb{U} \to \mathbb{R}$ is a decomposable literacy index which satisfies axioms **N, RM, PE** and **SI**.

Since $\eta : \mathbb{Z} \to \mathbb{R}$ is of the form (17.48), we must have for all $(r, s) \in \mathbb{Z}$ with $s \in \mathbb{N}$,

$$\begin{aligned} \frac{\eta(r, s)}{(r + s)} &= \frac{s}{(r + s)}\alpha(r/s) = \frac{(n - r)}{n}\alpha(r/(n - r)) \\ &= (1 - \frac{r}{n})\alpha\left(\frac{(r/n)}{1 - (r/n)}\right) = e(r/n) = \frac{E(r, n)}{n}, \end{aligned} \qquad (17.57)$$

where $n = r + s$, and we have used (17.54) and (17.55) in the second line of (17.57). Also, for all $(r, s) \in \mathbb{Z}$ with $s = 0$, we have:

$$\frac{\eta(r, s)}{(r+s)} = \frac{\eta(r, 0)}{r} = 0 = \frac{E(n, n)}{n} = \frac{E(r, n)}{n}, \qquad (17.58)$$

where $n = r + s = r$.

Since $\Lambda : \mathbb{V} \to \mathbb{R}$ is defined by (17.46), where $n_i = r_i + s_i$ for all $i \in \{1, ..., m\}$ and $n = (n_1 + \cdots + n_m)$, we must have:

$$\eta(r, s) = (r+s)\Lambda(\{(r, s)\}) - r \text{ for all } (r, s) \in \mathbb{Z}, \qquad (17.59)$$

so using this in (17.46), we see that (17.33) must hold. Further, we obtain by using (17.57) and (17.58),

$$\begin{aligned}\Lambda(\{(r_1, s_1), ..., (r_m, s_m)\}) &= \frac{\sum_{i=1}^{m}[r_i + \eta(r_i, s_i)]}{n} = \frac{\sum_{i=1}^{m}[r_i + E(r_i, n_i)]}{n} \\ &= L(\{(r_1, n_1), ..., (r_m, n_m)\}). \qquad (17.60)\end{aligned}$$

It remains to verify that Λ satisfies (17.37), (17.39) and (17.45). Since L satisfies axiom **N**, we obtain from (17.60) that $\Lambda(\{(1, 0)\}) = L(\{(1, 1)\}) = 1$ and $\Lambda(\{(0, 1)\}) = L(\{(0, 1)\}) = 0$, so that (17.37) must hold. Since L satisfies axiom **RM**, we obtain from (17.60) that for all $(r, s) \in \mathbb{Z}$ with $s \in \mathbb{N}$,

$$\Lambda(\{(r+1, s-1)\}) = L(\{(r+1, r+s)\}) > L(\{(r, r+s)\}) = \Lambda(\{(r, s)\})$$

so that (17.39) must hold. Since L satisfies axiom **SI**, when $(r_1, s_1) \in \mathbb{Z}$, and $(r_2, s_2) \in \mathbb{Z}$, with $(r_1, s_1) = (r_2, s_2)$, we obtain from (17.60):

$$\begin{aligned}&\Lambda(\{(r_1 + r_2), (s_1 + s_2)\}) \\ &\quad = \Lambda(\{(2r_1, 2s_1)\}) = L(\{(2r_1, 2r_1 + 2s_1)\}) \\ &\quad = L(\{(r_1, r_1 + s_1)\}) = (\tfrac{1}{2})L(\{(r_1, r_1 + s_1)\}) + (\tfrac{1}{2})L(\{(r_2, r_2 + s_2)\}) \\ &\quad = (\tfrac{1}{2})\Lambda(\{(r_1, s_1)\}) + (\tfrac{1}{2})\Lambda(\{(r_2, s_2)\}) = \Lambda(\{(r_1, s_1)\}, \{(r_2, s_2)\}),\end{aligned} \qquad (17.61)$$

where we have used the fact that Λ satisfies (17.33) in the third line of (17.61). Thus, (17.45)(1) must hold. Finally, since L satisfies axiom **PE**, when $(r_1, s_1) \in \mathbb{Z}$, and $(r_2, s_2) \in \mathbb{Z}$, with $(r_1, s_1) \neq (r_2, s_2)$, we obtain from (17.60):

$$\begin{aligned}\Lambda(\{(r_1 + r_2), (s_1 + s_2)\}) &= L(\{(r_1 + r_2), (n_1 + n_2)\}) \\ &\geq L(\{(r_1, n_1), (r_2, n_2)\}) = \Lambda(\{(r_1, s_1), (r_2, s_2)\}),\end{aligned}$$

where we have denoted $(r_1 + s_1)$ by n_1 and $(r_2 + s_2)$ by n_2. Thus, (17.45)(2) must hold.

Remark 5.

1. *The externality functions* $e : \mathbb{Y} \to \mathbb{R}_+$ *and* $\alpha : \mathbb{Q}_+ \to [0, 1)$ *appearing in Theorems 3 and 4 respectively can be interpreted 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* $\alpha(r/s) \equiv \alpha(r/(n-r))$ *measures the externality generated as a proportion of the fraction who are illiterate in the household. Thus, if the household is* $(r, n) = (2, 3)$, *then the fraction of the illiterates in the household is* $(1/3)$, *and* $\alpha(r/s) = \alpha(2)$ *measures* $[e(2/3)/(1/3)]$. *Note that* $\alpha(r/s)$ does not *measure the externality per illiterate. For the household* $(r, n) = (2, 3)$, *the externality per illiterate is* $e(2/3)$ *itself, since there is only one illiterate, while* $\alpha(r/s) = \alpha(2) = e(2/3)/(1/3)$.

   The relationship between the functions $e : \mathbb{Y} \to \mathbb{R}_+$ *and* $\alpha : \mathbb{Q}_+ \to [0, 1)$ *is used in the proof of Theorem 4, which relates them technically. This can be conveniently illustrated with the specific externality function* $\alpha : \mathbb{Q}_+ \to [0, 1)$ *proposed by Subramanian (2004: 456), where* $\alpha(y) = y/(1+y)$; *that is,*

$$\alpha(r/s) = \frac{r}{r+s} = \frac{(r/s)}{(r/s)+1}.$$

   The corresponding externality function $e : \mathbb{Y} \to \mathbb{R}_+$ *is then given by:*

$$e(z) = (1-z)\alpha\left(\frac{z}{1-z}\right) = (1-z)z \;\; \text{so that } e(r/n) = \frac{r}{n}\left(1 - \frac{r}{n}\right)$$

   and

$$ne(r/n) = \frac{r(n-r)}{n} = \frac{rs}{r+s} = s\alpha(r/s).$$

2. *The statement of the actual characterisation theorem presented in Valenti (2002: 14) differs from the statement of Theorem 4 in the following respects. The equality axiom (17.45) is replaced by the stronger equality axiom (17.41) in Valenti's Theorem. Further, for the class of externality functions* $\alpha : \mathbb{Q}_+ \to [0, 1)$, *property (b) is replaced by the stronger property:*

$$(b+) \text{ If } z, z' \in \mathbb{Q}_+, \text{ and } z' > z, \text{ then } \alpha(z') > \alpha(z)$$

and property (c) is replaced by the stronger property:

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

in Valenti's theorem.

3. *The traditional index of literacy, the literacy rate, arises by defining the externality function (in Theorem 4) to be:* $\alpha(z) = 0$ *for all* $z \in \mathbb{Q}_+$. *This* α *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 (in Theorem 4) to be:* $\alpha(z) = \alpha$ *for all* $z \in \mathbb{Q}_{++}$ *and* $\alpha(z) = 0$ *for* $z = 0$. *Thus, this* α *also violates property* (*b*+) *and property* (*c*+).
4. *Valenti's characterisation is in terms of externality functions,* α, *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 Theorem 2* directly *from her characterisation theorem, or from Theorem 4.*

Weak Externality

Dutta (2004) examines literacy indices which satisfy decomposability (that is, [17.33]) and the normalisation axiom (that is, [17.37]), and imposes in addition a *weak externality* axiom. We will find it convenient to present this axiom in two parts.

The first part of the weak externality axiom can be stated as follows.

$$\left.\begin{array}{l} 1.\ \Lambda(\{(r_1+r_2,0)\}) = \Lambda(\{(r_1,0),(r_2,0)\}) \text{ when } (r_1,0),(r_2,0) \in \mathbb{Z} \\ 2.\ \Lambda(\{(0,s_1+s_2)\}) = \Lambda(\{(0,s_1),(0,s_2)\}) \text{ when } (0,s_1),(0,s_2) \in \mathbb{Z} \end{array}\right\} \quad (17.62)$$

So, when a household consisting only of literates is split up into two households, this is an 'externality neutral' split and, therefore there is no loss of literacy from such a split. A similar explanation holds if a household consisting only of illiterates is split in two.

Given the decomposability of Λ, and the normalisation axiom, it is easy to verify by induction that (17.62) is equivalent to:

$$\Lambda(\{(r,0)\}) = 1 \text{ for all } r \in \mathbb{N};\ \Lambda(\{(0,s)\}) = 0 \text{ for all } s \in \mathbb{N} \quad (17.63)$$

that is, it combines the normalisation axiom (17.37) with a scale invariance property for some particular single household societies.

The second part of the weak externality axiom considers a split of a household which has both literates and illiterates. If the household is split up into two households in such a way that one of these two households has all the illiterates (that is, only the literates are split up), then this is an 'externality reducing' split and, consequently, there is a loss of literacy when such a split occurs.

$$\Lambda(\{(r_1 + r_2, s)\}) > \Lambda(\{(r_1, s), (r_2, 0)\}) \text{ when } (r_1, s), (r_2, 0) \in \mathbb{Z}, s \in \mathbb{N} \tag{17.64}$$

Note that this is a strictly weaker restriction than that imposed in (17.41)(2) (the second part of the equality axiom) by Valenti (2002). It is not directly comparable to (17.45)(2) because of the weak inequality in that version of the equality axiom.

Literacy indices satisfying (17.33), (17.37), (17.62) and (17.64) (that is, the decomposability, normalisation and weak externality axioms) have been characterised by Dutta (2004: 75; Theorem 1) in terms of a class of externality functions. We can state and prove his result as follows.

Theorem 5. [Dutta] *1. Suppose* $\Lambda : \mathbb{V} \to \mathbb{R}$ *is a literacy index satisfying the axioms (17.33), (17.37), (17.62) and (17.64). Then,* $\Lambda : \mathbb{V} \to \mathbb{R}$ *can be written as*

$$\Lambda(\{(r_1, s_1), ..., (r_m, s_m)\}) = \frac{\sum_{i=1}^{m}[r_i + a(r_i, s_i)]}{n}, \tag{17.65}$$

where $n_i = r_i + s_i$ *for all* $i \in \{1, ..., m\}$ *and* $(n_1 + \cdots + n_m) = n$ *and* $a : \mathbb{Z} \to \mathbb{R}$ *is defined by:*

$$a(r, s) = (r + s)\Lambda(\{(r, s)\}) - r \text{ for all } (r, s) \in \mathbb{Z}. \tag{17.66}$$

Further, $a : \mathbb{Z} \to \mathbb{R}$, *satisfies (a)* $a(r, 0) = 0 = a(0, s)$ *for all* $r, s \in \mathbb{N}$; *(b)* $a(r, s) \geq 0$ *for all* $(r, s) \in \mathbb{Z}$; *(c)* $a(r + r', s) > a(r, s)$ *for all* $(r, s) \in \mathbb{Z}$, *and* $(r', s) \in \mathbb{N}^2$.

2. Conversely, if there is a function $a : \mathbb{Z} \to \mathbb{R}$, *satisfying (a)* $a(r, 0) = 0 = a(0, s)$ *for all* $r, s \in \mathbb{N}$; *(b)* $a(r, s) \geq 0$ *for all* $(r, s) \in \mathbb{Z}$; *(c)* $a(r + r', s) > a(r, s)$ *for all* $(r, s) \in \mathbb{Z}$, *and* $(r', s) \in \mathbb{N}^2$, *then* $\Lambda : \mathbb{V} \to \mathbb{R}$ *defined by (17.65), where* $n_i = r_i + s_i$ *for all* $i \in \{1, ..., m\}$ *and* $n = (n_1 + \cdots + n_m)$, *is a literacy index which satisfies axioms (17.33), (17.37), (17.62) and (17.64).*

Proof: 1. Since $\Lambda : \mathbb{V} \to \mathbb{R}$ is a literacy index satisfying (17.33), we can write (using the notation $n_i = r_i + s_i$ for all $i \in \{1, ..., m\}$ and

$n = (n_1 + \cdots + n_m))$,

$$\Lambda(\{(r_1, s_1), ..., (r_m, s_m)\}) = \sum_{i=1}^{m} (n_i/n)\Lambda(\{(r_i, s_i)\}) = \frac{\sum_{i=1}^{m}[r_i + n_i\Lambda(\{(r_i, s_i)\}) - r_i]}{n} = \frac{\sum_{i=1}^{m}[r_i + a(r_i, s_i)]}{n}, \quad (17.67)$$

where $a : \mathbb{Z} \to \mathbb{R}$ is defined by (17.66). Thus, it only remains to verify the properties (a),(b) and (c) of the function a.

Using (17.33), (17.37) and (17.62), we know that (17.63) must hold. Using (17.63), (a) follows from (17.66). To verify (b), note that this is clearly true when $(r, s) \in \mathbb{Z}$ and either $r = 0$ or $s = 0$, by using (a). Thus, consider $(r, s) \in \mathbb{Z}$ with $(r, s) \in \mathbb{N}^2$. Then, by (17.33) and (17.64),

$$\Lambda(\{(r, s)\}) > \Lambda(\{(r, 0), (0, s)\}) = \frac{r}{r+s}\Lambda(\{(r, 0)\}) + \frac{s}{r+s}\Lambda(\{(0, s)\}) = \frac{r}{r+s}, \quad (17.68)$$

the last line of (17.68) following from (17.63). Now, it follows from (17.66) that $a(r, s) > 0$. This establishes (b).

To establish (c), note that for $(r, s) \in \mathbb{Z}$, and $(r', s) \in \mathbb{N}^2$, we have by (17.33) and (17.64),

$$\Lambda(\{(r + r', s)\}) > \Lambda(\{(r, s), (r', 0)\}) = \frac{r+s}{r+r'+s}\Lambda(\{(r, s)\}) + \frac{r'}{r+r'+s}\Lambda(\{(r', 0)\}) = \frac{r+s}{r+r'+s}\Lambda(\{(r, s)\}) + \frac{r'}{r+r'+s}. \quad (17.69)$$

Thus, we get, from (17.66) and (17.69),

$$a(r + r', s) = (r + r' + s)\Lambda(\{(r + r', s)\}) - (r + r') > (r + s)\Lambda(\{(r, s)\}) + r' - (r + r') = (r + s)\Lambda(\{(r, s)\}) - r = a(r, s)$$

which proves property (c).

2. Since $\Lambda : \mathbb{V} \to \mathbb{R}$ is defined by (17.65), where $n_i = r_i + s_i$ for all $i \in \{1, ..., m\}$ and $n = (n_1 + \cdots + n_m)$, we know that $a : \mathbb{Z} \to \mathbb{R}$ must

satisfy (17.66). Using (17.66) in (17.65), we obtain:

$$\Lambda(\{(r_1, s_1), ..., (r_m, s_m)\}) = \sum_{i=1}^{m} (n_i/n)\Lambda(\{(r_i, s_i)\})$$

so that Λ satisfies (17.33). Given property (a) of a, we get (17.63) directly by using (17.66). This establishes (17.37) as well as (17.62) since, as we have already noted under (17.33), the conditions (17.62) and (17.63) are equivalent.

It remains to establish (17.64). When $(r_1, s), (r_2, 0) \in \mathbb{Z}, s \in \mathbb{N}$, we have by property ($c$) of a,

$$\begin{aligned}(r_1 + r_2 + s)\Lambda(\{(r_1 + r_2, s)\}) - (r_1 + r_2) &= a(r_1 + r_2, s) \\ &> a(r_1, s) \\ &= (r_1 + s)\Lambda(\{(r_1, s)\}) - r_1 \end{aligned} \tag{17.70}$$

so that, by using (17.33), (17.63) and (17.70),

$$\begin{aligned}\Lambda(\{(r_1 + r_2, s)\}) &> \frac{(r_1 + s)}{(r_1 + r_2 + s)}\Lambda(\{(r_1, s)\}) + \frac{r_2}{(r_1 + r_2 + s)} \\ &= \frac{(r_1 + s)}{(r_1 + r_2 + s)}\Lambda(\{(r_1, s)\}) + \frac{r_2}{(r_1 + r_2 + s)}\Lambda(\{(r_2, 0)\}) \\ &= \Lambda(\{(r_1, s), (r_2, 0)\}),\end{aligned}$$

which proves (17.64).

Comparing Theorems 4 and 5, and, therefore, the functions $\eta : \mathbb{Z} \to \mathbb{R}$ and $a : \mathbb{Z} \to \mathbb{R}$, we see that the crucial difference is that in Theorem 4, one gets η in the form described in (17.48), where α is a concave, non-decreasing function on $\mathbb{Q}_+$, and there is no such corresponding restriction on the function a in Theorem 5. It is this feature of Theorem 4 which ensures that the literacy index $\Lambda : \mathbb{V} \to \mathbb{R}$ satisfies the monotonicity and equality axioms, and, consequently, the corresponding literacy index $L : \mathbb{U} \to \mathbb{R}$ satisfies axioms **N**, **SI**, **RM** and **PE**. This, in turn, implies that the literacy index L satisfies Theorem 1, and its principal application on the distribution of literacy rates (Theorem 2) holds. This important application, that a more equal distribution of literacy rates across households produces a higher overall literacy index for society, *cannot* be derived from Theorem 5.

We can make this point more precise by specifying a particular choice of the function $a : \mathbb{Z} \to \mathbb{R}$ as follows:

$$a(r, s) = \frac{sr^2}{r^2 + s^2} \quad \text{for all } (r, s) \in \mathbb{Z} \tag{17.71}$$

It is easy to check then that $a : \mathbb{Z} \to \mathbb{R}$ satisfies properties (a),(b) and (c) listed in Theorem 5. Thus, if we now define $\Lambda : \mathbb{V} \to \mathbb{R}$ by (17.35), where $n_i = r_i + s_i$ for all $i \in \{1, ..., m\}$ and $n = (n_1 + \cdots + n_m)$, then Λ is a literacy index which satisfies all of Dutta's axioms, namely, (17.33), (17.37), (17.62) and (17.64), by directly applying Theorem 5.

For a two-household society $\{(r_1, s_1), (r_2, s_2)\}$, we would then have (by (17.65)):

$$n\Lambda(\{(r_1, s_1), (r_2, s_2)\}) = \left[r_1 + \frac{s_1 r_1^2}{r_1^2 + s_1^2}\right] + \left[r_2 + \frac{s_2 r_2^2}{r_2^2 + s_2^2}\right]. \quad (17.72)$$

Consider now two societies specified as follows:

$$\left.\begin{array}{l} 1.\ \{(r_1, s_1), (r_2, s_2)\} = \{(1, 19), (3, 17)\} \\ 2.\ \{(r'_1, s'_1), (r'_2, s'_2)\} = \{(2, 18), (2, 18)\} \end{array}\right\} \quad (17.73)$$

Note that in the second society, there is complete equality in the literacy rates in the two households. In the first society, the literacy rates clearly differ among the households, with the second household having a higher literacy rate. Now, using the formula (17.72), one can easily verify that:

$$\Lambda(\{(r_1, s_1), (r_2, s_2)\}) = 4.56 \text{ and } \Lambda(\{(r'_1, s'_1), (r'_2, s'_2)\}) = 4.43$$

so that the society with the more *unequal* distribution of literacy rates has a higher overall literacy index.

Appendix: Mathematical Concepts and Results

Here, we provide an exposition of the mathematical concepts and results used in the second, third and fourth sections of the essay. We have tried to keep the exposition entirely elementary and self-contained. The reader who is familiar with the standard results on concave functions on the real line will be able to shorten the exposition considerably.

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.[9] 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

[9] See, for example, Green (1954), Roberts and Varberg (1973) and Rosenbaum (1950).

means that the function $f(z) = F(z, 1)$ is a convex function of z, for $z \in \mathbb{R}_+$. 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.[10]

Let $\mathbb{N}$ denote 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 the set of rational numbers, and $\mathbb{R}$ the set of real numbers. We denote 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:

$$f(z) = F(p, q)/q \quad \text{where} \quad z = (p/q), \text{ with } \quad (p, q) \in \mathbb{X} \qquad (17.74)$$

Note that the function, f, is well defined by (17.73). 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:[11]

Concavity (C): If $z, z' \in \mathbb{Y}$, and $t \in Y$, then $f(tz + (1 - t)z') \geq tf(z) + (1 - t)f(z')$.

---

[10] Properties of such functions, discussed later, 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.

[11] Here, as indicated in an earlier footnote, the terms '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.

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:

$$F(p, q) = qf(p/q) \quad \text{where} \quad (p, q) \in \mathbb{X} \tag{17.75}$$

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 (17.75), satisfies properties **H, SA, XM** and **OS.**

We summarise this discussion in the following proposition.

Proposition 1. *(a) 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 (17.74), satisfies properties* ***C, M*** *and* ***E****. (b) Suppose $f : \mathbb{Y} \to \mathbb{R}$ satisfies properties* ***C, M*** *and* ***E****. Then, the function, $F : \mathbb{X} \to \mathbb{R}$, defined by (17.75), satisfies properties* ***H, SA, XM*** *and* ***OS****.*

Majorisation 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 *majorisation* theory can be developed for it. The main result of this theory is known as the Tomic–Weyl theorem.[12] Something similar can be achieved for functions, $f : \mathbb{Y} \to \mathbb{R}$, which satisfy properties **C** and **M.** We summarise this theory in the following proposition.

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:*

$$\sum_{i=1}^{k} z'_i \leq \sum_{i=1}^{k} z_i. \tag{17.76}$$

Then, we must have:

$$\sum_{i=1}^{n} f(z'_i) \leq \sum_{i=1}^{n} f(z_i). \tag{17.77}$$

[12] See, for example, Mitrinovic-Vasic (1970: 165) for this result, as well as some of the important variations of it.

Remark: In Proposition 2, if there is a strict inequality in (17.76) when $k = n$, then we can infer that (17.77) must also hold with a strict inequality, by using property **M** of the function *f*.

Concave and Monotone Functions

In this subsection, we note two results relating to two classes of concave and monotone functions. These results are used to establish Theorem 4.

Proposition 3. *Suppose* $e : \mathbb{Y} \to \mathbb{R}_+$ *is a function such that (a)* $e(0) = e(1) = 0$, *(b)* $e(z) + z$ *satisfies property* **M** *on* $\mathbb{Y}$ *and (c) e satisfies property* **C** *on* $\mathbb{Y}$. *Then,* $\alpha : \mathbb{Q}_+ \to \mathbb{R}_+$, *defined by:*

$$\alpha(w) = (1 + w)e(w/(1 + w))$$

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

Proposition 4. *Suppose* α *is a function from* $\mathbb{Q}_+$ *to* $[0, 1)$, *such that: (a)* $\alpha(0) = 0$, *(b) if* $w, w' \in \mathbb{Q}_+$, *and* $w' > w$, *then* $\alpha(w') \geq \alpha(w)$ *and (c)* α *satisfies property* **C** *on* $\mathbb{Q}_+$. *If* $e : \mathbb{Y} \to \mathbb{R}_+$ *is a function defined as follows:*

$$e(z) = \begin{cases} \alpha(z/(1 - z))(1 - z) & \textit{for } z \neq 1 \\ 0 & \textit{for } z = 1 \end{cases}$$

then (a) $e(0) = e(1) = 0$, *(b)* $e(z) + z$ *satisfies property* **M** *on* $\mathbb{Y}$ *and (c) e satisfies property* **C** *on* $\mathbb{Y}$.

Proofs

In this section, the proofs of the mathematical results of the previous three subsections are presented.

Proof of Proposition 1. 1. Suppose $F : \mathbb{X} \to \mathbb{R}$ satisfies properties **H, SA, XM** and **OS**. Define the function, $f : \mathbb{Y} \to \mathbb{R}$ by (17.74). 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 $z, z' \in \mathbb{Y}$, 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:

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

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 previous computations.

2. 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 (17.75). 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,

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

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

Proposition 2 is established by using three lemmas, which we state and prove subsequently. The three lemmas are analogous to the basic results for concave functions defined on an interval of the real line.[13] 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:*

1. *If* $a, b, c \in \mathbb{Y}$, *and* $a < b \leq c$, *then* $[f(b) - f(a)]/[b - a] \geq [f(c) - f(a)]/[c - a]$.
2. *If* $a, b, c \in \mathbb{Y}$, *and* $a \leq b < c$, *then* $[f(c) - f(a)]/[c - a] \geq [f(c) - f(b)]/[c - b]$.
3. *If* $a, b, c \in \mathbb{Y}$, *and* $a < b < c$, *then* $[f(b) - f(a)]/[b - a] \geq [f(c) - f(a)]/[c - a] \geq [f(c) - f(b)]/[c - b]$.

Proof: We will only prove (1). The proof of (2) is similar; (3) follows directly from (1) and (2). 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:

$$\lambda = \frac{b - a}{c - a}. \tag{17.78}$$

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 rewritten as:

$$\lambda[f(c) - f(a)] \leq [f(b) - f(a)]. \tag{17.79}$$

Using (17.78) in (17.79), and dividing through by $(b - a) > 0$, we get:

$$\frac{[f(c) - f(a)]}{(c - a)} \leq \frac{[f(b) - f(a)]}{(b - a)}. \tag{17.80}$$

This establishes (1).

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:*

$$g(y) \equiv \lim_{\substack{\varepsilon \downarrow 0 \\ y+\varepsilon \in Y}} \{[f(y + \varepsilon) - f(y)]/\varepsilon\}. \tag{17.81}$$

Further, g is monotone non-increasing on Y.

---

[13] See, for example, Nikaido (1968: 47, Theorem 3.15) for comparison.

Proof: 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,

$$\frac{f(y) - f(y-\delta)}{\delta} \geq \frac{f(y+\varepsilon) - f(y)}{\varepsilon}. \tag{17.82}$$

Also, by Lemma 1, the right-hand side expression in (17.81) is non-decreasing as ε decreases to zero. Since it is bounded above by the left-hand side expression in (17.82), it must converge to a limit. Thus, $g : Y \to \mathbb{R}$, is well defined by:

$$g(y) \equiv \lim_{\substack{\varepsilon \downarrow 0 \\ y+\varepsilon \in Y}} \{[f(y+\varepsilon) - f(y)]/\varepsilon\}. \tag{17.83}$$

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,

$$\frac{f(y+(\delta/n)) - f(y)}{(\delta/n)} \geq \frac{f(z+(\delta/n)) - f(z)}{(\delta/n)}. \tag{17.84}$$

Letting $n \to \infty$ in (17.84), and using (17.83), we have $g(y) \geq g(z)$.

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:*

$$f(x) - f(y) \leq g(y)(x - y). \tag{17.85}$$

Proof: The result is trivial for $x = y$. So, we consider two cases (a) $x > y$ and (b) $x < y$. In case (*a*), 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:

$$\frac{f(x) - f(y)}{(x-y)} \leq \frac{f(y+(\delta/n)) - f(y)}{(\delta/n)} \tag{17.86}$$

by Lemma 1. Letting $n \to \infty$ in (17.86), and using Lemma 2, we have:

$$\frac{f(x) - f(y)}{(x-y)} \leq g(y) \tag{17.87}$$

Multiplying through in (17.87) by $(x - y) > 0$, we get the desired result.

In case (b), 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:

$$\frac{f(y) - f(x)}{(y - x)} \geq \frac{f(y) - f(y - (\delta/n))}{(\delta/n)} \geq \frac{f(y + (\delta/n)) - f(y)}{(\delta/n)} \tag{17.88}$$

by Lemma 1. Letting $n \to \infty$ in (17.90), and using Lemma 2, we get:

$$\frac{f(y) - f(x)}{(y - x)} \geq g(y). \tag{17.89}$$

Multiplying through in (17.89) by $(y - x) > 0$, we obtain:

$$f(y) - f(x) \geq g(y)(y - x). \tag{17.90}$$

Transposing terms in (17.90) yields the desired result.

The aforementioned lemmas, together with Abel's inequality, can be used to establish Proposition 2, which we now proceed to prove.

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 $i \in \{1, \ldots, n\}$,

$$f(z_i') - f(z_i) \leq g(z_i)(z_i' - z_i). \tag{17.91}$$

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

$$\sum_{i=1}^{n} [f(z_i') - f(z_i)] \leq \sum_{i=1}^{n} g(z_i)(z_i' - z_i). \tag{17.92}$$

Using Lemma 2, we have:

$$g(z_1) \geq g(z_2) \geq \cdots \geq g(z_n). \tag{17.93}$$

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

$$g(z_i) \geq 0 \textit{ for } i \in \{1, \ldots, n\}. \tag{17.94}$$

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

$$\sum_{i=1}^{n} g(z_i)(z_i' - z_i) \leq g(z_1)[\max_{1 \leq k \leq n} \sum_{i=1}^{k} (z_i' - z_i)]. \tag{17.95}$$

By (17.76), we have:

$$[\max_{1\leq k\leq n}\sum_{i=1}^{k}(z_i' - z_i)] \leq 0. \tag{17.96}$$

Thus, using (17.94) and (17.96) in (17.95), we have:

$$\sum_{i=1}^{n} g(z_i)(z_i' - z_i) \leq 0. \tag{17.97}$$

Using (17.97) in (17.92), we obtain (17.77), the desired result.

Turning now to the more general situation, we see that the following cases can arise: (*a*) $z_i \in Y$ for $i \in \{1, \ldots, n\}$; (*b*) $z_i = 0$ for $i \in \{1, \ldots, n\}$; (*c*) $z_i = 1$ for $i \in \{1, \ldots, n\}$; (*d*) there is $1 \leq p < n$ such that $z_i = 0$ for $i \in \{1, \ldots, p\}$ and $z_i = 1$ for $i \in \{p+1, \ldots, n\}$; (*e*) there is $1 \leq p < n$ such that $z_i = 0$ for $i \in \{1, \ldots, p\}$ and $z_i \in Y$ for $i \in \{p+1, \ldots, n\}$; (*f*) there is $1 \leq q < n$ such that $z_i \in Y$ for $i \in \{1, \ldots, q\}$ and $z_i = 1$ for $i \in \{p+1, \ldots, n\}$; (*g*) 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 (*a*). In case (*b*), (17.76) implies that $z_i' = 0$ for $i \in \{1, \ldots, n\}$, so (17.77) follows trivially. In case (*c*), $z_i' \leq z_i$ for $i \in \{1, \ldots, n\}$, so (17.76) follows from property **M** of *f*. In case (*d*), we have $z_i' = 0$ for $i \in \{1, \ldots, p\}$ by (17.76), while $z_i' \leq z_i$ for $i \in \{p+1, \ldots, n\}$, so (17.77) follows again from property **M** of *f*.

In case (*e*), we have $z_i' = 0$ for $i \in \{1, \ldots, p\}$ by (17.76), so we can define $j = i - p$ for $i = p+1, \ldots, n$, and $m = n - p$. Then, by (17.78), 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 (*a*), we obtain:

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

and this yields (17.77), since $z_i' = z_i = 0$ for $i = 1, \ldots, p$.

In case (*f*), using the analysis of case (*a*), we have:

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

and this yields (17.77), by using property **M** of f, since $z_i' \leq z_i = 1$ for $i = q+1, \ldots, n$.

In case (g), we have $z_i' = 0$ for $i \in \{1, \ldots, p\}$ by (17.76), so we can define $j = i - p$ for $i = p+1, \ldots, q$, and $m = q - p$. Then, by (17.76), 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 (a), we have:

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

This yields (17.77), by (A) using property **M** of f, and $z_i' \leq z_i = 1$ for $i = q+1, \ldots, n$, while noting that (A) $z_i' = z_i = 0$ for $i = 1, \ldots, p$.

We now present the proof of Proposition 3, which is used to establish part (1) of Theorem 4 in the text.

Proof of Proposition 3. Property (a) of the function α is clear from its definition and property (1) of the function e. We proceed to verify property (c) of α. 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:

$$(1 + w'') = t(1 + w) + (1 - t)(1 + w') \tag{17.98}$$

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

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

the inequality following from property (3) of e, after using (17.98).

We now turn to property (b) of the function, α. Let $w, w' \in \mathbb{Q}_+$, with $w' > w$. We claim that $\alpha(w') \geq \alpha(w)$. Suppose, on the contrary, that $\alpha(w') < \alpha(w)$. Define $\delta = [\alpha(w) - \alpha(w')]/(w' - w)$; then $\delta > 0$.

Using property (c) of α (which has already been established), and using the proof of Lemma 1, we have for all $n \in \mathbb{N}$,

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

This yields the inequality:

$$-\alpha(w) \leq \alpha(w' + n) - \alpha(w) \leq (w' + n - w)(-\delta) < -n\delta \qquad (17.99)$$

using the facts that $\alpha(w' + n) \geq 0$ and $(w' - w) > 0$. But (17.99) clearly leads to a contradiction for large n. This establishes our claim, and hence property (b) of α.

To verify that α maps from $\mathbb{Q}_+$ to $[0, 1)$, suppose on the contrary there is some $w \in \mathbb{Q}_+$, with $\alpha(w) \geq 1$. Define $z = w/(1 + w)$; then, we have $z \in (0, 1)$, and $(1 - z) = 1/(1 + w)$. Using the definition of α, we then have $e(z) = (1 - z)\alpha(w) \geq (1 - z)$, 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 (1) of e. This contradiction establishes the result.

We now present the proof of Proposition 4, which is used to establish part (2) of Theorem 4 in the text.

Proof of Proposition 4. Property (1) of e being clear from its definition, we proceed to prove property (2). Let $z, z' \in \mathbb{Y}$, with $z' > z$. There are two cases to consider: (I) $z' = 1$, (II) $z' < 1$. In case (I), $e(z') + z' = 1$, while $e(z) + z = \alpha(z/(1 - z))(1 - z) + z < (1 - z) + z$ (since $z < 1$, and $\alpha(z/(1 - z)) \in [0, 1)) = 1$.

In case (II), we have:

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

the first inequality following from property (b) of α and the fact that $z'/(1 - z') > z/(1 - z)$ (using $1 > z' > z$), and the second inequality following from the fact that $\alpha(z/(1 - z)) \in [0, 1)$ and $z' > z$.

We turn now to property (3) 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:

$$\Delta = t(1 - z) + (1 - t)(1 - z') \tag{17.100}$$

Then, we can write:

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

the inequality following from property **C** of the function, α, after using (17.100).

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, $z(n)$ and $z'(n)$ 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:

$$\begin{aligned} e(tz + (1-t)z') + [tz + (1-t)z'] &\geq e(tz(n) + (1-t)z'(n)) \\ &\quad + [tz(n) + (1-t)z'(n)] \\ &\geq te(z(n)) + (1-t)e(z'(n)) \\ &\quad + [tz(n) + (1-t)z'(n)], \end{aligned} \tag{17.101}$$

the first inequality following from property (2) of e (already established earlier), and the second inequality from the earlier analysis of property (3) of e, since $z(n), z'(n)$ are both less than 1. The inequality (17.101) yields:

$$\begin{aligned} e(tz + (1-t)z') &\geq te(z(n)) + (1-t)e(z'(n)) \\ &\quad + t(z(n) - z) + (1-t)(z'(n) - z') \\ &\geq te(z) + (1-t)e(z') \\ &\quad + t(z(n) - z) + (1-t)(z'(n) - z'), \end{aligned} \tag{17.102}$$

the second inequality following from the fact that $e(1) = 0$, while e maps to $\mathbb{R}_+$. Now, letting $n \to \infty$ in (17.102), we obtain the desired result.

References

Basu, K. and J. E. Foster (1998), 'On Measuring Literacy', *Economic Journal*, 108(451): 1733–49.

Basu, K., A. Narayan and M. Ravallion (2002), 'Is Literacy Shared within Households? Theory and Evidence for Bangladesh', *Labour Economics* 8(6): 649–65.

Dutta, I. (2004), 'Generalized Measures of Literacy', *Mathematical Social Sciences*, 48, 69–80.

Gibson, J. (2001), 'Literacy and Intrahousehold Externalities', *World Development* 29, 155–66.

Green, J.W. (1954), 'Recent Applications of Convex Functions', *American Mathematical Monthly*, 61(7): 449–54.

Green, S. E., T.A. Rich and E.G. Nesman (1985), 'Beyond Individual Literacy: the Role of Shared Literacy for Innovation in Guatemala', *Human Organisation* 44(4): 313–21.

Mitra, T. (2002), 'On Literacy Rankings', CAE Working Paper 02-16, Cornell University.

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

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

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

Rosenbaum, R.A. (1950), "Sub-Additive Functions", *Duke Mathematical Journal*, 17: 227–47.

Subramanian, S. (2004), 'Measuring Literacy: Some Extensions of the Basu–Foster Framework', *Journal of Development Economics*, 73(1): 453–63.

——— (2008), 'Externality and Literacy: A Note', *Journal of Development Studies*, 44(6): 839–48.

Valenti, P.M. (2002), "Should We Be Concerned about the Distribution of Literacy across Households? An Axiomatic Characterisation", CAE Working paper 02-15, Cornell University.

Page Intentionally Left Blank