CAE Working Paper #07-13

Benchmarking the Effective Literacy Rate

by

Travis Lee

July 2007

.

BENCHMARKING THE EFFECTIVE LITERACY RATE

TRAVIS LEE

ABSTRACT. By now the importance of literacy in the process of development is widely accepted. However, unlike measuring welfare or inequality, the problem of literacy measurement remains largely unexamined. Alternatives to the standard literacy rate, R, equal to the number of literate adults as a percentage of the adult population, are not well known, but this measure has its deficiencies. In particular, several authors have identified the externality accruing to proximate illiterates, that is, illiterate people with access to a literate person. The standard literacy rate ignores this externality; measures of *effective* literacy are sensitive to it. The present note offers two effective literacy rates (measures) and a set of axioms characterizing each. Nearly all measures of effective literacy appearing in the literature are greater than or equal to R. In fact, the best known of these, the Basu-Foster measure $\mathcal{L}^*$, is strictly greater in virtually every case (see Basu and Foster [1]). Although the inequality $\mathcal{L}^* \geq R$ is an unintended consequence of their construction, it amounts to setting a benchmark for the effective literacy rate. This notes examines Basu and Foster's framework and offers an alternative benchmark.

1. INTRODUCTION

The standard literacy rate, R, equal to the number of literate adults as a percentage of the adult population, is the only widely-used measure of literacy, in spite of the fact that it ignores significant aspects of literacy. In particular, several authors[1] have identified an externality accruing to the illiterate members of households which include at least one literate member, so called proximate illiterate people. Empirical evidence suggests this externality is large (see John Gibson [5]).

Measures of literacy which are sensitive to the intrahousehold externality are termed measures of *effective* literacy. There is a growing literature on the construction and characterization of such measures, beginning with Basu and Foster [1]. Their measure $\mathcal{L}^*$ is $R + \alpha P$, where P is the proximate illiteracy rate (percentage of proximate illiterate adults in the adult population) and α is a real number strictly between zero and one. From this formulation, it is clear that $\mathcal{L}^*$ is greater than or equal to R. In fact $\mathcal{L}^* > R$ for any society which includes a proximate illiterate person. This result is a consequence, not an assumption, of their axiomatic approach (more on this observation below), and yet adopting $\mathcal{L}^*$ to measure literacy amounts to setting a benchmark for the effective literacy rate, namely $\mathcal{L}^* \geq R$.

In addition, it is arguable that adopting $\mathcal{L}^*$, in place of the traditional R, raises practical concerns. Since $\mathcal{L}^*$ is greater than R, adopting $\mathcal{L}^*$ results in an immediate

[1]See, for example, Basu and Foster [1]; John Gibson [5]; Green, Rich, and Nesman [6]; and Bryan Maddox [8].

increase in the perceived literacy rate and a consequent sense of improvement, all without any real change taking place.

Other authors have extended Basu and Foster's ideas. Indranil Dutta [3], S. Subramanian [11], and Paola Valenti [12]) have given measures for which the extent of the externality depends on the rate of literacy within each household. In almost every case, the proposed measure of literacy is bounded below by the standard literacy rate. Additional authors have offered further extensions (see Basu, Narayan, and Ravallion [2] and Tapan Mitra [10]), but the problem of benchmarking remains unresolved.

For these reasons, there has been interest in obtaining new measures of effective literacy, $\mathcal{L}$, which are sensitive to the intrahousehold externality but which satisfy the opposite benchmark, that is $\mathcal{L} \leq R$. This note offers two such measures and a set of axioms characterizing each. These measures are $\mathcal{L}_1 = R - \alpha I$ for $\alpha > 0$ and $\mathcal{L}_2 = (1 - \alpha I)R$ for $0 < \alpha \leq 1$, where I is the percentage of isolated illiterate adults (illiterate adults who are not proximate).

S. Subramanian has already derived the measure of literacy $(1 - I)R$ from different motivations [11]. He provides an interesting characterization of $(1 - I)R$ as a function of R and a measure of efficiency loss $I/(1-R)$ satisfying certain properties. An alternate axiomatic characterization is provided here, and the presence of α in $\mathcal{L}_2 = (1 - \alpha I)R$ may permit empirical comparisons.

Section 2 introduces the Basu-Foster axiomatic framework and their measure $\mathcal{L}^*$. This section serves as a starting point for exploring the consistency of their framework with other possible benchmarks. Sections 3 and 4 concern the new measures $\mathcal{L}_1$ and $\mathcal{L}_2$ respectively. The main result, theorem 1, characterizes $\mathcal{L}_2$.

2. Notation and Axioms

Borrowing from [1], denote by Δ the set of all societies, where $x \in \Delta$ is a (finite-length) vector of households $x = (x^1, \ldots, x^m)$. In turn, each x^h is a (finite-length) vector of ones and zeroes, denoting literate and illiterate members of household h respectively. An illiterate person (denoted by a zero in some x^h) is *proximate illiterate* if his or her household x^h includes at least one literate member. Otherwise he or she is *isolated illiterate*; that is, an illiterate person is isolated illiterate if his or her household includes no literate members. Thus a household $x^1 = (1, 0, 0)$ includes one literate and two proximate illiterate members, while the household $x^2 = (0, 0, 0)$ consists of three isolated illiterate people. A household will be called *literate* if it includes a literate member and *illiterate* otherwise.

Throughout r_x, p_x, and i_x denote the number of literate, proximate illiterate, and isolated illiterate people in a society x. By the same token, R, P, and I denote the relative frequencies of these groups, that is $R(x) = r_x/n_x$, $P(x) = p_x/n_x$, and $I(x) = i_x/n_x$, where the symbol n_x denotes the total number of people in x. The number of households in x is denoted m_x, and the size of household h is n_h. Although the notation n_1 may refer to either the population of *society* 1 or the number of members of *household* 1, the context will make clear which meaning is intended. In fact the subscripts x and h in n_x, m_x, n_h, r_x, p_x, and i_x will be dropped when there is no ambiguity.

Basu and Foster give five axioms which characterize their measure of literacy $\mathcal{L}^*$. Quoting from [1], their axioms are as follows:

AXIOM A (Anonymity): If $x \in \Delta$ is obtained from $y \in \Delta$ by either a permutation of households or a permutation of individuals within a household, then $\mathcal{L}(x) = \mathcal{L}(y)$.

Society x is obtained from society y by a *simple increment* if for some (h, j) $x_j^h = 1$ and $y_j^h = 0$, while $x_{j'}^{h'} = y_{j'}^{h'}$ for all $(h', j') \neq (h, j)$.

AXIOM M (Monotonicity): If $x \in \Delta$ is obtained from $y \in \Delta$ by a simple increment, then $\mathcal{L}(x) > \mathcal{L}(y)$.

Society x is obtained from society y by a *household split* if y has m households, x has $m + 1$ households, y^m is the concatenation of x^m and x^{m+1}, and $x^h = y^h$ for all other households $h = 1, \ldots, m - 1$. The split is *externality-neutral* if either (i) both x^m and x^{m+1} contain a literate person or (ii) neither of x^m or x^{m+1} contain a literate person. It is called *externality-reducing* if exactly one of x^m or x^{m+1} contains a literate person.

AXIOM E (Externality): Suppose $x \in \Delta$ is obtained from $y \in \Delta$ by household split. If the split is externality-neutral, then $\mathcal{L}(x) = \mathcal{L}(y)$; and if the split is externality-reducing, then $\mathcal{L}(x) < \mathcal{L}(y)$.

A society x is *completely literate* if $x_j^h = 1$ for all h and j; x is *completely illiterate* if $x_j^h = 0$ for all h and j.

AXIOM N (Normalization): If $x \in \Delta$ is completely literate, then $\mathcal{L}(x) = 1$; if $x \in \Delta$ is completely illiterate, then $\mathcal{L}(x) = 0$.

Society x is *decomposed* into societies y and z if $y^h = x^h$ for all $h = 1, \ldots, m_y$ and $z^h = x^{h+m_y}$ for all $h = 1, \ldots, m_z$, where $m_y + m_z = m_x$.

AXIOM D (Decomposition): Suppose $x \in \Delta$ is decomposed into $y \in \Delta$ and $z \in \Delta$. Then $\mathcal{L}(x) = (n_y/n_x)\mathcal{L}(y) + (n_z/n_x)\mathcal{L}(z)$.

Any measure of literacy satisfying these five axioms is $\mathcal{L}^* = R + \alpha P$ for some $0 < \alpha < 1$. Obviously this measure is sensitive to the "the intrahousehold externality arising from the presence of a literate member" [1] through the parameter α. Essentially literate people contribute to this measure a value of one over the size of the population (as with the standard measure R) and, in addition, proximate illiterate people contribute α over the size of the population.

The inequality $\mathcal{L}^* \geq R$ is a consequence of axioms A, D, E, and N as follows: if x is a single household consisting of $r \geq 0$ literate and $i \geq 0$ illiterate members, then for any measure of literacy $\mathcal{L}$ satisfying these axioms

$$(2.1) \qquad \mathcal{L}(x) \geq \mathcal{L}((1, \ldots, 1), (0, \ldots, 0)) = \frac{r}{n}\mathcal{L}(1, \ldots, 1) + \frac{i}{n}\mathcal{L}(0, \ldots, 0)$$

$$= \frac{r}{n} \cdot 1 + \frac{i}{n} \cdot 0 = \frac{r}{n} = R(x),$$

where the inequality holds by axioms A and E, the first equality by axiom D, and the second by axiom N. Since both $\mathcal{L}$ and R satisfy axiom D, it follows that $\mathcal{L}(x) \geq R(x)$ for all societies x.

Thus the measure $\mathcal{L}^*$ implicitly assumes a benchmark for the effective literacy rate, namely $\mathcal{L}^* \geq R$; however there are other options. The two measures given below satisfy the polar opposite benchmark, that is $\mathcal{L} \leq R$. It is possible to imagine

additional measures of literacy which exceed R on some societies and fall below it on others. One could even devise a measure which is calibrated to match a region's standard literacy rate at a given point in time, with the complication that all future literacy comparisons between regions must account for this calibration.[2]

There are some practical reasons for concern that $\mathcal{L}^* \geq R$. Although switching to a new measure causes no change in literacy, reporting a new, higher "rate of literacy" creates a perception of improvement in literacy attainment. Hence reporting the new measure may ease pressure on policy makers to affect change. Examples of the misuse of statistics, particularly in questions of public policy, are plentiful; see, for instance, Mark Maier [9] and Jaffe and Spirer [7].

Since the externality axiom is essential to capturing the externality accruing to proximate illiterate members of literate households, and the anonymity axiom seems the least controversial, one is forced to consider measures $\mathcal{L}$ violating either the decomposition or normalization axioms in order to achieve the new benchmark $\mathcal{L} \leq R$.

3. Nixing Normalization

As observed in the previous section, $\mathcal{L}^*(x) \geq R(x)$ for all $x \in \Delta$. One might ask what conditions are necessary to achieve the new benchmark $\mathcal{L}(x) \leq R(x)$. The previous section explained why it is necessary to dispense with either axiom D or N. This section considers measures of literacy which violate the normalization axiom. What other restrictions, if any, are imposed by the new benchmark?

By axioms E and D respectively, $\mathcal{L}(0, \ldots, 0) = \mathcal{L}((0), \ldots, (0)) = \frac{1}{n} \cdot n \cdot \mathcal{L}(0) = \mathcal{L}(0)$, where n is the number of members of the one-household society $(0, \ldots, 0)$. The same argument shows $\mathcal{L}(1, \ldots, 1) = \mathcal{L}(1)$. Then given any one-household society x with r literate and p illiterate members, axioms A and E, axiom D, and the preceding computations respectively imply

$$\mathcal{L}(x) \geq \mathcal{L}((1, \ldots, 1), (0, \ldots, 0)) = \frac{r}{n}\mathcal{L}(1, \ldots, 1) + \frac{p}{n}\mathcal{L}(0, \ldots, 0) = \frac{r}{n}\mathcal{L}(1) + \frac{p}{n}\mathcal{L}(0)$$

If $\mathcal{L}$ satisfies axiom N, then $\mathcal{L}(x) \geq \frac{r}{n} = R(x)$. Again we see that it is necessary to do without this axiom. Notice if $\mathcal{L}(1) = 1$, then $\mathcal{L}(x) \geq R(x) + \frac{p}{n}\mathcal{L}(0)$ and $\mathcal{L}(0) < 0$ is a necessary condition for $\mathcal{L}$ to satisfy the new benchmark.[3] Now we are ready to formulate the new benchmark in the following axiom.

Axiom B (Benchmark): If $x \in \Delta$ has no isolated illiterate members, then $\mathcal{L}(x) = R(x)$.[4]

If $\mathcal{L}$ satisfies both axioms B and E, it is immediate that $\mathcal{L}(x) < R(x)$ when x includes at least one isolated illiterate member. Notice $\mathcal{L}^*(x) \geq R(x)$ for all $x \in \Delta$, and $\mathcal{L}^*(x) = R(x)$ exactly when society x includes no proximate illiterate members. The new benchmark has been defined in a precisely analogous way. This particular

---

[2]James Foster originated this idea for a solution.

[3]In fact there are various ways to go about weakening axiom N. The inequality $\mathcal{L} \leq R$ and axioms A, D, E, and N together imply $\mathcal{L}(x) = f_{\mathcal{L}}(R(x))$ where $f : [0,1] \to (-\infty, 1]$ is strictly increasing and $f_{\mathcal{L}}(\rho) \leq \rho$ for all $\rho \in [0,1]$. The main text considers the case where, in addition, $f_{\mathcal{L}}(1) = 1$.

[4]Technically it is unnecessary to explicitly mention the standard literacy rate since R is equivalent to axioms A, D, I, and N, where axiom I requires a measure to be indifferent to the externality enjoyed by the proximate illiterate. However the axiom seems more transparent in its current form.

formulation has the consequence that $\mathcal{L}(x) < 0$ for any completely illiterate society x. However as shown above, this restriction is already implied by axioms A, D, E, and $\mathcal{L}(1) = 1$. Another version of the benchmark axiom is given in the following section.

As the next proposition shows, axioms A, B, D, and E fully characterize a new measure of literacy $\mathcal{L}_1 = R - \alpha I$, where $\alpha > 0$.

Proposition 1. *A measure of literacy $\mathcal{L} : \Delta \to \mathbb{R}$ satisfies axioms A, B, D, and E if and only if $\mathcal{L} = \mathcal{L}_1$ for some $\alpha > 0$.*

Proof. $\mathcal{L}_1$ is simply an affine transformation of $\mathcal{L}^*$. To see this, set $\alpha' = \alpha/(1+\alpha)$ and make the following computation:

$$\mathcal{L}_1 + \alpha = R - \alpha I + \alpha = R - \alpha(1 - R - P) + \alpha = (1+\alpha)R + \alpha P.$$

Hence $(\mathcal{L}_1 + \alpha)/(1+\alpha) = \mathcal{L}^*_{\alpha'}$, where $\mathcal{L}^*_{\alpha'} = R + \alpha' P$ is the Basu-Foster measure of literacy $\mathcal{L}^*$ with parameter α'. So $\mathcal{L}_1$ satisfies axioms A, D, and E. (By the same reasoning $\mathcal{L}_1$ satisfies axiom M.) It is clear that $\mathcal{L}_1$ satisfies axiom B since $\alpha > 0$.

Conversely suppose $\mathcal{L} : \Delta \to \mathbb{R}$ satisfies axioms A, B, D, and E. Let x be any society, written $x = (x^1, \ldots, x^l, x^{l+1}, \ldots, x^h)$ where household x^k is literate for $k = 1, \ldots, l$, and illiterate for $k = l+1, \ldots, h$. By axiom A, this rearranging has no effect on $\mathcal{L}(x)$. As shown above, axioms A, B, D, and E imply there is some $\alpha > 0$ such that $\mathcal{L}(0) = -\alpha$.

Then axioms A, D, and B plus $\mathcal{L}(0) = -\alpha$ respectively imply

$$\mathcal{L}(x) = \mathcal{L}(x^1, \ldots, x^l, x^{l+1}, \ldots, x^h) = \sum_{k=1}^{l} \frac{n_k}{n_x} \mathcal{L}(x^k) + \sum_{k=l+1}^{h} \frac{n_k}{n_x} \mathcal{L}(x^k)$$

$$= \sum_{k=1}^{l} \frac{n_k}{n_x} R(x^k) - \alpha \sum_{k=l+1}^{h} \frac{n_k}{n_x} = \frac{r_x}{n_x} - \alpha \frac{i_x}{n_x} = R(x) - \alpha I(x).$$

Therefore $\mathcal{L} = \mathcal{L}_1$ with parameter $\alpha > 0$. □

As mentioned in the course of the preceding proof, $\mathcal{L}_1$ also satisfies axiom M. Thus it meets all the Basu-Foster axioms, except N, and fulfills the new benchmark $\mathcal{L}_1 \leq R$.

4. Dropping Decomposition

A legitimate criticism of the measure $\mathcal{L}_1$ is that it can take negative values, in violation of axiom N (more precisely axioms M and N imply $0 \leq \mathcal{L}(x) \leq 1$ for all $x \in \Delta$). Equation 2.1 exhibits the impossibility of finding a measure of literacy meeting the new benchmark and satisfying axioms A, D, E, and N. To find a measure respecting the intrahousehold externality, it is essential to keep axiom E. In addition axiom A is arguably the most reasonable and least controversial requirement. As the previous section concerned measures violating normalization, this section considers measures violating decomposition. The following is a reformulation of the benchmark axiom which is compatible with axiom N.

Axiom Q (Qualified Benchmark): If $x \in \Delta$ has no isolated illiterate members or x is completely illiterate, then $\mathcal{L}(x) = R(x)$.

One might ask whether there is a measure of literacy satisfying axioms A, E, N, and Q, and, if so, whether it is uniquely determined. As shown below, $\mathcal{L}_2 =$

$(1 - \alpha I)R$ meets these conditions when $0 \leq \alpha < 1$. So does $\mathcal{L}' = (1 - 1.5I)R$, although it violates axiom M. For $0 \leq \alpha < 1$, however, $\mathcal{L}'' = (1 - \alpha I^2)R$ satisfies axioms A, E, N, and Q, and monotonicity as well.[5] As these examples illustrate, it is necessary to impose additional constraints in order to uniquely characterize $\mathcal{L}_2$.

To set the stage, let $\mathcal{L}$ be any measure of literacy satisfying axioms A and E, and let x be any society. Then $\mathcal{L}(x) = \mathcal{L}(\tilde{x}^1, \tilde{x}^2)$ where the households $\tilde{x}^1$ and $\tilde{x}^2$ are the concatenations of all the literate and illiterate households in x respectively. Then $\tilde{x}^1$ consists of r literate and p proximate illiterate members, while $\tilde{x}^2$ is a household of i illiterate people. In other words,

$$\mathcal{L}(x) = \mathcal{L}((1, \ldots, 1, 0, \ldots, 0), (0, \ldots, 0)).$$

As this equation makes clear, $\mathcal{L}(x)$ is determined by r, p, and i. (Of course population size is $n = r + p + i$.) With this fact in mind, we will write $\mathcal{L}(x) = f_{\mathcal{L}}(r, p, i)$, where $f_{\mathcal{L}}$ is simply the function of these three variables which determines $\mathcal{L}$. More formally, denote by $\mathbb{Z}_+ = \{0, 1, 2, \ldots\}$ the set of nonnegative integers and by $\overline{\Delta} = \{(r, p, i) : r, p, i \in \mathbb{Z}_+\}$ the set of societies written as ordered triples. Then $f_{\mathcal{L}}$ is a particular function from $\overline{\Delta}$ to $\mathbb{R}$.

One way to impose an additional constraint on a measure $\mathcal{L}$ satisfying axioms A and E is to require a certain relationship between $f_{\mathcal{L}}(r, p, i)$ and $f_{\mathcal{L}}(r, p + 1, i - 1)$, provided $i \geq 1$. On this note, consider the following axiom.

AXIOM L (Linearity): If $x \in \Delta$, then $\mathcal{L}(x)$ is uniquely determined by r, p, and i. In other words $\mathcal{L}(x) = f_{\mathcal{L}}(r, p, i)$ for some function $f_{\mathcal{L}} : \overline{\Delta} \to \mathbb{R}$. Furthermore, whenever $r \geq 1$ and $i \geq 2$,

$$f_{\mathcal{L}}(r, p + 2, i - 2) - f_{\mathcal{L}}(r, p + 1, i - 1) = f_{\mathcal{L}}(r, p + 1, i - 1) - f_{\mathcal{L}}(r, p, i) > 0.$$

People who are isolated illiterate in a society x reduce $\mathcal{L}(x)$ for any measure of literacy $\mathcal{L}$ satisfying axiom E. In other words, moving an isolated illiterate person into a literate household produces a gain in the literacy measure. Intuitively the new axiom says the same increase in the effective literacy rate is obtained from moving the first, second, third, and so on, isolated illiterate person into a literate household.

The axiom does not say the increase in effective literacy obtained by moving one isolated illiterate person into a literate household is the same for all societies. In other words, if $x, y \in \Delta$ with $r_x, r_y \geq 1$ and $i_x, i_y \geq 2$, it may be that

$$f_{\mathcal{L}}(r_x, p_x + 1, i_x - 1) - f_{\mathcal{L}}(r_x, p_x, i_x) \neq f_{\mathcal{L}}(r_y, p_y + 1, i_y - 1) - f_{\mathcal{L}}(r_y, p_y, i_y).$$

In fact if $\mathcal{L}$ also satisfies axiom Q, it is easy to see that

$$f_{\mathcal{L}}(r_x, p_x + 1, i_x - 1) - f_{\mathcal{L}}(r_x, p_x, i_x) = (1/i_x)(R(x) - \mathcal{L}(x))$$

so the gain depends on the number of isolated illiterate people in a given society.

As shown above, $\mathcal{L}(x) = f_{\mathcal{L}}(r, p, i)$ is already implied by axioms A and E. Thus for any measure satisfying these two axioms, only the second part of axiom L,

---

[5]This example illustrates an additional point. The literacy measures considered here, as well as many of those considered elsewhere, are polynomial functions in the society-wide per capita rates of literacy, proximate illiteracy, and isolated illiteracy. Thus many of the supporting axioms can be reformulated as conditions on these polynomials. Exceptions include works by I. Dutta [3], S. Subramanian [11], and P. Valenti [12], all of whom consider measures in which the size of the externality accruing to a proximate illiterate person depends on the distribution of literacy within his or her household.

concerning linearity, is new. Conversely a little thought shows axiom L implies the anonymity and externality axioms. The stipulation $r \geq 1$ permits $\mathcal{L}(x) = 0$ when x is completely illiterate.

Axiom L strengthens axiom E by requiring the effective literacy rate to increase linearly as any given society x (with $r \geq 1$ and $i \geq 2$) moves to a society resembling x except that isolated illiteracy has been eliminated. Axiom L is quite strong, but it is meant to formalize a strong idea, namely that the externality accruing to proximate illiteracy be measured in a linear fashion.[6]

Theorem 1. *A measure of literacy $\mathcal{L} : \Delta \to \mathbb{R}$ satisfies axioms A, L, N, and Q if and only if $\mathcal{L} = \mathcal{L}_2$ for some $0 < \alpha \leq 1$.*

Proof. Since $\mathcal{L}_2 = (1 - \alpha I)R$, it is easy to check that it satisfies axioms A, N, and Q. ($\mathcal{L}_2$ also satisfies the monotonicity axiom.) In fact

$$\mathcal{L}_2(x) = f_{\mathcal{L}}(r, p, i) = \left(1 - \alpha \frac{i}{n}\right)\frac{r}{n},$$

where n is an abbreviation for $r + p + i$. Further if $r \geq 1$ and $0 < k \leq i$, then

$$f_{\mathcal{L}}(r, p + k, i - k) - f_{\mathcal{L}}(r, p + k - 1, i - k + 1) = \frac{\alpha r}{n}$$

so $\mathcal{L}_2$ satisfies axiom L.

Conversely suppose $\mathcal{L}$ satisfies these axioms. By axiom N, $\mathcal{L}(x) = 0$ for any completely illiterate society x. Now suppose $x \in \Delta$ includes at least one literate member. Let $\hat{x}^2$ be any literate household in x, and let $\hat{x}^1$ be the concatenation of all the other literate households. Lastly let $\hat{x}^3 = (0_{i_x})$ be the concatenation of all the illiterate households where 0_k denotes a finite sequence of k illiterate individuals. Then by axioms A and L, there is an $\alpha > 0$ such that

$$\mathcal{L}(x) = \mathcal{L}(\hat{x}^1, \hat{x}^2, \hat{x}^3) = \mathcal{L}(\hat{x}^1, (\hat{x}^2, 0_1), (0_{i_x - 1})) - \frac{\alpha}{n}R(x),$$

where $(\hat{x}^2, 0_k)$ is a single household which is the concatenation of $\hat{x}^2$ and k illiterate individuals. By the linearity axiom, $\mathcal{L}(x) = \mathcal{L}(\hat{x}^1, (\hat{x}^2, 0_k), (0_{i_x - k})) - \frac{\alpha k}{n}R(x)$ for any $k \leq i_x$, and in particular

$$\mathcal{L}(x) = \mathcal{L}(\hat{x}^1, (\hat{x}^2, 0_{i_x})) - \frac{\alpha i_x}{n}R(x) = R(x) - \alpha I(x)R(x),$$

where the last equality follows by axiom Q. Axiom N guarantees $\alpha \leq 1$. Thus $\mathcal{L} = \mathcal{L}_2$ for some $0 < \alpha \leq 1$. $\qquad\square$

Axiom D has proved extremely useful in practice. Although $\mathcal{L}_2$ does not satisfy axiom D, it is possible to make similar computations. In fact if x is the concatenation of y and z, then

$$\mathcal{L}_2(x) = \frac{n_y(1 - \alpha I(x))}{n_x(1 - \alpha I(y))}\mathcal{L}_2(y) + \frac{n_z(1 - \alpha I(x))}{n_x(1 - \alpha I(z))}\mathcal{L}_2(z).$$

---

[6]Writing down the precise constraints imposed by this idea requires some attention to fine detail, but a little thought makes clear that it imposes numerous restrictions, such as, for example, $\mathcal{L}((1,0),(1,0)) = \mathcal{L}((1,0,0),(1))$.

This formula[7] gives a decomposition of the literacy rate as a weighted sum where, admittedly, the weights are more complicated than in axiom D. Note that to compute $\mathcal{L}_2$ for some region, it is necessary to know both $\mathcal{L}_2$, R, and population size for each subregion. Thus $\mathcal{L}_2$ satisfies one particular decomposability axiom, of which there are many. In [4], James Foster and Amartya Sen give a detailed survey of decomposability in the context of measuring inequality.

As mentioned in the introduction, S. Subramanian [11] gives an interesting characterization of $R(1-I)$ as a function of R and a an efficiency loss indicator $I/(1-R)$. Similarly the measure $\mathcal{L}_2$ is a function of R and a slightly more general indicator of efficiency loss $Q' = \alpha I/(1-R)$, for $0 < \alpha \leq 1$. This generalization is hardly worth mentioning, except that the presence of $0 < \alpha \leq 1$ may provide the opportunity for an empirical comparison of $\mathcal{L}^*$ and $\mathcal{L}_2$ (see, for example, Gibson [5]).

5. Summary

This note offers two measures of literacy which account for the positive externality arising from proximate illiteracy. Both are less than or equal to the standard measure of literacy R. Furthermore each measure is characterized by a set of axioms.

References

1. Kaushik Basu and James Foster, *On measuring literacy*, Economic Journal **108** (1998), no. 451, 1733–1749.
2. Kaushik Basu, Ambar Narayan, and Martin Ravallion, *Is literacy shared within households? Theory and evidence from Bangladesh*, MIT Department of Economics Working Paper 01-44, October 2001.
3. Indranil Dutta, *Generalized measures of literacy*, Mathematical Social Sciences **48** (2004), 69–80.
4. James Foster and Amartya Sen, *On economic inequality after a quarter century*, Clarendon Press, Oxford, England, 1997, in *On economic inequality* by Amartya Sen, enlarged edition with a substantial annexe *On economic inequality after a quarter century*.
5. John Gibson, *Literacy and intrahousehold externalities*, World Development **29** (2001), 155–166.
6. S.E. Green, , T.A. Rich, and E.G. Nesman, *Beyond individual literacy: the role of shared literacy for innovation in Guatemala*, Human Organization **44** (1985), no. 4, 313–321.
7. A. J. Jaffe and Herbert F. Spirer, *Misused statistics: Straight talk for twisted numbers*, Popular Statistics series, no. 5, Dekker, New York, 1987.
8. Bryan Maddox, *Worlds apart? Ethnographic reflections on "effective literacy" and intrahousehold externalities*, World Development **35** (2007), no. 3, 532–541.
9. Mark H. Maier, *The data game: Controversies in social science statistics*, second ed., Sharpe, Armonk, NY, 1995.
10. Tapan Mitra, *On literacy rankings*, Center for Analytical Economics Working Paper 02-16, November 2002.
11. S. Subramanian, *Measuring literacy: Some extensions of the Basu-Foster framework*, Journal of Development Economics **73** (2004), 453–463.
12. Paola M. Valenti, *Should we be concerned about the distribution of literacy across households? An axiomatic investigation*, Center for Analytical Economics Working Paper 02-15, October 2002.

Department of Economics, Cornell University, Ithaca, New York, USA
E-mail address: JTL29@Cornell.edu

---

[7] Given $\mathcal{L}_2(x)$ and $R(x)$ for any society x, it is always possible to compute $I(x)$ via the formula $I(x) = (1/\alpha)(1 - \mathcal{L}_2(x)/R(x))$.