Working Paper No. 154

Assessing the `Femaleness' of a population

by

D. Jayaraj
S. Subramanian

Madras Institute of Development Studies
79, Second Main Road, Gandhi Nagar
Adyar, Chennai 600 020

July 1998

TITLE OF PAPER: ASSESSING THE 'FEMALENESS' OF A POPULATION

AUTHORS' NAMES
AND INSTITUTIONAL
AFFILIATION: D. Jayaraj and S. Subramanian
Madras Institute of Development Studies
79, Second Main Road
Gandhinagar, Adyar
CHENNAI - 600 020
INDIA

ABSTRACT OF PAPER: Conventional measures of the 'femaleness' of a population -- measures such as the female headcount ratio F and the sex- ratio S -- are insensitive to the precise age-specific distribution of sex-ratios in the population. This paper attempts to correct for this neglect of variability in the age-wise distribution of sex-ratios, by advancing variants of F and S, called F^* and S^* respectively, which go some way in mitigating the misleadingness of wholly 'mean-centred' approaches to the measurement of 'femaleness' in a population.

ACKNOWLEDGEMENTS: The authors are indebted to R. Dharumaperumal for his assistance with the word-processing of the paper, and for seeing them through the graphics.

ASSESSING THE 'FEMALENESS' OF A POPULATION

by

D. Jayaraj and S. Subramanian

1. MOTIVATION

A key demographic indicator of the status of women in a society, which can also be regarded as reflecting a crucial socio-economic aspect of human development, is constituted by the weight of women in the society's population. There is, clearly, more than one way in which the 'weight of women in a society's population' can be reckoned. This note is concerned, in a preliminary and elementary way, with advancing one particular method for assessing the degree of 'femaleness' of a population. This method, we contend, is sensitive to the age-specific distribution of 'femaleness' in a way which conventional measures fail to reflect.

Specifically, it is reasonable to attach a greater weight to any given level of 'female intensity' of a population the greater is the age at which it obtains. It is possible for two societies to share the same overall sex-ratio and yet display vastly differing profiles of the age-related distribution of sex-ratios, and a measure of 'femaleness' which fails to discriminate between these two demographic regimes is quite misleading. Concretely, a typical feature of the pattern of age-specific sex composition in a 'demographically developed' regime is that the sex-ratio at birth is less than unity but is compensated for to such an extent at later ages that the overall sex-ratio exceeds unity. In principle, one could have a regime which inverts this pattern of age-specific sex composition and yet displays the same overall sex-ratio. Failure to detect the crucial difference between the two regimes must be seen to be a serious failure. Briefly, standard measures of the 'femaleness' of a population tend to gloss over the specific pattern of age-wise distribution of sex-ratios in a population. This paper is concerned to offer a corrective to this inadequacy.

The paper is organized as follows. Section 2 briefly reviews two standard measures of the 'femaleness' of a population: the female headcount ratio F and the sex-ratio S. Section 3 introduces and discusses a graphical representation of the age-specific distribution of the sex composition of a population which we call the A-curve. Section 4 advances two real-valued measures of femaleness, F^* and S^*, which are derived from the A-curve; the difference between the measures F^* and S^* on the one hand and conventional indices like F and S on the other is stressed, by reference to the 'transfer' property which F^* and S^* satisfy and F and S violate. Section 5 presents some empirical illustrations -- based on demographic data from India -- of the measurement-related concerns reviewed in the earlier sections. Concluding observations are offered in Section 6.

2. TWO ELEMENTARY MEASURES OF 'FEMALENESS'

Consider a population whose total size is designated by P. Let P^f and P^m stand, respectively, for the sizes of the female and male populations (so that $P^f+P^m = P$). The most elementary index of 'femaleness' one can think of is what we shall call the *female headcount ratio*, F, which simply measures the proportion of females in the population:

(2.1) $F = P^f/P$.

For future reference, let us also define the *male headcount ratio*, M, which measures the proportion of males in the total population:

$$(2.1') \quad M = P^m/P.$$

The most widely-employed measure of 'femaleness' of a population is the so-called *sex-ratio*, S, which is also called the *female-to-male ratio*[1], and is given by:

$$(2.2) \quad S = P^f/P^m.$$

The indices S and F convey exactly the same information, and it is readily apparent that for any two societies 1 and 2, $S_1 \gtreqless S_2$ according as $F_1 \gtreqless F_2$ where S_i [respectively, F_i] is the sex-ratio [respectively, the female headcount ratio] for society i (i =1,2). (From the definitions of F and S furnished in (2.1) and (2.2) respectively, it is immediate that $S = F/(1-F)$ and $F = S/(1+S)$).

The overall weight of women in a population is clearly some aggregation of their *age-specific weights* in the population. For example, the index F is 'age-wise decomposable' in the sense of being a (population-share) weighted average of the age-specific values of F. Specifically, let α be a random variable denoting age, and assuming values in the interval $[0,\bar{\alpha}]$. For every $\alpha\in[0,\bar{\alpha}]$, define $F(\alpha) := P^f(\alpha)/P(\alpha)$ where $P^f(\alpha)$ is the female population at age α and $P(\alpha)$ is the total population at age α. If $t(\alpha) := P(\alpha)/P$ is the proportion of the total population at age α, then it is clear that the overall proportion of females in the total population, F, is simply the population-share weighted average of the age-specific proportions of females in the population:

$$(2.3)\ F = \int_0^{\bar{\alpha}} F(\alpha)t(\alpha)d\alpha.$$

The distribution of 'femaleness' by age is an important demographic datum; and the assessment of the 'femaleness' of a population is greatly facilitated by a graphic depiction of this distribution. Such a depiction is made possible through a contrivance which we call the 'age-distributed gender composition curve' -- or A-curve for short -- for a population. This curve is introduced in the following section.

3. THE A-CURVE

For every age $\alpha \in [0, \bar{\alpha}]$ define

$$(3.1)\ G^f(\alpha) := (1/P)\int_0^{\alpha} P^f(\beta)d\beta$$

and

$$(3.1')\ G^m(\alpha) := (1/P)\int_0^{\alpha} P^m(\beta)d\beta.$$

Also, for every $\alpha \in [0, \bar{\alpha}]$, let

$$(3.2)\ g^f(\alpha) := dG^f(\alpha)/d\alpha = P^f(\alpha)/P;$$

and

$$(3.2')\ g^m(\alpha) := dG^m(\alpha)/d\alpha = P^m(\alpha)/P.$$

That is, for every age α, $G^f(\alpha)$ [respectively, $G^m(\alpha)$] is the cumulative number of females [respectively, males] of age not exceeding α, expresed as a proportion of the total population. The derivative of $G^f(\alpha)$ [respectively, $G^m(\alpha)$], which we call $g^f(\alpha)$

[respectively, $g^m(\alpha)$], is simply the number of females [respectively, males] of age α expressed as a proportion of the total population. [Note that $g^f(\alpha)/t(\alpha)$ and $g^m(\alpha)/t(\alpha)$ are regular density functions, just as $G^f(\alpha)/t(\alpha)$ and $G^m(\alpha)/t(\alpha)$ are regular cumulative density functions].

The *age- distributed gender composition curve*, or *A-curve* for short, is simply the graph of $L^f(\alpha) := [F - G^f(\alpha)]$ plotted against $L^m(\alpha) := [M - G^m(\alpha)]$ over the range $[0, \bar{\alpha}]$ of α, *with the sequence of α running from largest age to smallest age, i.e. in descending order*. The first point on the A-curve will therefore be $(M - G^m(\bar{\alpha}), F-G^f(\bar{\alpha})]$ while the last point will be $(M-G^m(0), F-G^f(0))$. The following features of the A-curve merit attention.

(i) Since $G^m(\bar{\alpha}) = M$ and $G^f(\bar{\alpha}) = F$, the first point on the A-curve is the natural origin (0,0); further, since $G^m(0) = G^f(0) = 0$, the final point on the A-curve is the point (M,F). The A-curve is therefore a nondecreasing curve going from (0,0) to some point on the hypotenuse of the right-angled triangle constituted by the lower left half of the unit square (see Figure 3.1)

[Figure 3.1 to be inserted here]

(ii) Making use of (3.2) and (3.2'), one can see that the slope of the A-curve, at any given age α, $\sigma(\alpha)$, is given by:
$\sigma(\alpha) = dL^f(\alpha)/dL^m(\alpha) = [d(F-G^f(\alpha))/d\alpha]/[d(M-G^m(\alpha))/d\alpha] = [-g^f(\alpha)]/[-g^m(\alpha)] = [P^f(\alpha)/P]/[P^m(\alpha)/P]$ or, in view of (2.2)

(3.3) $\sigma(\alpha) = S(\alpha)$.

That is, the slope of the A-curve at any age α is simply the sex-ratio at that age. The A-curve is therefore a useful device for obtaining a composite visual picture of the age-specific distribution of sex-ratios in a population. In a 'demographically developed' society, it will typically be the case that the sex-ratio at birth is less than unity but exceeds unity as we climb up the age-ladder, so that the A-curve is an increasing, (possibly) strictly concave function lying above the 45° line. Wherever it occurs, any contrast from this picture of the 'norm' can be readily inferred from the slope and location of the A-curve.

(iii) Theoretically, the case of *complete female disadvantage* in the sex composition of a population obtains when the A-curve coincides with the base of the right-angled triangle: this is a situation in which the population has no females. Contrarily, the case of *complete female advantage* obtains when the A-curve coincides with the vertical side of the right-angled triangle: this is a situation in which there are no males in the population.

(iv) The 45° line drawn from the origin to the hypotenuse of the right-angled triangle could be called 'the line of gender-equality': along this line the number of females exactly balances the number of males at every age. If the A-curve lies everywhere above [respectively, everywhere below] the 45° line, then this reflects *unambiguous female advantage* [respectively, *unambiguous female disadvantage*] in terms of the sex composition of the population. This is illustrated in Figure 3.2.

[Figure 3.2 to be inserted here].

(v) The A-curve can also be written as a function $p^f(p^m(\alpha))$: for the oldest p^mth fraction of males in the total population, with p^m ranging from 0 to M, the A-curve plots the corresponding oldest p^fth fraction of females in the total population, with p^f ranging from 0 to F. For any two populations 1 and 2, we shall say that the A-curve for society 1, A_1, *weakly dominates* the A-curve for society 2, A_2, whenever A_1 lies nowhere below A_2. Formally, letting $\succeq$ stand for the weak dominance relation, we have:

$$(3.4)\quad A_1 \succeq A_2 \text{ iff } p_1^f(p^m) \geq p_2^f(p^m)\ \forall p^m \in [0, M_1],$$

where the subscripts 1 and 2 stand for populations 1 and 2 respectively. Clearly, $A_1 \succeq A_2$ implies that the female advantage in terms of the sex composition of the population is unambiguously at least as much in society 1 as in society 2.

For any two societies 1 and 2 we shall say that A_1 *(strictly) dominates* A_2 if A_1 lies nowhere below and somewhere above A_2: that is, letting $\succ$ stand for the (strict) dominance relation, we have:

$$(3.4')\quad A_1 \succ A_2 \text{ iff } A_1 \succeq A_2 \text{ and } \textit{not}\ [A_2 \succeq A_1].$$

Finally, for any two societies 1 and 2, we shall say that *A_1 is indistinguishable from A_2* if A_1 everywhere coincides with A_2: that is, letting ~ stand for the indistinguishability relation, we have:

$$(3.4'')\quad A_1 \sim A_2 \text{ iff } A_1 \succeq A_2\ \&\ A_2 \succeq A_1.$$

Figure 3.3 illustrates a case of A-dominance.

[Figure 3.3 to be inserted here].

(vi) In discussing A-dominance, it is frequently more convenient to deal with a *discrete* distribution than with a continuous one which latter is the one we have worked with thus far. A *discrete age-distributed sex profile* is a 3K-vector $q = (a_1,\ldots,a_K;n_1^f,\ldots,n_K^f;n_1^m,\ldots,n_K^m)$, which conveys the following information: $a_1,\ldots,a_K$ stand for K distinct ages *indexed in descending order*, viz. $a_1>a_2>\ldots>a_K$; for all $i=1,\ldots,K$, n_i^f stands for the number of females at age a_i; and for all $i=1,\ldots,K$, n_i^m stands for the number of males at age a_i. Clearly, $\Sigma_{i=1}^K n_i^f$ stands for the total number of females P^f in the society; $\Sigma_{i=1}^K n_i^m$ stands for the total number of males P^m in the society; and $(\Sigma_{i=1}^K n_i^f + \Sigma_{i=1}^K n_i^m)$ stands for the size P of the total population. We describe below how to obtain an A-curve for a discrete distribution[2].

Recalling that M [respectively, F] stands for the male [respectively, female] headcount ratio, for every $i=1,\ldots,K$, define $p_i^m := (M - \Sigma_{j=i}^K n_j^m/P)$: p_i^m stands for the cumulative number of males of age exceeding a_i, expressed as a fraction of the total population. For the female population we can, analogously, define $p_i^f := (F - \Sigma_{j=i}^K n_j^f/P)$, $i=1,\ldots,K$. Define $p_{K+1}^m := M$ and $p_{K+1}^f := F$.

Then, given any discrete age-distributed sex profile q, we can see that the corresponding age-distributed sex composition curve A(q) is given by a plot of the points (p_1^m,p_1^f), $(p_2^m,p_2^f),\ldots,(p_K^m,p_K^f)$, (p_{K+1}^m,p_{K+1}^f), these points being connected by 'piece-wise' linear segments.

We now derive a result on A-dominance for societies with the same population size and the same overall sex-ratio. Before stating the result, we first describe the notion of one sex profile being derived from another through a *favourable transfer*. To this end, consider any two profiles q and $\bar{q}$ such that

$$q = (a_1,\ldots,a_K;n_1^f,\ldots,n_K^f;n_1^m,\ldots,n_K^m);\text{ and}$$

$$\bar{q} = (a_1,\ldots,a_K;\bar{n}_1^f,\ldots,\bar{n}_K^f;\bar{n}_1^m,\ldots,\bar{n}_K^m),$$

where $\bar{n}_i^m=n_i^m$ for all $i=1,\ldots,K$; and $\bar{n}_i^f=n_i^f$ for all $i\in\{1,\ldots,K\}\setminus\{j,k\}$

for some j,k satisfying $a_j>a_k$; $\bar{n}_j^f=n_j^f+1$ and $\bar{n}_k^f=n_k^f-1$.

Of a pair of profiles q and $\bar{q}$ exhibiting the properties described above, we shall say that *$\bar{q}$ is derived from q through a favourable transfer*; that is, $\bar{q}$ is derived from q through a favourable transfer whenever, *ceteris paribus*, $\bar{q}$ contains one female more at age a_j than q and one female less at age a_k than q, with a_j being greater than a_k: it is 'as if' one female at age a_k has been 'transferred' to the population at age a_j.

The following proposition is now true. For all sex profiles q,$\bar{q}$ such that $\bar{q}$ is derived from q through a favourable transfer, it is the case that the A-curve $\bar{A}$ corresponding to $\bar{q}$ dominates the A-curve A corresponding to q. To see this, consider WLOG, a situation in which K = 4, j = 2 and k = 3, so that:

$$q = (a_1,a_2,a_3,a_4;n_1^f,n_2^f,n_3^f,n_4^f;n_1^m,n_2^m,n_3^m,n_4^m);\text{ and}$$

$$\bar{q} = (a_1,a_2,a_3,a_4;n_1^f,n_2^f+1,n_3^f-1,n_4^f;n_1^m,n_2^m,n_3^m,n_4^m).$$

It is immediately clear that the total population -- call it P -- is the same for both q and $\bar{q}$, just as the male and the female headcount ratios -- call them M and F respectively -- are the same for both q and $\bar{q}$; and $\bar{q}$ is derived from q through a favourable transfer (from the population at age a_3 to the population at age a_2). Letting p_i^m [respectively, $\bar{p}_i^m$] stand for the value of the cumulative fraction of males in the total population corresponding to the profile q [respectively, $\bar{q}$], we have:

$p_1^m=\bar{p}_1^m=0$; $p_2^m=\bar{p}_2^m=M-\Sigma_{j=2}^4 n_j^m/P$; $p_3^m=\bar{p}_3^m=M-\Sigma_{j=3}^4 n_j^m/P$; $p_4^m=\bar{p}_4^m=M-n_4^m/P$; and

$p_5^m=\bar{p}_5^m=M$.

Further,

$p_1^f=\bar{p}_1^f=0$; $p_2^f=\bar{p}_2^f=F-\Sigma_{j=2}^4 n_j^f/P$; $p_3^f=F-\Sigma_{j=3}^4 n_j^f/P$ and $\bar{p}_3^f=F-\Sigma_{j=3}^4 n_j^f/P+1/P$;

$p_4^f=\bar{p}_4^f=F-n_4^f/P$; and $p_5^f=\bar{p}_5^f=F$.

The curves A and $\bar{A}$ are obtained, respectively, by plotting the following sets of pairs of points:

A : (0,0); $(M-\Sigma_{j=2}^4 n_j^m/P, F-\Sigma_{j=2}^4 n_j^f/P)$; $(M-\Sigma_{j=3}^4 n_j^m/P, F-\Sigma_{j=3}^4 n_j^f/P)$;

$(M-n_4^m/P, F-n_4^f/P)$; (M,F); and

$\bar{A}$: (0,0); $(M-\Sigma_{j=2}^4 n_j^m/P, F-\Sigma_{j=2}^4 n_j^f/P)$; $(M-\Sigma_{j=3}^4 n_j^m/P, F-\Sigma_{j=3}^4 n_j^f/P+1/P)$;

$(M-n_4^m/P, F-n_4^f/P)$; (M,F).

It is immediately clear from the above that the point $(\bar{p}_3^m, \bar{p}_3^f)$ lies vertically above the point (p_3^m, p_3^f) while the pairs of points

$[(p_1^m, p_1^f), (\bar{p}_1^m, \bar{p}_1^f)]$, $[(p_2^m, p_2^f), (\bar{p}_2^m, \bar{p}_2^f)]$, $[(p_4^m, p_4^f), (\bar{p}_4^m, \bar{p}_4^f)]$ and

$[(p_5^m, p_5^f), (\bar{p}_5^m, \bar{p}_5^f)]$ coincide with each other. That is to say, $\bar{A}$ dominates A. We have therefore demonstrated that if one sex profile is derived from another through a favourable transfer, then the A-curve corresponding to the former profile strictly dominates the A-curve corresponding to the latter profile. We shall have occasion to return to this result at a later stage.

(vii) Does the A-curve reveal anything about the magnitudes of the real-valued indices F and S? It does so in a very straightforward way, as Figure 3.4 suggests. The slope of the straight line connecting the points (0,0) and (M,F) is M/F which is just the sex-ratio; while the height of the point (M,F) is just the female headcount ratio.

[Figure 3.4 to be inserted here].

(viii) A few elementary facts regarding the connection between the A-dominance relation and dominance in terms of the indices F and S are worth noting. First, let us say that for any two societies 1 and 2 with corresponding A curves A_1 and A_2 respectively, A_1 *super-dominates* A_2 -- written $A_1 \succ\succ A_2$ -- if A_1 lies *everywhere* above A_2 (except of course at the origin where the two curves coincide). Then, it is immediate that if for any pair of populations 1 and 2, $A_1 \succ\succ A_2$, it must be the case that $S_1 > S_2$ and $F_1 > F_2$: this is revealed clearly in Figure 3.5(a). As the figure indicates, whenever $A_1 \succ\succ A_2$, $F_1 > F_2$; and as we have already seen, $F_1 > F_2$ implies $S_1 > S_2$.

Second, however, the converse is not true; for any two populations 1 and 2, if $F_1 > F_2$ (and therefore $S_1 > S_2$), then this is no guarantee that $A_1 \succ A_2$ (much less that $A_1 \succ\succ A_2$): with intersecting A-curves, no unambiguous ranking in terms of the A-dominance relation may be possible even if we have an unambiguous dominance ranking in terms of the indices F and S, as Figure 3.5(b) makes clear.

Thirdly, it is always possible in principle that for a pair of populations 1 and 2, $F_1 = F_2$ (and therefore $S_1 = S_2$), but $A_1 \succ A_2$. Such a possibility is illustrated in Figure 3.5(c). As the figure makes clear, the sex-ratio of the entire population -- obtained as a population-share weighted average of all age-specific sex-ratios -- may be the same for two different societies; but the A-dominance relation will favour that society in which the sex-ratios at the older ages are relatively higher than the sex-ratios at the younger ages.

A related point concerning the general significance of the shape of the A-curve can be analyzed along the following lines. For any given A-curve there exists a distinguished counterpart -- the A^*-curve -- which one may call the 'linearized A-curve', and which is obtained by assigning the overall female headcount ratio to the population at *each* age in the interval $[0, \bar{\alpha}]$, so that the resulting A-curve is the linear graph connecting the origin to the point (M,F) on the hypotenuse of the right-angled triangle (see Figure 3.5(d)). If it is the case that $A \succ A^*$, then -- in view of the discussion in paragraph (vi) earlier -- it is clear that A can be seen to have been derived from A^* through at least one favourable transfer; and if it is the case that $A^* \succ A$, then A^* can be seen to have been derived from A through at least one favourable transfer. As Figure 3.5(d) makes clear, strict concavity of A ensures that $A \succ A^*$; linearity of A ensures that $A \sim A^*$; and strict convexity of A ensures that $A^* \succ A$. The general point which emerges is that a given overall female headcount ratio F (or sex-ratio S) is compatible with many different patterns of the age-specific sex composition of the population: a neglect of the specific pattern which obtains in a given case could cause F (or S) to be an unreliable guide to the 'extent of femaleness' of the population. In particular, if A is strictly concave we have a situation in which the slope of the A-curve is declining: that is, the age-specific sex-ratio declines as we move from largest to smallest age. Conversely, if the A-curve is strictly convex, we have a situation in which the slope of the A-curve is increasing: that is, the age-specific sex-ratio increases as we move from largest to smallest age. Briefly, if our underlying 'welfare' calculus suggests that a given level of femaleness acquires greater significance the greater the age at which it obtains, then this is reflected -- *via* the A-dominance relation -- in the judgment that, other things equal, a strictly concave A-curve is 'welfare-superior' to a linear one, while a linear A-curve is 'welfare-superior' to a strictly convex one.

[Figures 3.5(a)-3.5(d) to be inserted here].

(ix) It should have occurred to the reader that the device we call the A-curve is reminiscent of the Lorenz curve so widely invoked in the literature on the measurement of income-inequality, and perhaps even more proximately reminiscent of the 'segrgation curve' which is commonly employed in the sociological literature dealing with the assessment of discrimination. Proceeding with the analogy, it is tempting to consider the possibility of deriving a real-valued measure of female advantage in the gender composition of a population which is based on the A-curve in much the same way in which the Gini coefficient of income inequality is based on the Lorenz curve. To this issue we now turn[3].

4. TWO ALTERNATIVE MEASURES OF FEMALENESS: F^* AND S^*

A natural measure of the extent of 'femaleness' of a population -- call it $\hat{F}$ -- is given by the area to the right of the A-curve in the right-angled triangle. This area, as can be seen from Figure 4.1, can be broken up into two areas: the area A under the A-curve, and the area B of the triangle to the right of the area A.

[Figure 4.1 to be inserted here].

It is easy to see that

$$\text{Area A} = \int_{\bar{\alpha}}^{0}(F-G^f(\alpha))(-g^m(\alpha))d\alpha$$

$$= \int_{0}^{\bar{\alpha}}(F-G^f(\alpha))g^m(\alpha)d\alpha$$

$$= F\int_{0}^{\bar{\alpha}}g^m(\alpha)d\alpha - \int_{0}^{\bar{\alpha}}G^f(\alpha)g^m(\alpha)d\alpha$$

$$= FM - \int_{0}^{\bar{\alpha}}G^f(\alpha)g^m(\alpha)d\alpha$$

$$= F(1-F) - \int_{0}^{\bar{\alpha}}G^f(\alpha)g^m(\alpha)d\alpha.$$

Further, we have:

$$\text{Area B} = F^2/2.$$

Putting these results together, we obtain:

$$\hat{F}\ (= \text{Area A} + \text{Area B}) = F(1-F) - \int_{0}^{\bar{\alpha}}G^f(\alpha)g^m(\alpha)d\alpha + F^2/2,$$

or, simplifying,

(4.1) $\hat{F} = [F(2-F) - 2\int_0^{\bar{\alpha}} G^f(\alpha)g^m(\alpha)d\alpha]/2.$

To obtain a *normalized* index of 'femaleness', we just divide $\hat{F}$ by the maximum value the latter can attain. $\hat{F}$ attains its maximum value when the A-curve coincides with the vertical side of the right-angled triangle, i.e. when the population has no males at all. In this event, the value of $\hat{F}$ -- call it $\hat{F}_{max}$ -- is just the area of the right-angled triangle, which is one-half. The normalized index of 'femaleness', in view of (4.1), is then given by

(4.2) $F^* \ (= \hat{F}/\hat{F}_{max}) = F(2-F) - 2\int_0^{\bar{\alpha}} G^f(\alpha)g^m(\alpha)d\alpha.$

Clearly, F^* lies in the closed interval [0,1], with a larger value of F^* signifying a greater degree of 'femaleness' of the population.

The relationship between the index F^* and the index F, as mediated by the slope of the A-curve, is straightforwardly apparent. Specifically, if the A-curve is linear, then it can be shown that F^* equals F. To see this, note that the linearity of the A-curve implies that the sex-ratio at every age α in the interval $[0,\bar{\alpha}]$ is the same, and equal to the overall sex-ratio S. As we have seen earlier, the sex ratio at age α is given by $S(\alpha) = g^f(\alpha)/g^m(\alpha)$, and the overall sex-ratio S is given by $S = F/(1-F)$. Consequently, for a linear A-curve, one has: $g^f(\alpha)/g^m(\alpha) = F/(1-F) \ \forall\alpha\in[0,\bar{\alpha}]$, i.e.,

(4.3) $g^m(\alpha) = [(1-F)/F]g^f(\alpha) \ \forall\alpha\in[0,\bar{\alpha}].$

Substituting for $g^m(\alpha)$ from (4.3) into the last term on the Right Hand Side of (4.2) yields:

$$\int_0^{\bar{\alpha}} G^f(\alpha)g^m(\alpha)d\alpha = [(1-F)/F]\int_0^{\bar{\alpha}} G^f(\alpha)g^f(\alpha)d\alpha = [(1-F)/F][F^2/2],$$

whence

$$(4.4)\quad \int_0^{\bar{\alpha}} G^f(\alpha)g^m(\alpha)d\alpha = F(1-F)/2.$$

Substituting for the integral from (4.4) into (4.2) yields:

$$(4.5)\quad F^* \; [= F(2-F) - F(1-F)] = F,$$

as desired.

Given (4.5), inspection of Figure 4.2 assures us that (i) if the A-curve is strictly concave (which guarantees that $A \succ A^*$), then $F^* > F$; (ii) if the A-curve is linear (which guarantees that $A \sim A^*$), then $F^* = F$; and if the A-curve is strictly convex (which guarantees that $A^* \succ A$), then $F^* < F$. What does this imply? Figure 4.2 depicts three age-distributed gender composition profiles which yield the same overall female headcount ratio F; but F^* -- in relation to F -- 'rewards' the strictly concave profile 2 and 'penalizes' the strictly convex profile 3, thereby realizing the underlying 'welfare' judgment that 'favourable transfers' are welcome and 'unfavourable transfers' are unwelcome. This is largely an echo of our earlier discussion in paragraph (viii) of Section 3.

[Figure 4.2 to be inserted here].

A significant feature which emerges from the preceding discussion is that the index F^* satisfies a property embodied in an axiom we shall call the *transfer axiom*. In defining the axiom we again take recourse to the convenience of a discrete distribution.

The Transfer Axiom. Let Φ be any real-valued index of 'femaleness' defined on the set of discrete age-distributed sex profiles. Φ will be said to satisfy the *transfer axiom* if, for all pairs of profiles $(q,\bar{q})$ such that $\bar{q}$ is derived from q through a favourable transfer, it is the case that $\bar{\Phi}$ $(:= \Phi(\bar{q})) > \Phi$ $(:= \Phi(q))$.

That F^* satisfies the transfer axiom is very easily demonstrated. Consider a pair of profiles $q,\bar{q}$ such that $\bar{q}$ is derived from q through a favourable transfer. Define $\bar{F}^* := F^*(\bar{q})$ and $F^* := F^*(q)$. Since $\bar{q}$ is derived from q through a favourable transfer, we know from earlier discussion (paragraph (vi) in section 3) that $\bar{A} \succ A$, where $\bar{A}$ [respectively, A] is the A-curve corresponding to the profile $\bar{q}$ [respectively, q]. By definition of F^*, $\bar{A} \succ A$ implies $\bar{F}^* > F^*$, as desired.

The index F, of course, fails the transfer axiom. (This is immediately clear from the fact that no matter what the actual shape of the A-curve, F will always be computed as the area to the right of the (linear) A^*-curve: so whether A is derived from A^* through a favourable or an unfavourable transfer will not affect the value of F). Briefly, the 'standard' indices F and S are concerned only with the overall 'femaleness' of a population averaged over age-specific levels of 'femaleness', so that any age-specific deviations from the average simply 'come out in the wash'. By contrast, the index F^* *is* sensitive to variations in age-specific female headcount ratios, and any deviation of F^* from F is attributable to the presence of such variations. (It should be emphasized that while it is true that if $F^* \neq F$ then this is because of a non-linearity in the A-curve, the converse does not hold; that is, it is not necessarily true that if A is non-linear then it must be the case that $F^* \neq F$).

Our discussion also suggests that it is conceivable in principle that one can have a pair of populations 1 and 2 such that $F_1 > F_2$ but $F_1^* < F_2^*$: the possibility of rank-reversal suggests that how we choose to measure the 'femaleness' of a population can have non-trivial implications for our judgments regarding 'more' and 'less'. Specifically -- and as we have already seen -- indices like F and S are measures of 'central tendency' which ignore the question of variability about the mean; to the extent that dispersion is also considered to be important in the scheme of things, an index such as F^* clearly has some interpretational advantage to offer. F^*, concretely, can be viewed as a 'corrected' version of F, where the correction assumes the form of adjustment for variability in the age-specific distribution of the population's gender composition. Indeed, it is easy to see that a normalized measure of the extent of dispersion around the mean is given by the quantity $D = (F^* - F)/F$, so that F^* itself can be written as $F^* = F(1+D)$: F^*, that is, is just the overall female headcount ratio enhanced or reduced by the extent of variability of age-specific female headcount ratios about their mean value, with $F^* \gtreqless F$ according as $D \gtreqless 0$[4].

In this connection one can, by analogy, advance a 'corrected' version S^* of the standard sex-ratio measure S. Specifically, recall that the index S is the ratio of the female headcount ratio F to the male headcount ratio M; the 'corrected' version of S, by analogy, can be seen as the ratio of the 'corrected' female headcount ratio F^* to the 'corrected' male headcount ratio M^*

(4.6) $S^* \; [= F^*/M^*] = F^*/(1-F^*)$.

In the Appendix to this paper, we consider the pragmatic question of how to actually compute the value of the index F^* from data on the distribution of age-specific sex-ratios as these are typically available in data sources, namely in *grouped* form.

It remains to provide a few examples from actual data sets of some of the measurement concerns reviewed in this note. Accordingly, some simple empirical applications based on demographic data from India are furnished in the following section.

5. SOME EMPIRICAL ILLUSTRATIONS FROM INDIAN DATA

In this section, we shall employ data provided in the Age Tables of the Census of India in order to give empirical content to some of the measurement issues that have been discussed in earlier sections. The accent will be only on providing illustrative examples of this or that aspect of measurement, and not at all on any substantive analysis of findings. In what follows, we shall be drawing on information provided in four Censuses - those of 1961, 1971, 1981 and 1991. Since the context of the present discussion does not warrant any sort of detailed critique of the reliability of the data source, we shall for the most part be *accepting the data at face value.*

We shall be concerned to obtain estimates of the indices F^* and S^* from grouped data for various states of the Indian Union and for the country as a whole. To this end we classify the population into the following fifteen non-overlapping age groups:

0-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69 and 70 and above. We shall, of course, be assuming throughout that the population within any age group is uniformly distributed. (It may be added that while single-year age data are available, we have resorted to grouping in order to mitigate biases that would arise from the clustering of populations around particular ages owing to 'digit preference' in the matter of age-reporting). Given these preliminary remarks, here are a few illustrations of the use to which data on the age-distribution of sex-ratios can be put.

Example 5.1. Kerala and Uttar Pradesh: The Good and the Bad

Tables 5.1(a) and 5.1(b) provide information, from the Census of 1991, on the age-wise sex composition of the population for two states of the Indian Union: Kerala in the south and Uttar Pradesh in the north. Our choice of these two states is intended to highlight the contrast that obtains between them in terms of one aspect of the status of women in a society, namely the 'femaleness' of the population: Kerala has long been recognized as a 'front-runner' in this respect, just as Uttar Pradesh has been recognized as a 'laggard'.

In Table 5.1(a), which provides data for Kerala, we have fifteen age-groups, indexed in descending order: the age-group '70 and above' is group 1 and the age-group '0-4' is group 15. For each group $i=1,\ldots,15$, columns 2, 3 and 4 provide information on the number of males, females and all persons — P_i^m, P_i^f and P_i — respectively, while columns 5 and 6 furnish information on the fraction of the total population accounted for by males and females — g_i^m $(:=P_i^m/P)$ and g_i^f $(:=P_i^f/P)$ — respectively. In columns 7 and 8 we have the cumulative fractions in total population of males and females of age not exceeding the upper limit of age-group i — r_i^m $(:= \sum_{j=i}^{15} g_j^m)$ and r_i^f $(:= \sum_{j=i}^{15} g_j^f)$ — respectively, for each group i=1,...,15. Finally, in columns 9 and 10, for each group i=1,...,15, information is available on the cumulative fraction of the total population of males and females of age exceeding the upper limit of age-group i — p_i^m $(:= M-r_i^m)$ and p_i^f $(:= F-r_i^f)$ — respectively. Columns 9 and 10 are crucial for the construction of the A-curve: the coordinates of the A-curve are constituted by the set of points $(p_1^m, p_1^f), \ldots, (p_{16}^m, p_{16}^f)$ where, by definition, $p_{16}^m := M$ and $p_{16}^f := F$. Table 5.1(b) for U.P is identical in construction to Table 5.1(a) for Kerala, and calls for no comment.

[Tables 5.1 (a) and 5.1(b) to be inserted here]

From the data presented on the coordinates of the A-curve in columns 9 and 10 of Tables 5.1(a) and 5.1(b), we have, in Figure 5.1, plotted the A-curves for Kerala and U.P. respectively. The contrast between the two states is immediately visually striking. Figure 5.1 is an illustrative example of (a) the notion of 'A-dominance' and (b) the notions of 'unambiguous female advantage' and 'unambiguous female disadvantage'. Specifically, the A-curve for Kerala strictly dominates the A-curve for U.P. Further, since the Kerala A-curve lies everywhere above the 45° line, we have here a case of 'unambiguous female advantage' in the sex composition of the population; contrarily, in the case of U.P. the A-curve lies everywhere below the 45° line, and this reflects an instance of 'unambiguous female disadvantage'.

[Figure 5.1 to be inserted here]

Next, using the data provided in Columns 9 and 10 of Tables 5.1(a) and 5.1(b) respectively, we can compute the values of F^* for Kerala and U.P., recalling from equation (A.6) in the Appendix that the formula for the 'trapezoidal' approximation of F^* is given by: $F^* \cong \sum_{i=2}^{16} (p_i^m - p_{i-1}^m)(p_i^f + p_{i-1}^f) + F^2$. The corrected version S^* of the sex-ratio S, further, is given by $S^* = F^*/(1-F^*)$: the S^*-values for Kerala and U.P. can be routinely estimated once their respective F^*-values have been computed. Table 5.1(c) furnishes the values of F, F^*, S and S^* for each of the states of Kerala and U.P. It can be seen from Table 6.1(c) that F^* (respectively, S^*), in relation to F (respectively, S), *rewards* Kerala and *penalizes* U.P.: thus, while the contrast between Kerala and U.P. in terms of S is striking enough (the S-values for the two states are 1036 and 882 respectively), the contrast in terms of S^* becomes even more pronounced (the S^*-values for the two states are 1058 and 871 respectively).

[Table 5.1(c) to inserted here]

<u>Example 5.2: West Bengal and Rajasthan: A Case of Rank-reversal</u>

Tables 5.2(a) and 5.2(b), which provide information on the age-specific sex composition of the populations in the states of West Bengal and Rajasthan for the year 1991, duplicate the construction of Table 5.1(a). From the data available in columns 9 and 10 of the two tables, we can plot the A-curves for the two states, and also compute the 'corrected' values F^* and S^* respectively of the female headcount ratio F and the sex-ratio S. Figure 5.2 depicts the A-curves for West Bengal and Rajasthan, and Table 5.2(c) furnishes the values of F, F^*, S and S^* for the two states.

[Tables 5.2(a),5.2(b) and 5.2(c) and Figure 5.2 to be inserted here]

In section 3, it was remarked that one could in principle have a pair of societies 1 and 2 such that F_1 [respectively, S_1] > F_2 [respectively, S_2], but F_1^* [respectively, S_1^*] < F_2^* [respectively, S_2^*]. Precisely such an instance of rank-reversal is discernible in the case of West Bengal and Rajasthan: the sex-ratio for West Bengal, at 918, is higher than that for Rajasthan, at 911; but the S^*-value for West Bengal, at 895, is *less* than that for Rajasthan, at 913. As we have seen earlier, F^* (and therefore S^*) would — in relation to F (and therefore S) — tend to favour a distribution in which the sex-ratios at the upper ages are relatively larger than the sex-ratios at the lower ages. This is precisely what happens in the case of Rajasthan *vis-a-vis* West Bengal: as figure 5.2 makes clear, there is a range of ages in the upper age-groups over which the A-curve for Rajasthan clearly lies above the A-curve for West Bengal — and this to a point where F^* reverses the ranking of the two states according to F.

5.3 A Cross-sectional View of the 'Femaleness' of India's Population: 1991

Using state-wise data from the Census of 1991 on the age-distributed gender composition of the population, we present in Table 5.3 the values of the conventional indices of 'femaleness' F and S, and those of their corresponding 'corrected' versions F^* and S^*, for thirteen states of the Indian Union. Some salient features of the figures reported in Table 5.3 are the following. First, the ranking of states according to S (or F) and that according to S^* (or F^*) are quite similar: Spearman's rank correlation coefficient for the two sets of rankings is high, at

0.9725. Second, an examination of pair-wise rankings of states reveals that there are only four instances of rank reversals from a possible 121 pairwise comparisons: these are the cases pertaining to Tamilnadu and Andhra Pradesh, West Bengal and Bihar, Bihar and Rajasthan, and Rajasthan and West Bengal. While the particular data set we have employed suggests that the probability of rank-reversal is small, this in itself is not a cause for excessive complacence about the use of indices such as F and S. Specifically, if for some reason our concern were confined to the subset of five states constituted by Tamilnadu, Andhra Pradesh, West Bengal, Bihar and Rajasthan, then the ranking of the states in descending order according to F [respectively, F^*] would be: Tamilnadu, Andhra Pradesh, West Bengal, Bihar, Rajasthan [respectively, Andhra Pradesh, Tamilnadu, Rajasthan, Bihar, West Bengal], so that Spearman's rank correlation coefficient for the two sets of rankings, at 0.50, is now only moderately positive (and indeed not significantly different from zero at the 5 per cent level of significance). Finally, if we are interested in obtaining a picture of the extent of inter-state variation in the degree of 'femaleness' of the population, then a means to this end would be to employ the squared coefficient of variation as a measure of dispersion. It turns out that if 'femaleness' is measured by S, then for our set of thirteen states the squared coefficient of variation in state-specific S-values is 0.0019, while if 'femaleness' is measured by S^*, then the squared coefficient of variation is higher, at 0.0025, by a factor of 132 per cent. Briefly, for more than one reason, it cannot be a matter of indifference whether we employ S or S^* as the 'appropriate' measure of 'femaleness' of a population.

5.4 A Time-Series View of the 'Femaleness' of India's Population: 1961-1991

From data in the Censuses of 1961, 1971, 1981 and 1991, we present in Table 5.4(a), for each of these years, information on the age-specific sex composition of the Indian population. Based on these data, Table 5.4(b) furnishes the computed values of the 'femaleness' indices F, F^*, S and S^* for each of the four years under review. As can be seen from Table 5.4(b), S^* is less than S in each year, but the difference between S and S^* as a proportion of S follows a generally declining pattern over time. A result of this is that we have two rank-reversals: while S for 1961 is greater than S for 1981, S^* for 1981 exceeds S^* for 1961; and similarly, while S for 1971 is greater than S for 1991, the 1991 S^* - value exceeds the 1971 S^* - value.

To see how variations in the age-specific distribution of sex-ratios can bring about such rank-reversals, consider the following somewhat gross calculations for the years 1961 and 1981. Let us divide the population into just two non-overlapping age-groups: 0-29 and 30+, and let us call these Groups I and II respectively. From the data provided in Table 5.4(a), it can be verified that Group I's sex-ratio has *declined* from 961 in 1961 to 939 in 1981, while Group II's sex-ratio has *increased* from 904 in 1961 to 927 in 1981. Since S^* attaches greater significance to a given level of the sex-ratio the higher the age at which it obtains, the 'favourable' impact of the rise in the sex-ratio at the upper age-group II swamps the 'unfavourable' impact of the fall in the sex-ratio at the lower age-group I, resulting in S^* for 1981 exceeding S^* for 1961.

What is the overall effect of these rank-reversals on our appreciation of temporal trends? In Table 5.4(c) we present a ranking of the four Censal years 1961, 1971, 1981, and 1991 in terms of (a) their vintage (viz. the year - 1961 - which is most distant in the past from the present receives a rank of 1, the next most distant year - 1971 - receives a rank of 2, etc.); (b) their S-values (in descending order); and (c) their S^*-values (also in descending order). If we had a monotonous decline in the 'femaleness' index over time, then Spearman's coefficient of rank correlation between the ranking according to vintage and the ranking according to the 'femaleness' index should be unity. As it happens, the coefficient of rank correlation between the ranking according to vintage and that according to S is highly positive, at 0.8, but drops to zero when we seek to correlate the ranking according to vintage with the ranking according to S^*. This example again is an illustration of the simple point that F^* and F (and therefore S^* and S) are not just trivial variants of each other.

6. CONCLUDING OBSERVATIONS

In this note we have argued that the usual convention of assessing the 'femaleness' of a population in terms of a simple measure of central tendency could be inadequate and misleading. Accordingly, we have advanced alternative measures which 'correct' the conventional measures for the dispersion of age-specific sex-ratios around their mean value. This correction, broadly speaking, assumes the form of 'rewarding' distributions in which the sex-ratios at the upper age-groups are relatively large and 'penalizing' distributions in which the sex-ratios at the upper age-groups are relatively small. We have sought to provide empirical illustrations of our measurement concerns, with a view to providing examples of how the precise way in which we assess the 'femaleness' of a population can and does affect the outcome of our evaluative exercises.

APPENDIX
COMPUTING F^* FROM GROUPED DATA

We present below, analogously to the computation of the Gini coefficient of income-inequality through the 'trapezoidal approximation method', a method for computing the value of F^* from grouped data. (For a similar derivation of the Gini coefficient of inequality from grouped income-distribution data, see Kakwani (1980)). Specifically, suppose that there are K distinct age-groups $a_1, \ldots, a_K$ *which are indexed in descending order*, so that $a_1 > a_2 > \ldots a_K$. Let $p^m(a_i)$, or more compactly p_i^m, stand for the cumulative proportion of the male population in total population with ages exceeding the upper limit of a_i, $i=1,\ldots,K$. Similarly, we let p_i^f stand for the cumulative proportion of the female population in total population with ages exceeding the upper limit of a_i, $i=1,\ldots,K$. From grouped data which provide sex-wise frequencies of the population in each age-group, we can obtain a 'piece-wise linear' version of the A-curve, derived as a plot of the points $(p_1^m, p_1^f), \ldots, (p_K^m, p_K^f)$, (p_{K+1}^m, p_{K+1}^f) where $p_{K+1}^m := M$ (the male headcount ratio) and $p_{K+1}^f := F$ (the female headcount ratio). This is depicted in Figure A.1 for the special case in which K=4.

[Figure A.1 to be inserted here]

Under the trapezoidal approximation procedure, the value of F^* can be approximated by twice the sum of the areas A, B_1, B_2, C_1, C_2, D_1, D_2 and E in figure A.1. It is routine that

(A.1) Area $A = p_2^m . p_2^f/2$.

Area $B_1 = (p_3^m - p_2^m) . p_2^f$;

Area $B_2 = (p_3^m - p_2^m)(p_3^f - p_2^f)/2$; so

(A.2) $(\text{Area } B_1 + \text{Area } B_2) = (p_3^m - p_2^m)(p_2^f + p_3^f)/2$.

Area $C_1 = (p_4^m - p_3^m) . p_3^f$;

Area $C_2 = (p_4^m - p_3^m) . (p_4^f - p_3^f)/2$; so

(A.3) $(\text{Area } C_1 + \text{Area } C_2) = (p_4^m - p_3^m)(p_3^f + p_4^f)/2$.

Area $D_1 = (p_5^m - p_4^m) . p_4^f$;

Area $D_2 = (p_5^m - p_4^m) . (p_5^f - p_4^f)/2$; so

(A.4) $(\text{Area } D_1 + \text{Area } D_2) = (p_5^m - p_4^m)(p_4^f + p_5^f)/2$.

(A.5) Area $E[=(p_5^f)^2/2] = F^2/2$.

We now have from (A.1)-(A.5):

$$F^* \cong 2 \, [(\text{Area } A) + (\text{Area } B_1 + \text{Area } B_2) + (\text{Area } C_1 + \text{Area } C_2) + (\text{Area } D_1 + \text{Area } D_2) + (\text{Area } E)]$$

$$= 2.(1/2)[(p_2^m - p_1^m)(p_1^f + p_2^f) + (p_3^m - p_2^m)(p_2^f + p_3^f) + (p_4^m - p_3^m)(p_3^f + p_4^f) + (p_5^m - p_4^m)(p_4^f + p_5^f) + F^2]$$

$$= \sum_{i=2}^{5} (p_i^m - p_{i-1}^m)(p_i^f + p_{i-1}^f) + F^2.$$

In general, given $K \geq 2$ age-groups, the formula for the trapezoidal approximation to F^* is given by:

(A.6) $$F^* \cong \sum_{i=2}^{K+1} (p_i^m - p_{i-1}^m)(p_i^f + p_{i-1}^f) + F^2.$$

It is to be noted that F^* will understate the area under the ('true') A-curve over its strictly concave range, and overstate this area over its strictly convex range.

NOTES

1. See, for example, Sen (1985; Appendix B).

2. There are clear similarities between obtaining an A-curve for a discrete age-distributed sex profile and obtaining a Lorenz curve for a discrete income distribution; on the latter, see, for example, Anand (1983; Appendix B).

3. On the Lorenz curve see, among others, Lorenz (1905), Atkinson (1970) and Sen (1973). As far as the 'segregation curve' is concerned, the *locus classicus* is Duncan and Duncan (1955).

4. There is an analogy here with Sen's (1976) index of 'real national income', given by $W = \mu(1-G)$ where μ is *per capita* GNP and G is the Gini coefficient of inequality in the interpersonal distribution of incomes. Further, analogously to Atkinson's (1970) notion of 'the equally distributed equivalent income', F^* could be viewed in the light of an 'equally distributed equivalent female headcount ratio'.

REFERENCES

A nand, S. (1983): *Inequality and Poverty in Malaysia: Measurement and Decomposition.* New York: Oxford University Press.

Atkinson, A.B. (1970): 'On the Measurement of Inequality', *Journal of Economic Theory*, 2.

Duncan, O. and B. Duncan (1955): 'A Methodological Analysis of Segregation Indices', *American Sociological Review*, 20.

Kakwani, N.C. (1980): *Income Inequality and Poverty: Methods of Estimation and Policy Applications.* New York: Oxford University Press.

Lorenz, M.O. (1905): 'Methods for Measuring the Concentration of Wealth', *Journal of the American Statistical Association*, 9.

Sen, A.K. (1973). *On Economic Inequality.* Oxford: Clarendon Press.

Sen, A.K. (1976): 'Real National Income', *Review of Economic Studies*, 43.

Sen, A.K. (1985): *Commodities and Capabilities.* Amsterdam: North-Holland.

Table 5.1(a): The Age-wise Sex Composition of the Population in Kerala: 1991

Age group i (in descending order)	Number of Males in group i P^m_i	Number of Females in group i P^f_i	Number of Persons in group i P_i	Proportion in Total Population of Males in group i g^m_i (:= P^m_i/P)	Proportion in Total Population of Females in group i g^f_i (:= P^f_i/P)	Cumulative Proportion in Total Population of Males of age ≤ upperlimit of age-group i r^m_i (:= $\Sigma^{15}_{j=i} g^m_j$)	Cumulative Proportion in Total Population of Females of age ≤ upper limit of age-group i r^f_i (:= $\Sigma^{15}_{j=i} g^f_j$)	Cumulative Proportion in Total Population of Males of age > upper limit of age-group i p^m_i (:=M-r^m_i)	Cumulative Proportion in Total Population of Females of age > upper limit of age-group i p^f_i (:=M-r^f_i)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
70 & above	449258	547801	997059	0.0155	0.0189	0.4911	0.5089	0.0000	0.0000
65-69	324559	374175	698734	0.0112	0.0129	0.4441	0.4641	0.0470	0.0447
60-64	417045	454527	871572	0.0144	0:0157	0.3938	0.4150	0.0973	0.0938
55-59	465404	510493	975897	0.0160	0.0176	0.3399	0.3622	0.1512	0.1467
50-54	523257	538602	1061859	0.0180	0.0186	0.2894	0.3087	0.2018	0.2002
45-49	657954	678628	1336582	0.0227	0.0234	0.2385	0.2523	0.2526	0.2565
40-44	764030	724101	1488131	0.0263	0.0249	0.1950	0.2041	0.2961	0.3047
35-39	1013437	1036484	2049921	0.0349	0.0357	0.1590	0.1676	0.3322	0.3413
30-34	1046690	1060612	2107302	0.0361	0.0365	0.1241	0.1319	0.3671	0.3770
25-29	1262730	1398554	2661284	0.0435	0.0482	0.0977	0.1069	0.3934	0.4019
20-24	1475802	1636830	3112632	0.0508	0.0564	0.0751	0.0836	0.4161	0.4253
15-19	1467374	1551965	3019339	0.0506	0.0535	0.0571	0.0650	0.4341	0.4439
10-14	1564752	1534328	3099080	0.0539	0.0529	0.0410	0.0474	0.4501	0.4614
5-9	1459473	1425373	2884846	0.0503	0.0491	0.0267	0.0318	0.4645	0.4771
0-4	1365090	1298620	2663710	0.0470	0.0447	0.0155	0.0189	0.4757	0.4900
Total	14256855 (:=P^m)	14771093 (:=P^f)	29027948 (:=P)	0.4911 (:=M)	0.5089 (:=F)				

Source: Based on state-wise data (available on tape) in Census of India 1991 (Table C-5): Office of the Registrar General, Census of India.

Table 5.1(b): The Age-wise Sex Composition of the Population in Uttar Pradesh : 1991

Age-group i (in descending order)	Number of Males in group i P^m_i	Number of Females in group i P^f_i	Number of Persons in group i P_i	Proportion in Total Population of Males in group i g^m_i $(:=P^m_i/P)$	Proportion in Total Population of Females in group i g^f_i $(:=P^f_i/P)$	Cumulative Proportion in Total Population of Males of age ≤ upper limit of age-group i r^m_i $(:= \Sigma^{15}_{j=i}\, g^m_j)$	Cumulative Proportion in Total Population of Females of age ≤ upper limit of age-group i r^f_i $(:= \Sigma^{15}_{j=i}\, g^f_j)$	Cumulative Proportion in Total Population of Males of age > upper limit of age-group i p^m_i $(:=M-r^m_i)$	Cumulative Proportion in Total Population of Females of age > upper limit of age-group i p^f_i $(:=M-r^f_i)$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
70 & above	2060424	1542935	3603359	0.0149	0.0112	0.5317	0.4683	0.0000	0.0000
65-69	1100721	978938	2079659	0.0080	0.0071	0.5167	0.4572	0.0149	0.0112
60-64	2181123	1682802	3863925	0.0158	0.0122	0.5088	0.4501	0.0229	0.0183
55-59	1694107	1704238	3398345	0.0123	0.0123	0.4930	0.4379	0.0387	0.0304
50-54	2966928	2149689	5116617	0.0215	0.0156	0.4807	0.4256	0.0509	0.0428
45-49	2967067	2749666	5716733	0.0215	0.0199	0.4592	0.4100	0.0724	0.0583
40-44	3719737	3114270	6834007	0.0269	0.0225	0.4378	0.3901	0.0939	0.0782
35-39	4128962	3755950	7884912	0.0299	0.0272	0.4108	0.3676	0.1208	0.1008
30-34	4563374	4408328	8971702	0.0330	0.0319	0.3809	0.3404	0.1507	0.1280
25-29	5252576	4997476	10250052	0.0380	0.0362	0.3479	0.3084	0.1837	0.1599
20-24	5821885	5537923	11359808	0.0421	0.0401	0.3099	0.2723	0.2218	0.1961
15-19	7357772	5626551	12984323	0.0533	0.0407	0.2677	0.2322	0.2639	0.2362
10-14	9233769	7703847	16937616	0.0668	0.0558	0.2145	0.1915	0.3172	0.2769
5-9	10604306	9479126	20083432	0.0768	0.0686	0.1476	0.1357	0.3840	0.3327
0-4	9789525	9263528	19053053	0.0709	0.0671	0.0709	0.0671	0.4608	0.4013
Total	73442276 $(:=P^m)$	64695267 $(:=P^f)$	138137543 $(:=P)$	0.5317 $(:=M)$	0.4683 $(:=F)$				

Source: Based on state-wise data (available on tape) in Census of India 1991 (Table C-5): Office of the Registrar General, Census of India.

Table 5.1(c): Alternative Measures of the 'Femaleness' of the Population in Kerala and Utter Pradesh: 1991

State	Female Headcount Ratio F	'Corrected' Female Headcount Ratio F*	Sex-Ratio S	'Corrected' Sex-Ratio S*
(1)	(2)	(3)	(4)	(5)
Kerala	0.5089	0.5140	1036	1058
Uttar Pradesh	0.4683	0.4655	882	871

Source: Based on state-wise data (available on tape) in Census of India 1991 (Table C-5): Office of the Registrar General, Census of India.

Table 5.2(a): The Age-wise Sex Composition of the Population in West Bengal : 1991

Age-group i (in descending order)	Number of Males in group i P^m_i	Number of Females in group i P^f_i	Number of Persons in group i P_i	Proportion in Total Population of Males in group i g^m_i $(:= P^m_i/P)$	Proportion in Total Population of Females in group i g^f_i $(:= P^f_i/P)$	Cumulative Proportion in Total Population of Males of age ≤ upper limit of age-group i r^m_i $(:= \Sigma^{15}_{j=i} g^m_j)$	Cumulative Proportion in Total Population of Females of age ≤ upper limit of age-group i r^f_i $(:= \Sigma^{15}_{j=i} g^f_j)$	Cumulative Proportion in Total Population of Males of age > upper limit of age-group i p^m_i $(:= M - r^m_i)$	Cumulative Proportion in Total Population of Females of age > upper limit of age-group i p^f_i $(:= M - r^f_i)$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
70 & above	785913	747519	1533432	0.0116	0.0110	0.5214	0.4786	0.0000	0.0000
65-69	497832	476316	974148	0.0074	0.0070	0.5098	0.4676	0.0116	0.0110
60-64	822956	785037	1607993	0.0122	0.0116	0.5024	0.4605	0.0190	0.0181
55-59	943881	801836	1745717	0.0139	0.0118	0.4903	0.4489	0.0311	0.0297
50-54	1297233	1101358	2398591	0.0192	0.0163	0.4763	0.4371	0.0451	0.0415
45-49	1607110	1307493	2914603	0.0237	0.0193	0.4572	0.4208	0.0642	0.0578
40-44	1835614	1487530	3323144	0.0271	0.0220	0.4334	0.4015	0.0880	0.0771
35-39	2469467	1989018	4458485	0.0365	0.0294	0.4063	0.3795	0.1151	0.0991
30-34	2700576	2342186	5042762	0.0399	0.0346	0.3698	0.3501	0.1516	0.1285
25-29	3089590	3050477	6140067	0.0457	0.0451	0.3299	0.3155	0.1915	0.1631
20-24	3214747	3092250	6306997	0.0475	0.0457	0.2842	0.2704	0.2372	0.2082
15-19	3322584	2984759	6307343	0.0491	0.0441	0.2367	0.2247	0.2847	0.2539
10-14	4140707	3929596	8070303	0.0612	0.0581	0.1876	0.1806	0.3337	0.2980
5-9	4594479	4440741	9035220	0.0679	0.0656	0.1265	0.1226	0.3949	0.3560
0-4	3964709	3855325	7820034	0.0586	0.0570	0.0586	0.0570	0.4628	0.4216
Total	35287398 $(:= P^m)$	32391441 $(:= P^f)$	67678839 $(:= P)$	0.5214 $(:= M)$	0.4786 $(:= F)$				

Source: Based on state-wise data (available on tape) in Census of India 1991 (Table C-5): Office of the Registrar General, Census of India.

Table 5.2(b): The Age-wise Sex Composition of the Population in Rajasthan: 1991

Age-group i (in descending order)	Number of Males in group i P^m_i	Number of Females in group i P^f_i	Number of Persons in group i P_i	Proportion in Total Population of Males in group i g^m_i $(:=P^m_i/P)$	Proportion in Total Population of Females in group i g^f_i $(:=P^f_i/P)$	Cumulative Proportion in Total Population of Males of age ≤ upper limit of age-group i r^m_i $(:=\Sigma^{15}_{j=i} g^m_j)$	Cumulative Proportion in Total Population of Females of age ≤ upper limit of age-group i r^f_i $(:=\Sigma^{15}_{j=i} g^f_j)$	Cumulative Proportion in Total Population of Males of age > upper limit of age-group i p^m_i $(:=M-r^m_i)$	Cumulative Proportion in Total Population of Females of age > upper limit of age-group i p^f_i $(:=M-r^f_i)$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
70 & above	485749	497578	983327	0.0111	0.0113	0.5233	0.4767	0.0000	0.0000
65-69	291517	309199	600716	0.0066	0.0070	0.5122	0.4654	0.0111	0.0113
60-64	623895	559932	1183827	0.0142	0.0128	0.5056	0.4583	0.0177	0.0184
55-59	498326	524742	1023068	0.0114	0.0120	0.4913	0.4456	0.0319	0.0312
50-54	876134	692389	1568523	0.0200	0.0158	0.4800	0.4336	0.0433	0.0431
45-49	874641	847813	1722454	0.0199	0.0193	0.4600	0.4178	0.0633	0.0589
40-44	1122952	953870	2076822	0.0256	0.0217	0.4401	0.3985	0.0832	0.0782
35-39	1329614	1170893	2500507	0.0303	0.0267	0.4145	0.3768	0.1088	0.1000
30-34	1537858	1433959	2971817	0.0351	0.0327	0.3842	0.3501	0.1391	0.1267
25-29	1769595	1703040	3472635	0.0403	0.0388	0.3491	0.3174	0.1742	0.1593
20-24	1871346	1836442	3707788	0.0427	0.0419	0.3088	0.2786	0.2145	0.1982
15-19	2252584	1856895	4109479	0.0513	0.0423	0.2661	0.2367	0.2572	0.2400
10-14	2963281	2597482	5560763	0.0675	0.0592	0.2148	0.1944	0.3085	0.2824
5-9	3381853	3049705	6431558	0.0771	0.0695	0.1472	0.1352	0.3761	0.3416
0-4	3076375	2879852	5956227	0.0701	0.0656	0.0701	0.0656	0.4531	0.4111
Total	22955720 $(:=P^m)$	20913791 $(:=P^f)$	43869511 $(:=P)$	0.5233 $(:=M)$	0.4767 $(:=F)$				

Source: Based on state-wise data (available on tape) in Census of India 1991 (Table C-5): Office of the Registrar General, Census of India.

Table 5.2 (c): Alternative Measures of the 'Femaleness' of the Population in West Bengal and Rajasthan : 1991

State	Female Headcount Ratio F	'Corrected' Female Headcount Ratio F*	Sex-Ratio S	'Corrected' Sex-Ratio S*
(1)	(2)	(3)	(4)	(5)
West Bengal	0.4786	0.4724	918	895
Rajasthan	0.4767	0.4777	911	913

Source: Based on state-wise data (available on tape) in Census of India 1991 (Table C-5): Office of the Registrar General, Census of India.

Table 5.3: Alternative Measures of the 'Femaleness' of the Population For the States of the Indian Union: 1991

State	Female Headcount Ratio F	'Corrected' Female Headcount Ratio F*	Sex Ratio S	'Corrected' Sex Ratio S*	Rank (in descending order) of state according to S	Rank (in descending order) of state according to S*
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Andhra Pradesh	0.4931	0.4931	973	973	3	2
Bihar	0.4774	0.4748	914	904	10	10
Gujarat	0.4832	0.4846	935	940	6	6
Karnataka	0.4899	0.4883	960	954	5	5
Kerala	0.5089	0.5140	1036	1058	1	1
Madhya Pradesh	0.4827	0.4818	933	930	8	8
Maharashtra	0.4829	0.4836	934	936	7	7
Orissa	0.4929	0.4907	972	956	4	4
Punjab	0.4687	0.4682	882	881	12	12
Rajasthan	0.4767	0.4777	911	913	11	9
Tamilnadu	0.4935	0.4919	974	968	2	3
Uttar Pradesh	0.4683	0.4655	881	871	13	13
West Bengal	0.4786	0.4724	918	895	9	11

Source: Based on state-wise data (available on tape) in Census of India 1991 (Table C-5): Office of the Registrar General, Census of India.

Table 5.4(a): The Age-Specific Sex Composition of the Indian Population at the Censuses of 1961, 1971, 1981 and 1991

Age-grouping (in descending order)	1961		1971		1981		1991	
	Number of Males in group i	Number of Females in group i	Number of Males in group i	Number of Females in group i	Number of Males in group i	Number of Females in group i	Number of Males in group i	Number of Females in group i
(1)	(2)	(3)	(2)	(3)	(2)	(3)	(2)	(3)
70 & above	4169203	4432359	5745232	5579218	7843217	7642187	10962858	10111307
65-69	2468273	2373156	3644372	3356877	4793728	4720693	6493630	6364869
60-64	5698649	5519641	7484721	6889311	9385925	8781636	11907237	10841739
55-59	5274399	4538314	6876424	5951965	8498074	7918507	10941747	10530755
50-54	9118155	7963261	11115887	9415037	13794219	11602685	16904890	14208702
45-49	9722590	8307088	12467510	10417273	15372729	13866000	18954561	17179237
40-44	12068944	10754185	15058117	13229867	18033269	16154363	22842245	19714094
35-39	13584549	11841585	17236348	15661954	19899070	18959698	27558300	24840570
30-34	15963939	14832236	18321341	17867076	21579342	20800402	29917765	28486719
25-29	18503600	18024631	20339371	20481079	25754744	24969871	34546587	34692671
20-24	18164552	19105755	21573419	21527935	29000075	28337783	37514223	36958481
15-19	18567957	17255857	25221778	22246454	34026988	30111820	42231074	36803855
10-14	26234676	22994194	36493304	32274530	45265292	40646075	51947630	46744268
5-9	33031120	31557696	42211297	39796175	48267570	45418306	57419164	53875568
0-4	33147895	32876926	40203916	39355600	42227766	41282041	52360652	50017380
Total	225718501	212376884	283993037	264050351	343742008	321212067	432502563	401370215

Source: (1) Census of India, 1961: Age Tables (Paper No.2 of 1963).

(2) Census of India, 1971: Age Tables (Paper No.3 of 1977).

(3) Census of India, 1981: Social and Cultural Tables (Table C-5 Single Year Age Returns).

(4) Census of India, 1991: Table C-5 (available on tape): Office of the Registrar General, Census of India.

Table 5.4 (b): Alternative Measures of the 'Femaleness' of the Indian Population at the Censuses of 1961,1971,1981 and 1991

Year	Female Headcount Ratio F	'Corrected' Female Headcount Ratio F*	Sex-ratio S	'Corrected' Sex-Ratio S*
(1)	(2)	(3)	(4)	(5)
1961	0.4848	0.4817	941	929
1971	0.4818	0.4789	930	919
1981	0.4831	0.4818	935	930
1991	0.4813	0.4800	928	923

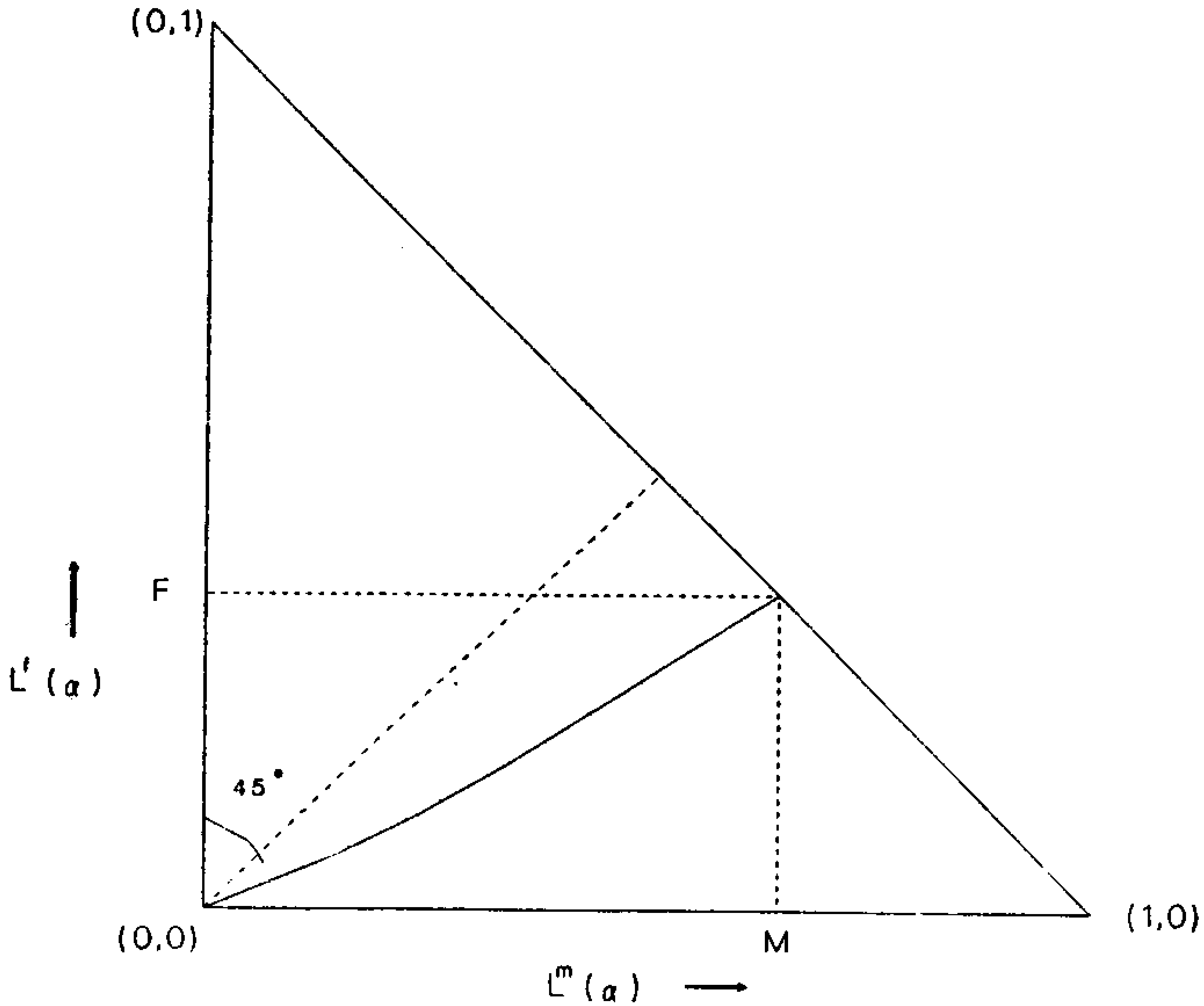
Source: Computed from data in Table 5.4 (a).

Table 5.4(c): Ranking of the Censal Years 1961-1991 in terms of Vintage and in terms of the 'Femaleness' of the Indian Population

Censal Year	Ranking of year in terms of vintage (in descending order)	Ranking of year in terms of the sex- ratio S (in descending order)	Ranking of year in terms of the 'corrected' sex-ratio S* (in descending order)
(1)	(2)	(3)	(4)
1961	1	1	2
1971	2	3	4
1981	3	2	1
1991	4	4	3

Source: Based on data in Table 5.4 (b).

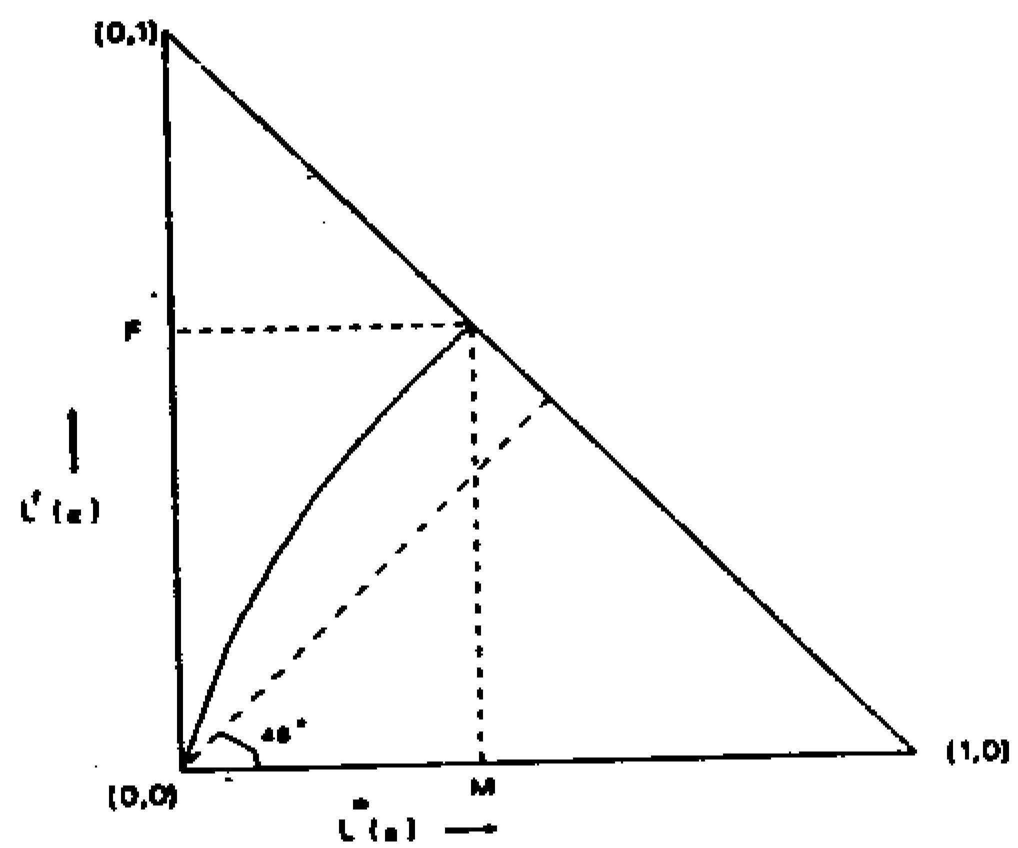
Figure 3.1: A Possible A-Curve



The A-curve is a non-decreasing graph going from the origin to some point on the hypotenuse of the right-angled triangle.

Figure 3.2: Two Possible A-Curves

a) The A - Curve displays unambiguous female advantage



b) The A - Curve displays unambiguous female disadvantage

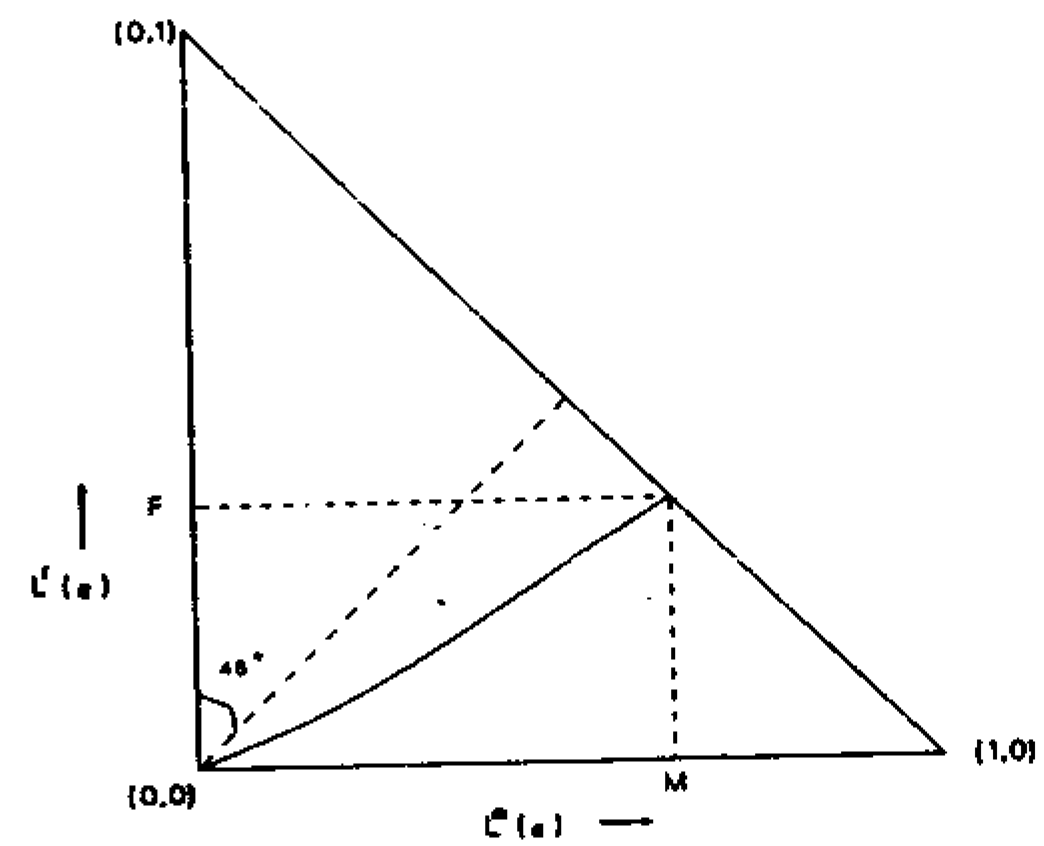
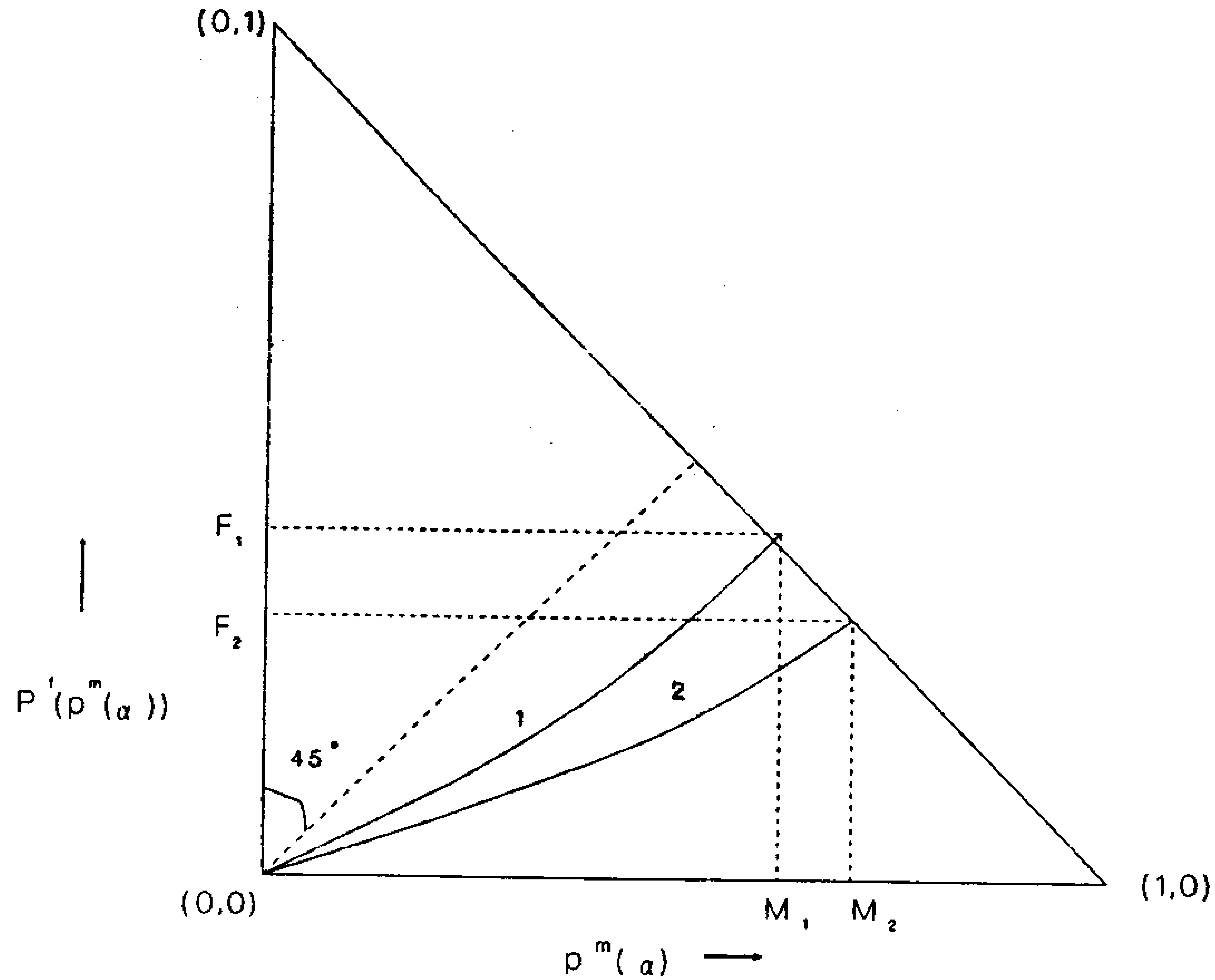
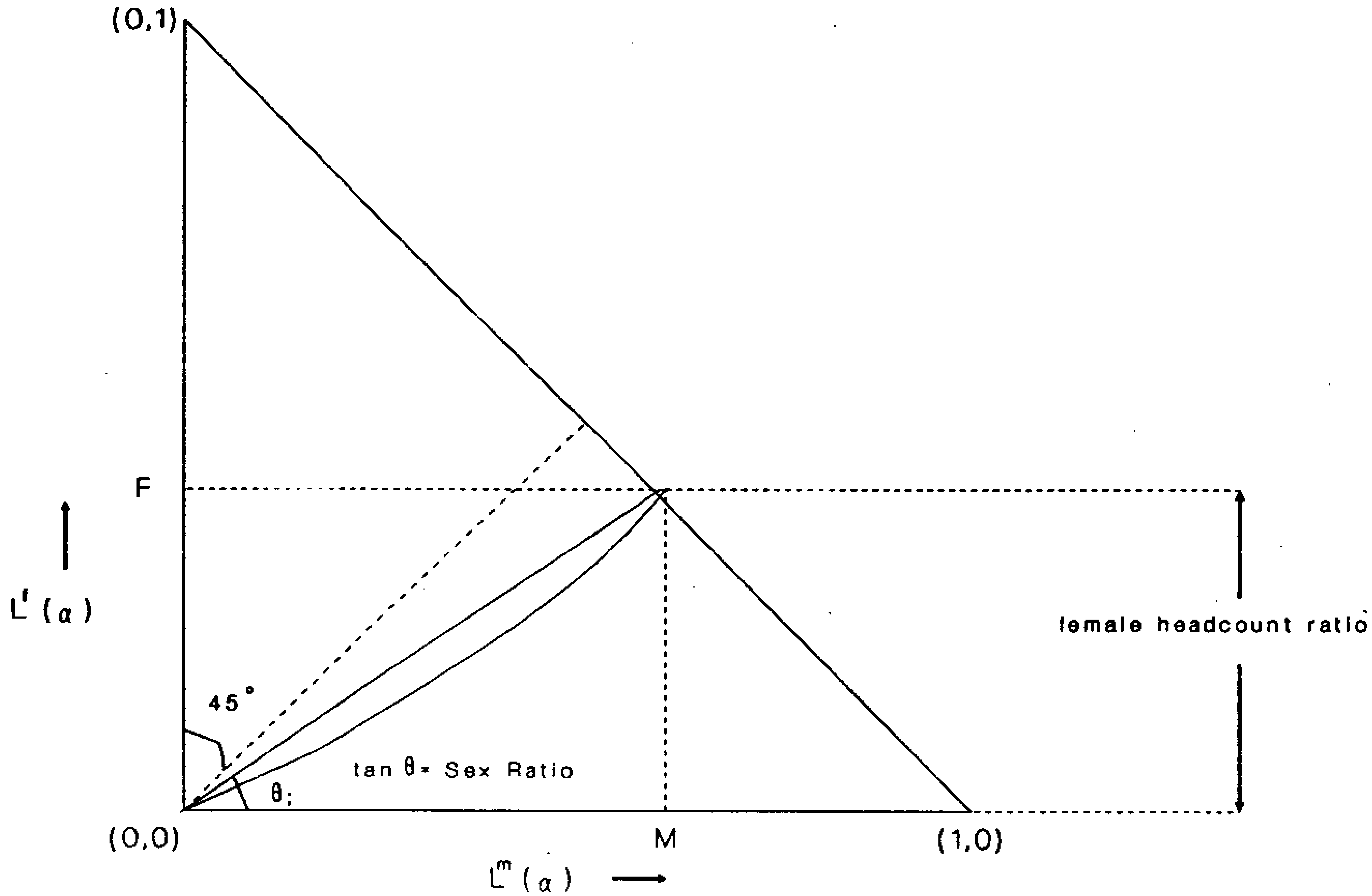


Figure 3.3: A Case of A-Dominance



The A-Curve for population 1 (strictly) dominates the A-Curve for population 2

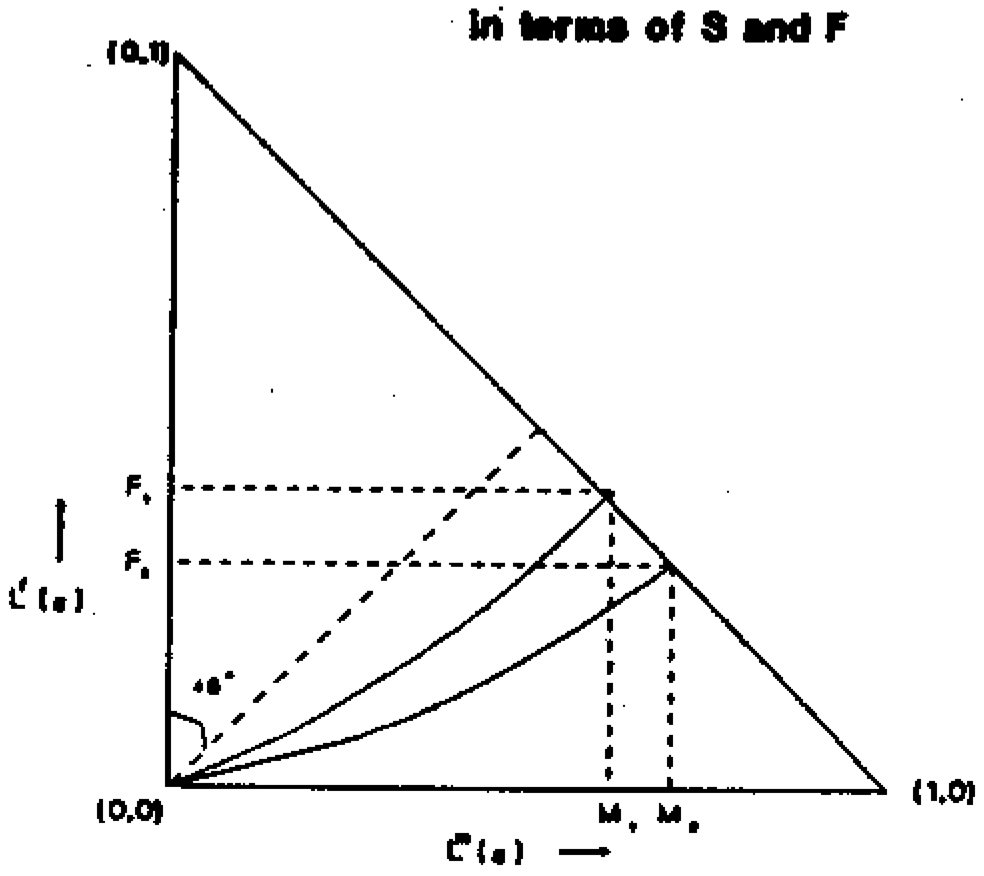
Figure 3.4: The A - Curve and the Indices S and F



The sex-ratio is the slope of the straight line connecting the origin to the point (M,F) on the hypotenuse of the right-angled triangle. The female headcount ratio is the height of the point (M,F).

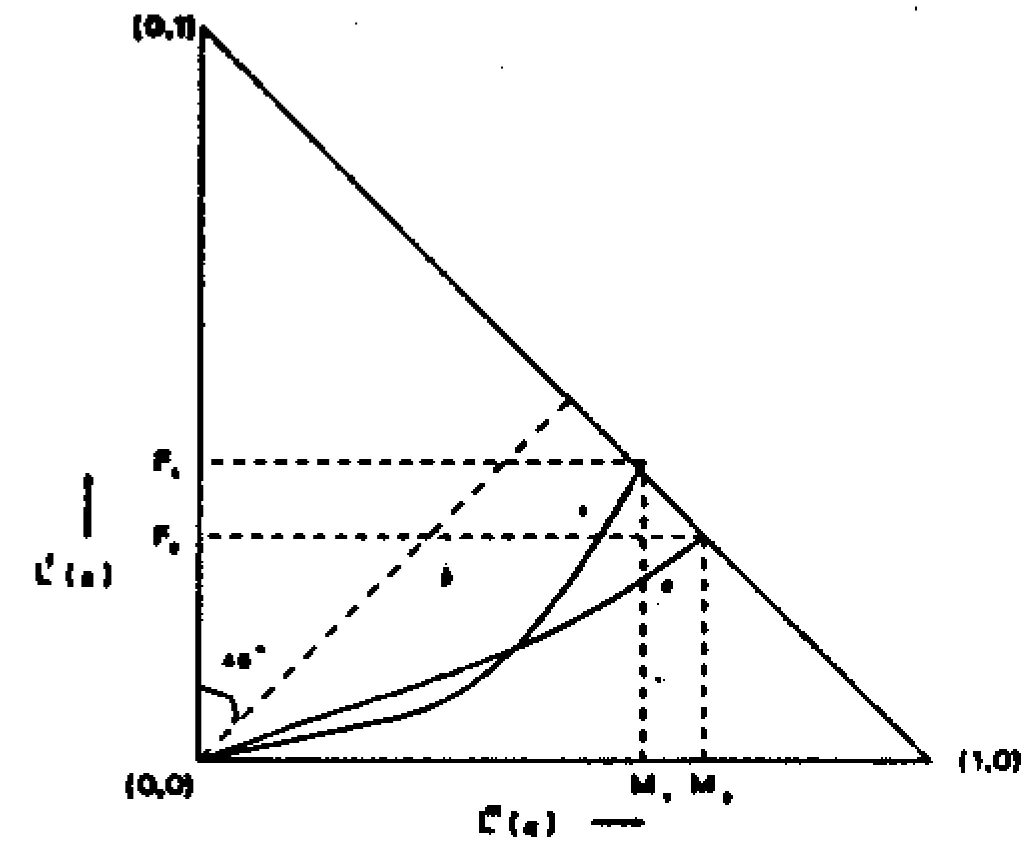
Figure 3.5: A-Dominance and Dominance in terms of the Indices F and S

Figure 3.5 (a): Super-dominance of A implies Dominance in terms of S and F



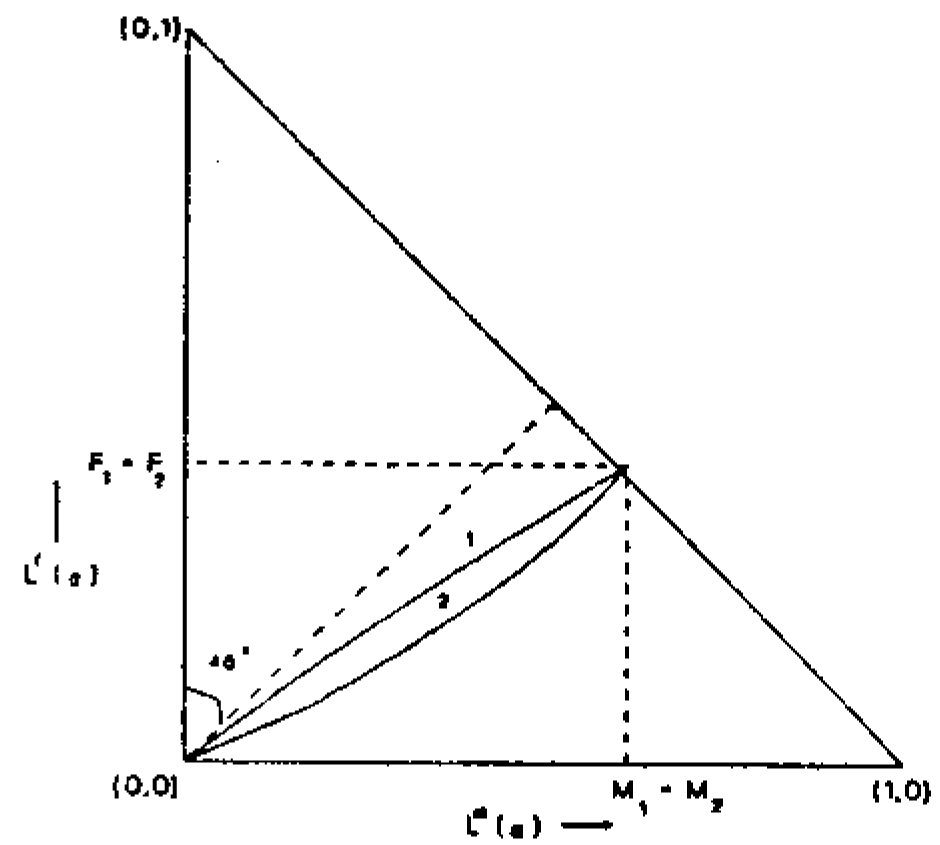
A1 ≻≻ A2 ; F1 > F2 & S1 > S2

Figure 3.5(b): Dominance in Terms of S and F does not imply A-Dominance



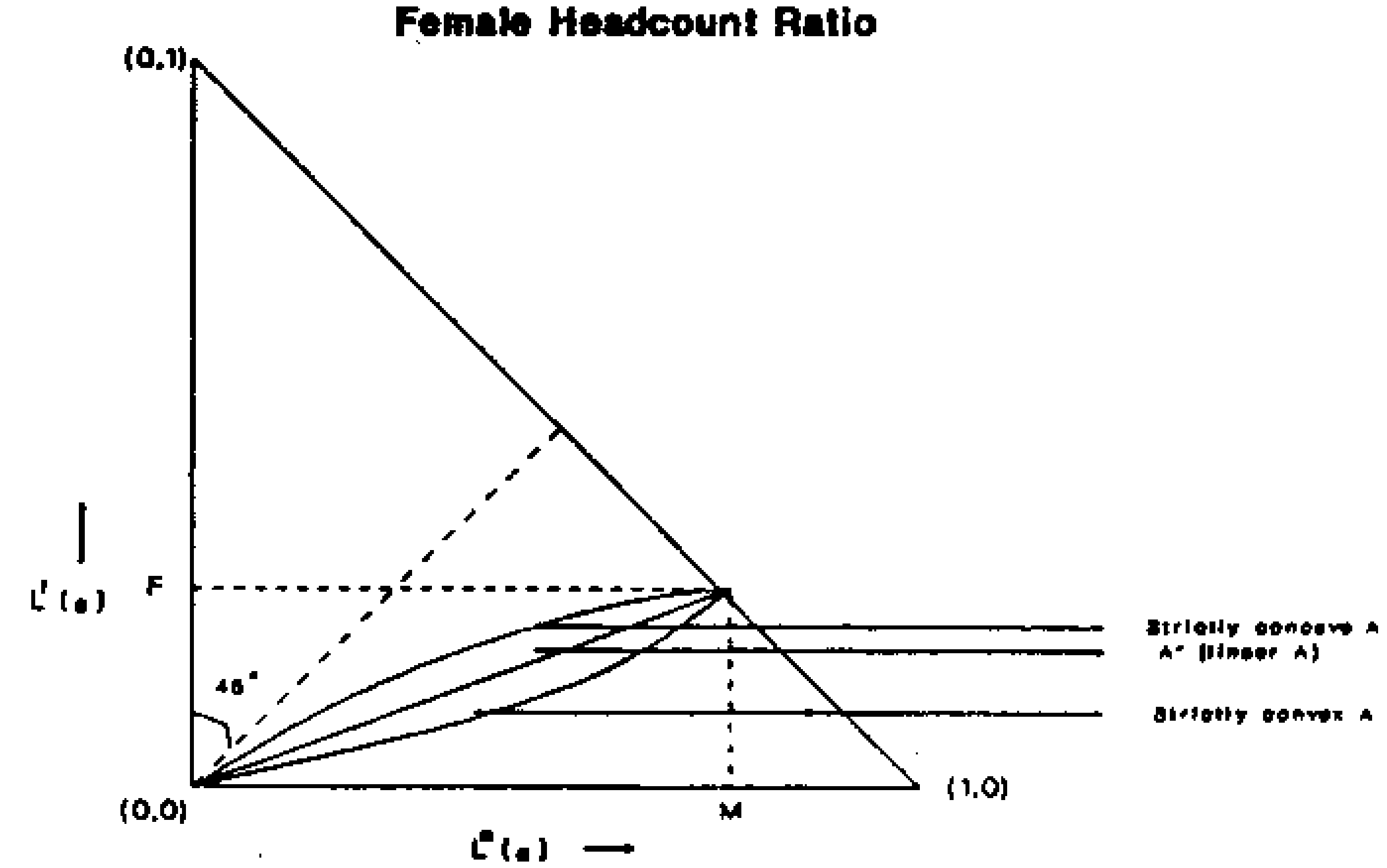
F1 > F2 & S1 > S2 ; but ~ (A1 ≻ A2)

Figure 3.5(c): Equality in Terms of S and F, But Dominance in terms of the A-Curve



F1=F2;S1=S2: and A1≻ A2

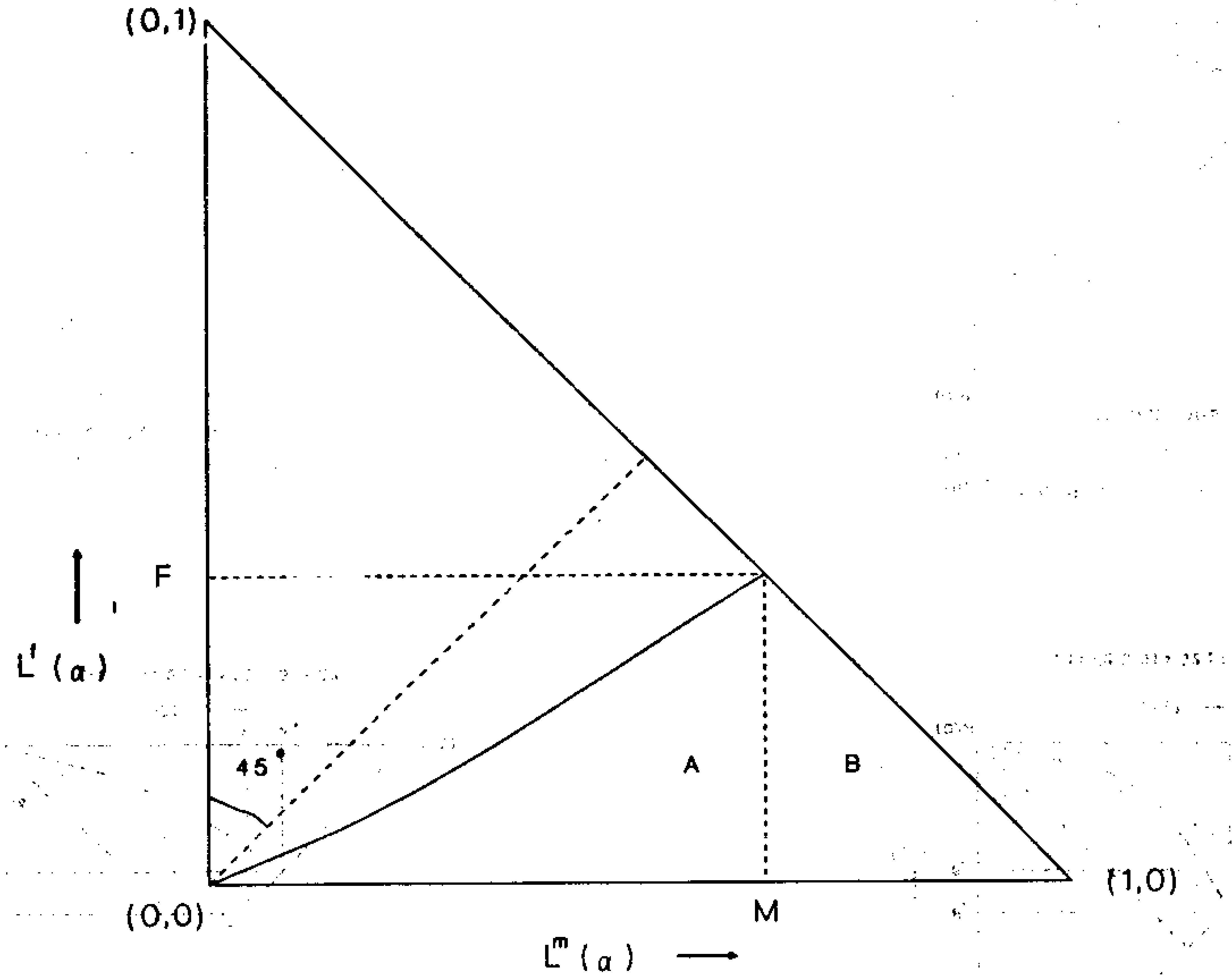
Figure 3.5(d): Three Possible Age-Distributed Sex Composition Curves for a given level of the Female Headcount Ratio



A≻ A* if A is strictly concave

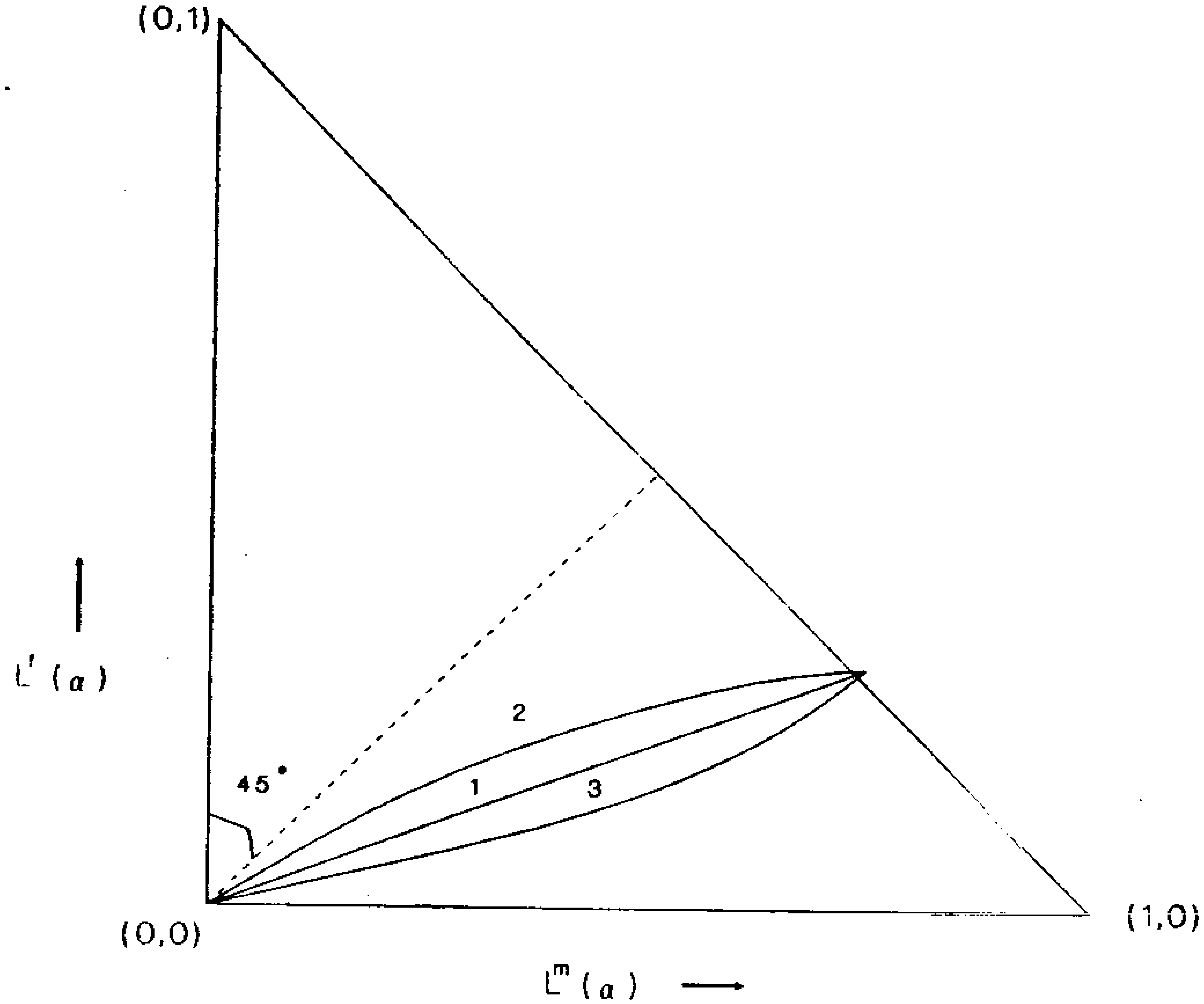
A*≻ A if A is strictly convex

Figure 4.1: The A-Curve and the Index F^*



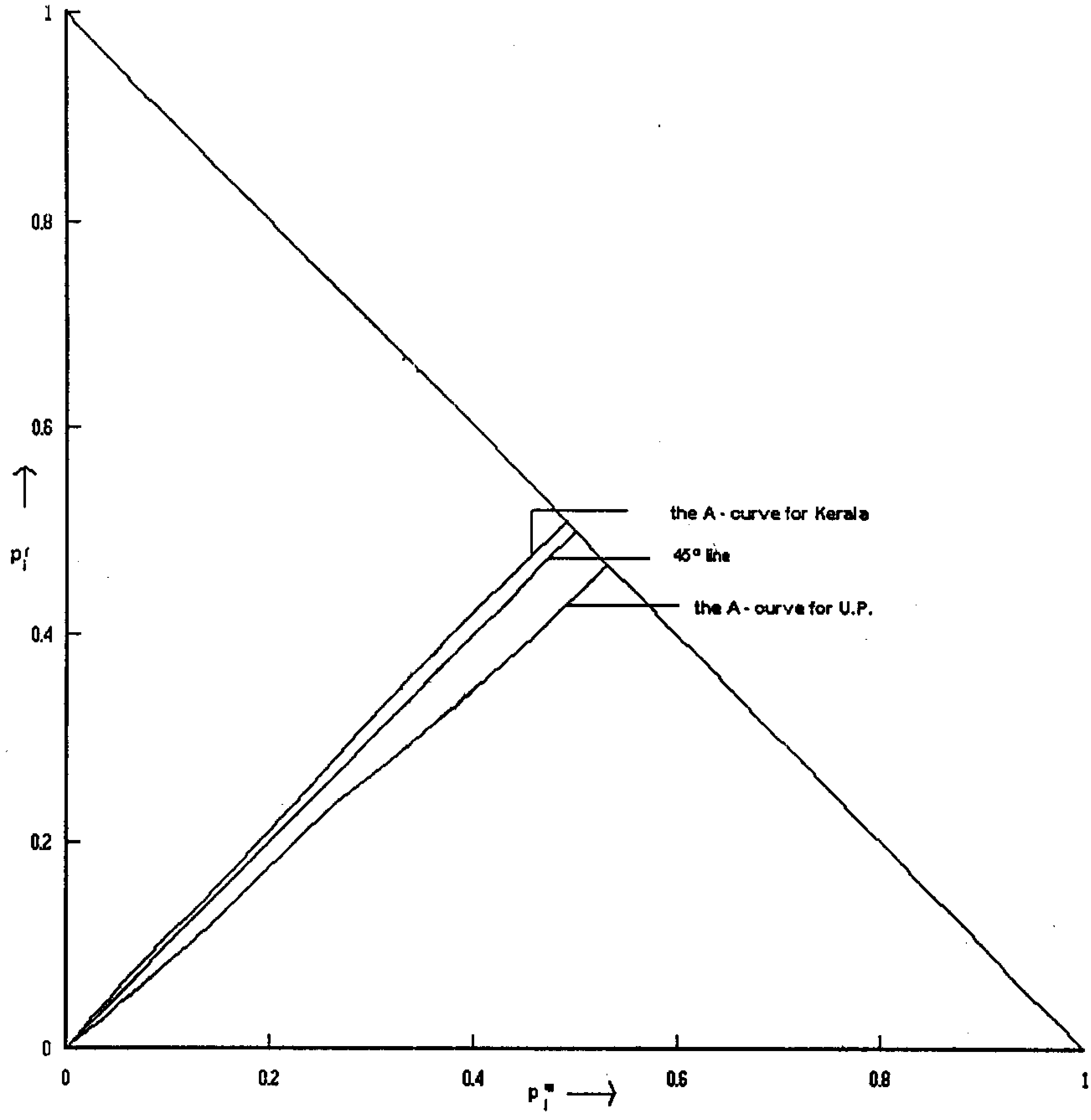
The Index F^* is given by twice the area to the right of the A-curve i.e. by twice the sum of the areas A and B

Figure 4.2: The Shape of the A-Curve and the Value of F* in relation to F



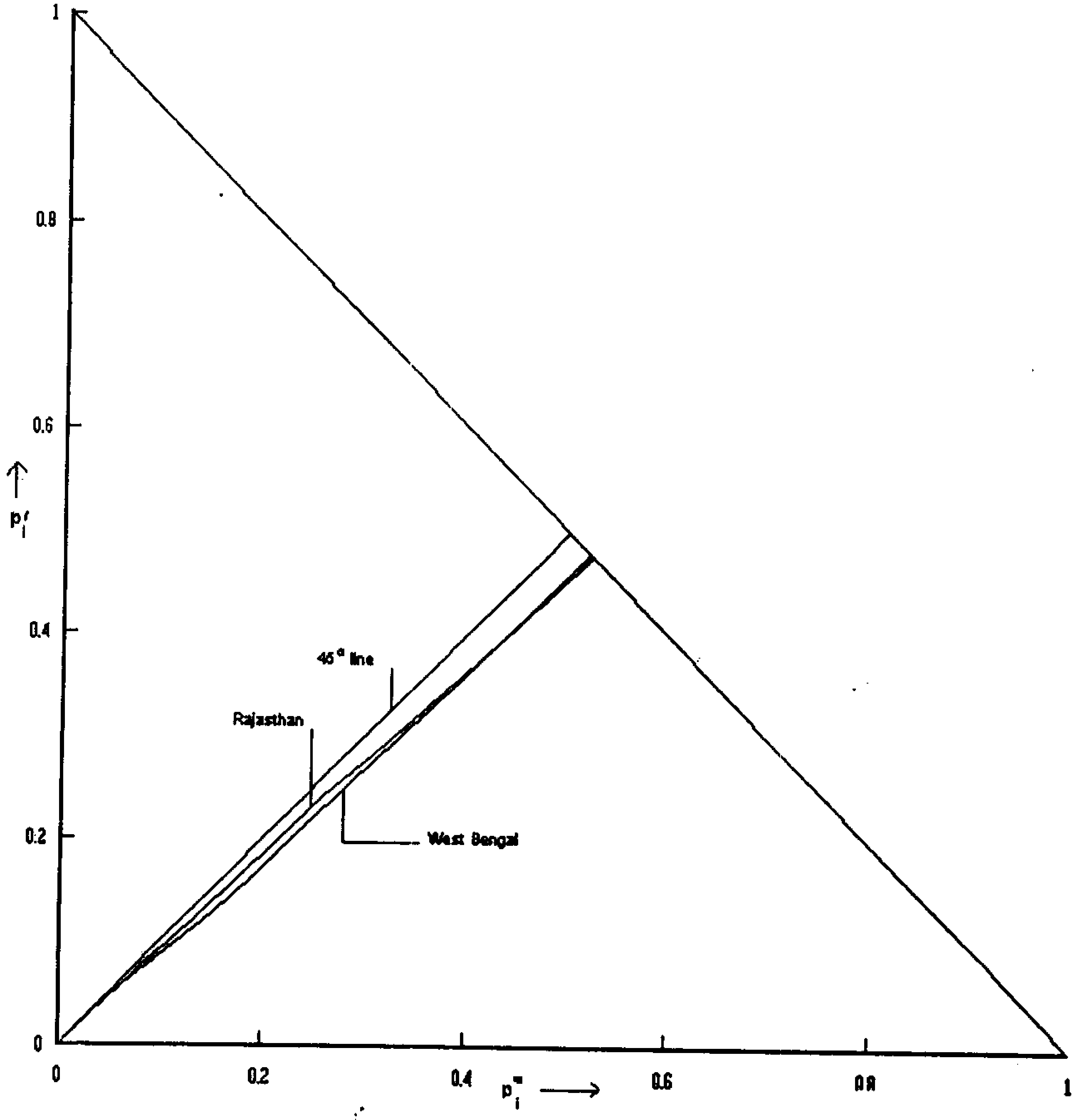
Three societies 1,2 and 3 have the same overall female headcount ratio F. The A-Curve is linear for society 1, strictly concave for society 2 and strictly convex for society 3; and F* = F for society 1, F* > F for society 2 and F* < F for society 3.

Figure 5.1: The Age-distributed Gender Composition Curves for the states of Kerala and Uttar Pradesh: 1991



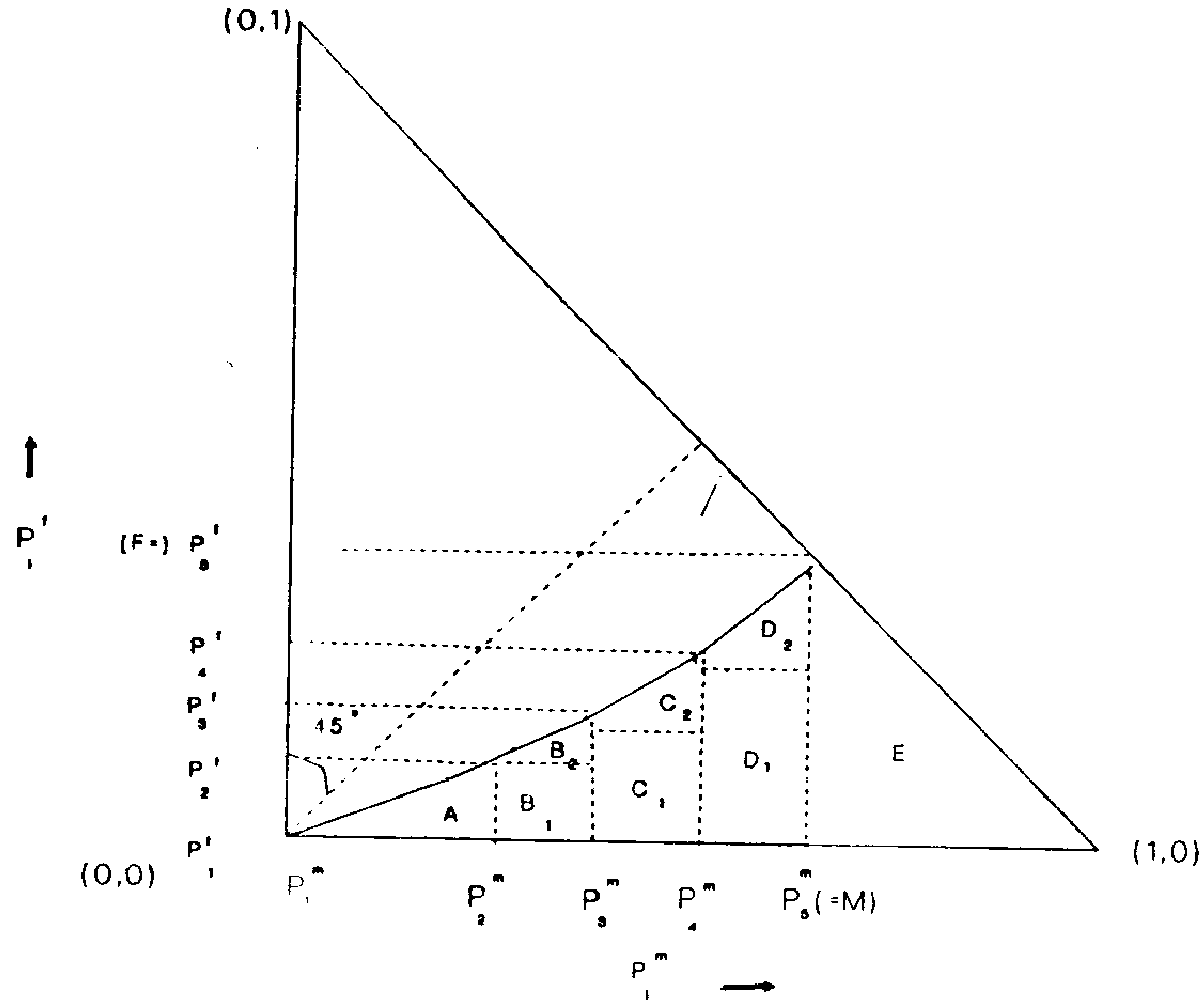
The 'Piece-wise linear' A-curves for Kerala and U.P. have been generated from the coordinates of the two A-curves, data on which are available in columns 9 and 10 of Tables 6.1(a) and 6.1(b). Notice that here we have an example of A-dominance (of Kerala over U.P.); further, we also have an example of 'unambiguous female advantage' (Kerala) and one of 'unambiguous female disadvantage' (U.P.) in terms of the location of the respective A-curves in relation to the 45° line.

Figure 5.2: The Age-distributed Gender Composition Curves for the States of West Bengal and Rajasthan: 1991



The 'Piece-wise linear' A-curves for West Bengal and Rajasthan has been generated from the coordinates of the two A-curves, data on which are available in columns 9 and 10 of Tables 6.2(a) and 6.2(b). We have here an instance of intersecting A-curves. Although the F-value for West Bengal is higher than that for Rajasthan, the raking is inverted by F*: notice that there is a range of ages in the upper age-groups over which the A-curve for Rajasthan dominates that for West Bengal.

Figure A.1: A Piece-wise Linear A-Curve from Grouped Data



The 'trapezoidal approximation' of the index F* is given by twice the sum of the areas A, B_1, B_2, C_1, C_2, D_1, D_2, and E