Two Issues Involved in Measuring the Concentration of Farmland Ownership

By

G. Andrew Bernat, Jr.
U.S. Department of Agriculture,
ERS/NRED/Land Economics Branch

Paper Presented at the 1986 Annual Meeting of the American Agricultural Economics Association

July 27-30, 1986
Reno, Nevada

1986

DRAFT: Do Not Quote

Two Issues Involved in Measuring the Concentration of Farmland Ownership

I. Introduction

Underlying much of U.S. agricultural policy since the nation's founding is a belief in the importance of farmers owning the land they cultivate. For this reason, the fact that, up to the Second World War, a large number of farmers did not own any of the land they operated was considered to be a serious problem. In the postwar era, American agriculture has undergone a tremendous amount of structural change, a salient feature of which is a progressive decline in the rate of tenancy (defined as the proportion of farmers who operate only rented land) and a dramatic increase in the importance of the part-owner (defined as a farmer who operates both rented and owned land). As a result, nearly 90% of all farmers own at least some of the land they operate. However, the fact that landownership among farm operators is now relatively widespread does not mean that the pattern of farmland ownership is no longer an important issue because land ownership concentratation may still be considered too high even when every farm operator owns some portion of the land he or she operates.

Before any progress can be made in discovering the causes and consequences of landownership concentration, it is necessary that satisfactory ways of measuring concentration be found. The purpose of this paper is to discuss two problems that are encountered in attempts to measure the concentration of landownership. The first problem is the choice of concentration or inequality index. This is a problem that is faced in the analysis of any type of economic inequality. The second problem, deciding on how to calculate the chosen index (or indices), arises only when it is necessary to use grouped data. Thus, while not inherent to the problem of inequality measurement, this is an issue researchers face whenever secondary sources are utilized.

The paper is organized into five sections. In the section following this introduction we discuss the problem of choosing an index to measure concentration. We also present the four indices that are used in subsequent analyses. In the third section we discuss the problem that arises whenever grouped data is used to calculate concentration indices. In section IV, concentration indices are calculated to assess the empirical importance of the issues raised in the preceding two sections. The fifth and final section summarizes the results and discusses some of the implications of the analysis.

II. Choice of Concentration Indices.

Concentration indices can be thought of as measures of the distance between a given distribution and a reference distribution. For every index, the measured distance, or inequality, is a function of the difference between the corresponding elements of the distribution of interest and the reference distribution. The way in which each individual element of a distribution is weighted is really a definition of inequality. Using the uniform distribution as the reference distribution, each index measures how "far" a given distribution is from complete equality. The standard practice is to use indices normalized to range between 0 and 1. A value of 0 indicates the minimum distance, viz., the distribution coincides with the uniform distribution, whereas a value of 1 indicates that the distribution in question is as far away from the uniform distribution as is possible.

Two types of comparisons can be made with concentration indices, ordinal and cardinal. Ordinal comparsions can tell us whether one distribution (distribution A) is more or less concentrated than another (distribution B) whereas a cardinal comparison purports to tell us how much more or less concentrated is distribution A than distribution B. Thus, if we are interested in the ranking of distributions, that is, if we are interested in whether landownership is more concentrated today than twenty years ago, all we need is an ordinal comparison. If, on the other hand, we want to know how much the concentration of landownership today differs from twenty years ago, we need to make cardinal comparisons.

A common theme in much of the literature on the measurement of inequality is that every inequality index has a corresponding social welfare function in the sense that the index and the social welfare function would always rank any set of distributions in identical order (Atkinson, Blackorby and Donaldson). To see this correspondence, note that both a social welfare function and an inequaltiy index consist of weighted combinations of attributes of individuals in the population. A social welfare function, as commonly defined, is a weighted sum of the utility of the individuals in society. The attribute is thus utility. The attribute that is combined to create an inequality index of land ownership is land owned by individual ownership units. Thus, if an index and a social welfare function employ the same weighting schemes, they will rank all distributions in the same order.

It is therefore desirable to look closely at the properties of any prospective concentration index so that indices representing unacceptable definitions of inequality can be ruled out. Before discussing how indices differ with regard to weighting, we turn to a brief discription of the four indices examined here.

The Gini index. The Gini index (G) is one of the most popular inequality indices. A primary reason for the Gini's popularity is that it can be graphically represented because of its relationship to the Lorenz curve. Mathematically, the Gini coefficient can be calculated from equation (1).

$$(1) \quad G = \frac{(N + 1)}{N} - \frac{2(\Sigma a_i y_i)}{TY \cdot N}$$

where y_i is the landholding of the i^{th} ownership unit, N is the number of ownership units, TY is total landholdings of the group in question, and $a_i = N + 1 - i$, the position in the distribution when all ownership units are ranked such that $y_1 < y_2 < \ldots < y_n$.

Thiel's Index. Thiel's index (T) is derived from his entropy index, the general formula of which is

$$(2)\quad \max H(x) = x\ln(1/x).$$

Defining x_i as the share of land owned by the i^{th} ownership unit, equation (2) is equal to 0 when a single unit owns all the land, and is equal to ln(N) when landownership is divided equally among all N units. Rearranging and normalizing (2) so that the index equals 1 at maximum inequality and 0 at equality gives us the form we will use in this paper.

$$(3)\quad T = \frac{\Sigma(y_i/\mu)\ \ln(y_i/\mu)}{\ln(N)}$$

where μ is the average land holding.

Atkinson's Index. The third index we will examine was proposed by Atkinson (Atkinson). As pointed out above, it can be shown that each inequality index has a corresponding social utility function. Atkinson uses this correspondence to show how the problem of ranking income, wealth, or, in the present case, landownership distributions is quite similar to the well-studied problem of ranking probability distributions that appears in the literature on decision making under uncertainty. This similarity leads him to an index that incorporates the concept of inequality aversion, a concept closely resembling the concept of risk aversion. Thus, lower levels of ownership concentration would be preferred, the greater is the level of inequality aversion. This index is presented in equation 4.

$$(4)\quad A = 1 - \Sigma\left[\frac{(y_i/\mu)^{(1-\varepsilon)}}{N}\right]^{1/(1-\varepsilon)}$$

The parameter ϵ represents the degree of inequality aversion. The greater the value of ϵ, the more sensitive is the index to the level of concentration. For example, at $\epsilon = \infty$, the index is sensitive only to the landholdings of the lowest ownership unit. This is consistent with a Rawlsian maxi-min social welfare function in which a change is evaluated with regard solely to its effect on the lowest member of the distribution. At the other extreme, $\epsilon = 0$, the index is completely insensitive to the degree of concentration. All that matters in this case is the total amount of land owned.

Coefficient of Variation. The fourth index is the coefficient of variation (C, equation 5).

$$(5) \quad C = \left[\frac{\Sigma\,(y_i - \mu)^2}{N \cdot \mu^2} \right]^{1/2}$$

Used relatively infrequently as an inequality index, it is included to give some indication of how the distributions compared in this paper differ with regard to what might be called statistical variation.

III. Two Properties of Inequality Indices.

In this section, we compare the four indices with regard to (1) what we will call the elasticity of concentration, and (2), how measured concentration changes in response to a small transfer from one individual to another. These comparsions give us important insights regarding the way in which each index weights different landholding units and thereby gives us a better understanding of the underlying definition of inequality represented by each index.

The elasticity of concentration is defined as the percent change in measured concentration due to a percent change (small) in the landholdings of a single individual, y_i. Equations (6) - (9) present the elasticities for each of the five indices.

$$(6) \quad \pi_G = \left[\frac{Y_i}{TY} \right] \left[\frac{2i - N - 1}{(N \cdot G)} - 1 \right]$$

$$(7) \quad \pi_T = \left[\frac{y_i}{TY} \right] \left[\frac{N \ln (y_i/\mu)}{T \cdot \ln(N)} - 1 \right]$$

$$(8) \quad \pi_A = \left[\frac{y_i}{TY \cdot A \cdot N} \right] \left[N(1-A)^{(1-\varepsilon^2)} - (1-A)^{\varepsilon(1-\varepsilon)} \cdot (\mu/y_i)^{\varepsilon} \right]$$

$$(9) \quad \pi_C = \left[\frac{y_i}{TY} \right] \left[\frac{y_i - \mu}{\mu^2 C^2} - \mu \right]$$

In all cases, the percent change in measured concentration resulting from a percentage change in the landholdings of the i^{th} landownership unit is a function of the individual's share of total land owned. In addition, the initial level of concentration affects the size of this change. However, the precise way in which the initial level of concentration appears in each of the above elasticities is an important difference between the indices. In the expressions for both the Gini and the Thiel (equations 6 and 7), the initial level of inequality appears in the denominator. This means that the greater the initial level of inequality, the smaller will be the effect on concentration of a given percentage change in size of the i^{th} land-holding. Initial inequality also appears in the denominator of the coefficient of variation (equation 9). However, unlike the Gini and the Thiel, initial concentration is squared in equation 9. This means that the coefficient of variation gives greater weight than either the Gini or the Thiel to changes in y_i, the lower is initial inequality. The Atkinson index differs from the other indices in that the elasticity of concentration is a function of the parameter as well as of the initial level of concentration.

Another important way to view concentration indices is to derive expressions for the change in concentration due to a small transfer from one ownership unit to another. Equations (10) - (13) tell us, for each index, how measured concentration of landownership would change due to a marginal transfer from ownership unit i to ownership unit j, asuming that the size of the transfer is less than $y_i - y_j$.

$$(10) \quad G_t = \left[\frac{2(i - j)}{TY \cdot N} \right]$$

$$(11) \quad T_t = \left[\frac{\ln(y_i/y_j)}{\ln(N)\mu} \right]$$

$$(12) \quad A_t = \left[\frac{\mu^{\varepsilon} \; (1-A)^{\varepsilon(1-\varepsilon)}}{N \cdot TY} \right] \left[y_i^{-\varepsilon} - y_j^{-\varepsilon} \right]$$

$$(13) \quad C_t = \frac{(y_i - y_j)}{\mu^2 TY}$$

As measured by the Gini, the effect on the concentration of landownership of a small transfer of land from ownership unit i to j is a function only the relative rank of the two ownership units and of total land holdings (equation 10). The initial level of concentration is irrelevant, as is the amount of land owned by i relative to j. The same transfer would also effect the coefficient of variation in a linear fashion (equation 13). Unlike these two indices, the effect on concentration of a transfer between two ownership units is nonlinear when measured by either the Atkinson or the Thiel (equations 12 and 13, respectively). In the case of the Thiel, the change in measured concentration is a function of the ratios of the landholdings. Like the Gini, initial concentration does not affect the size of the change in the Thiel. In the case of the Atkinson, the change in measured concentration due to a mean preserving transfer is a complex function of the relative land holdings of the two ownership units, the level of initial ownership concentration, and the value of .

IV. Grouped Data.

The second problem that we address in this paper arises whenever grouped data must be used in calculating concentration indices. This problem is essentially an aggregation problem. When dealing with grouped data, all observations falling within each group are aggregated so that we have no information on the variation within each group.

Two cases are encountered. In the less serious case, we have data on the number of individuals within each group and data on the total land owned by individuals in each group, or what is equivalent, data on the average land holdings for each group. In this case, the problem is simply that we do not know what the variation is within each group. However, because we know the total amount of land owned by ownership units within each group, we know what the intergroup variation is. The standard procedure in such a situation is to calculate the concentration index under the assumption that ownership within each group is equally distributed. Clearly, this is a lower bound on the overall level of concentration. It is also possible to calculate the maximum degree of concentration consistent with the group means by assuming that landownership within each group has the maximum variance consistent with known information. The amount of error in concentration estimates introduced by this type of aggregation is a function of the number of groups: the greater the number of groups, the smaller the error.

The more serious grouping problem arises whenever group averages are unavailable. We thus have data only on the number of individuals within each group and the total land holdings of the population. In this case we do not have information that will permit calculatation of the total or average land holdings within each group. One way to proceed is to assume that the midpoint of each group is the average for that group. Multiplying the number of individuals in the group by this "average" land holding gives an estimate of the total holdings of the group. This proceedure will under- or over-estimate the total holdings of the group as the actual mean land holding for the group exceeds or falls short of the boundary midpoint. If the distribution is skewed to the right, this proceedure may result in an estimate of total land holdings that exceeds the actual amount of land held.

If the upper bound of the highest group is open, as is usually the case, the total holdings of the highest group is estimated by taking the difference between actual land holdings of the entire population and the total estimated land holdings of other groups. Estimating the landholdings of the upper group in this way can result in an average landholding for the group that is less than the lower bound for the group, or, in extreme cases, even result in an average which is negative.

V. The Concentration of Landownership: 1946-1982.

In the first part of the preceeding section, a comparison of four concentration indices showed some important differences in the way the indices weight indivdual ownership units. The second part of the section pointed out some difficulties that arise whenever grouped data are used in calculating land ownership concentration. The purpose of this section is to assess the empirical importance of both these issues.

Estimates of the concentration of farmland ownership among farm operators for the years 1946, 1979, and 1982 as measured by the Gini, the Thiel, the coefficient of variation (C.V.), and two versions of the Atkinson, one with $\epsilon = 0.5$ (Atkin1), the other with $\epsilon = 1.5$ (Atkin2) are presented in Table 1. Also in Table 1 are estimates of farmland ownership concentration among landlords for the years 1946 and 1979. Grouped data were used in calculating these indices. The minimum variance assumption was used for all three years. For 1946, group means were not available, so the class midpoints were used as group means. Group means were available for the other two years. The minimum variance assumption in these cases means that all

Table 1. Concentration of Farmland Ownership: 1946, 1979, and 1982: Minimum Estimates.

	Year	Thiel	CV	Atkin1	Atkin2	Gini
Operators:						
	1946	0.082566	1.497072	0.302329	0.665180	0.598050
	1979	0.133081	2.355679	0.374497	0.666127	0.656102
	1982	0.139281	2.360840	0.389698	0.687219	0.670808
Landlords:						
	1946	0.081414	1.371662	0.267382	0.615274	0.562663
	1979	0.165447	3.069102	0.425419	0.722112	0.691429

Sources:
Indices for 1946 were calculated from data in Inman and Fippin. Indices for 1979 and 1982 were calculated from U.S. Department of Commerce 1982 and 1984, respectively.

ownership units were assumed to own the average number of acres for the group.

Looking first at the question of choice of index, this table shows that the choice of index would not matter if the goal were an ordinal ranking because all indices indicate that the concentration of farmland ownership increased between 1946, 1979, and 1982 among farm operators and between 1946 and 1979 among landlords. If, on the other hand, our purpose is to measure how much ownership concentration has increased, the choice of index makes a big difference in what conclusions we would draw. For example, the Atkinson index with $\epsilon = 1.5$ shows only a 3% increase in concentration among operators from 1946 to 1982. In contrast, the Thiel indicates an increase of nearly 170% over the same period.

Turning now to the grouping problem, a comparison of Tables 1 and 2 shows the importance of assumptions regarding within-group variances. The only difference between the calculation of the concentration indices presented in Table 2 and those in Table 1 is the assumption regarding within group variances. Table 1 was calculated under the assumption of no within-group variance. Table 2 was calculated under the assumption of maximum within-group variance consistent with the data. For 1946, the only constraints are the number of landowners within each group and total land owned. Under these constraints, the maximum within-group variance is attained when one landowning unit owns the maximum number of acres for the group and all remaining units own the minimum number of acres. Consequently, the differences between Table 1 and Table 2 for 1946 are very large. Because group means were available for 1979 and 1982, the differences between the two assumptions are much smaller.

The above comparison also brings out differences between the indices. In calculating the indices in Table 2, within-group variances were the maximum consistent with the data. However, for 1979 and 1982, the Gini and both versions of the Atkinson indicate a lower level of concentration when within-group variances are maximized than when within-group variances are zero.

Table 2. Concentration of Farmland Ownership: 1946, 1979, and 1982: Maximum Estimates.

	Year	Thiel	CV	Atkin1	Atkin2	Gini
Operators:						
	1946	0.286648	18.324600	0.550077	0.938388	0.745580
	1979	0.141509	2.417121	0.359112	0.664510	0.651975
	1982	0.147748	2.420361	0.373949	0.685635	0.666287
Landlords:						
	1946	0.284716	12.508916	0.491807	0.872085	0.705522
	1979	0.165689	3.082353	0.425640	0.722275	0.691665

Sources:

Indices for 1946 were calculated from data in Inman and Fippin. Indices for 1979 and 1982 were calculated from U.S. Department of Commerce 1982 and 1984, respectively.

VI. Summary and Conclusions.

The evidence presented in Tables 1 and 2 demonstrate the importance of the choice of concentration index in measuring landownership concentration. While a discussion of why one index might be preferred over another was beyond the scope of the paper, the analysis of Part II presented two ways to compare and evaluate different indices. In the absence of any justification for using one index over another, the use of multiple indices in any analysis of concentration is recommended. Tables 1 and 2 also demonstrate that whenever grouped data is used, the assumption regarding within-group variances can have rather large effects on measured concentration.

References

Atkinson, Anthony B. "On the Measurement of Inequality." Journal of Economic Theory. 2(1970): 244-263.

Blackorby, Charles and David Donaldson. "Measures of Relative Equality and Their Meaning in Terms of Social Welfare." Journal of Economic Theory. 18(1978): 59-80.

Inman, Buis T. and William H. Fippin. Farm Land Ownership in the United States. Miscellaneous Publication No. 699. Washingtion, D.C.: United States Department of Agriculture. December 1949.

United States Department of Commerce. 1978 Census of Agriculture: 1979 Farm Finance Survey. Washington, D.C. 1982.

United States Department of Commerce. 1982 Census of Agriculture: United Sates Summary and State Data. Washington, D.C. 1984.