Staff Paper No. 294 May 1991

Single- and Multi-Dimensional Theil Inequality Measures: An Application to Farm Households

by

Stephan J. Goetz

Single- and Multi-Dimensional Theil Inequality Measures: An Application to Farm Households

by

Stephan J. Goetz*

Assistant Professor
Department of Agricultural Economics
University of Kentucky

May 14, 1991

---

* Selected Paper presented at the American Agricultural Economics Association Annual Meeting, Manhattan, Kansas, 4–7 August, 1991. The author gratefully acknowledges valuable comments made by Jerry Skees, but is solely responsible for the contents of this paper.

Single- and Multi-Dimensional Theil Inequality Measures:

An Application to Farm Households

Abstract

Earlier analyses of farm-level inequality are extended by incorporating measures of inequality *between* groups of farmers and multiple attributes, including years of education. Single- and multi-dimensional inequality declined between 1985 and 1989. A sharp decrease occurred in inequality among gross farm sales between full- and part-time farmers.

Single- and Multi-Dimensional Theil Inequality Measures: An Application to Farm Households

I. Introduction

The debate surrounding the 1990 Farm Bill, stemming in part from the perception that benefits of farm programs are concentrated among a few farmers, provides a renewed stimulus for evaluating the extent of economic inequality among agricultural households. This evaluation is particularly relevant given the exodus of farmers from agriculture which, some allege, is leading to an increasingly bimodal size distribution of farms as measured by gross sales of agricultural output (see, e.g., Lin *et al*; U.S. Congress). A more subtle manifestation of the exodus is increased reliance by farm operators on non-farm income.

These structural changes lead to the question of how the distribution of net income and other measures of economic well-being is changing among those who remain in the farm sector, and among those who do or do not rely on the non-farm sector for a portion of their income. For example, is net farm household income from all sources becoming more concentrated in the hands of fewer and larger farm operations over time? Is income distributed more evenly among those who work off the farm, and how is the distribution of income *between* full- and part-time farmers changing? Other measures reflecting economic well-being include the value of gross farm sales (an important measure since the transfers or benefits received from farm programs in most cases increase proportionally with the value of gross sales), the net worth of farm households, and years of education (which serve as a proxy of human capital).

Earlier analyses of the distribution of farm income and other welfare measures were generally based on the Gini index of concentration (Ahearn and El-Osta; Ahearn, Johnson and Strickland; Gould and Saupe; and Carlin and Reinsel). These analyses are limited in two respects. First, they do not address the question of how income is distributed between different groups of farmers. Second, the papers do not provide a straightforward method of aggregating welfare attributes which are measured in different units (such as income and education) to arrive at a comprehensive multi-dimensional measure of inequality. This paper presents an application which rectifies these shortcomings. In addition, data from more than one year are used, so that it is possible to discuss trends.

The paper begins by discussing the measurement of inequality using both single- and multi-dimensional measures of economic inequality. The consideration of multiple dimensions or attributes is motivated by the fact that focusing on only one attribute at a time may provide an incomplete and potentially misleading representation of the direction and extent of changes in composite inequality over time. Section III describes the farm household panel data used for the empirical analysis, while Section IV presents and discusses the results from the application of the inequality measures.

II. Measuring Inequality

We will consider a collection of $p \in [1,P]$ categories or sets of a total of $i \in [1,I] \subset \mathcal{F}$ farm households, partitioned as $\mathcal{F}_1 \cup \mathcal{F}_2 \cup \ldots \cup \mathcal{F}_p = \{1,2,\ldots,I\}$ and $\mathcal{F}_p \cap \mathcal{F}_q = \varnothing$, $p \neq q$. Furthermore, we will be concerned with $n \in [1,N]$ different welfare attributes W, where W_{in} denotes the amount of the nth attribute possessed by the ith household and w_{in} the share of the household in the total of the attribute ($w_{in} = W_{in}/\Sigma_{i \in I} W_{in}$, $\Sigma_{i \in I} w_{in} = 1$, and $w_{in} > 0 \; \forall \; i,n$). For each set, I_p denotes the number of households in the set ($\Sigma_{p \in P} I_p = I$).

Theil's entropy-based index is used to compare the evolution of inequality of the welfare attributes, along with changes in within- and between-set inequality, over time. Assuming con-

tinuously distributed attributes, let $F(W_n)$ denote the cumulative distribution function of attribute W_n. Then $F(W_n^o)$ is the proportion of households possessing a level $W_n < W_n^o$. Further, the arithmetic mean of W_n is $\bar{W}_n = \int W_n dF$, while the geometric mean is $\bar{W}_n^g = \exp \int \log W_n dF$ (see also Cowell).[1] Theil's inequality index is calculated from (suppressing n),

$$(1) \qquad \mathbb{I}_u = \int (W/\bar{W}) \log(W/\bar{W}) dF$$

As discussed, for example, in Maasoumi (1989; see also the references therein for an axiomatic treatment), Theil's index is independent of the mean of the distribution investigated (i.e., it is homogenous), it is symmetric, and it satisfies the Transfer Principle of Pigou and Dalton. The discrete analogue of (1) is

$$(1') \qquad \mathbb{I}_u = (1/I\bar{W}) \Sigma_{i \in I} W_i \log(W_i/\bar{W}) = \Sigma_{i \in I} w_i \log I w_i .$$

This total inequality can be decomposed into an inequality within set $\mathcal{S}_p$,

$$(2) \qquad \mathbb{I}_u^{wp} = \Sigma_{i \in \mathcal{S}_p} w_i / W_p \log I_p w_i / W_p \quad \forall\, p.$$

The average within-set inequality can be calculated from (2) as $\mathbb{I}_u^0 = \Sigma_{p \in P} W_p \mathbb{I}_u^{wp}$, while the between-set inequality is

$$(3) \qquad \mathbb{I}_u^b = \Sigma_{p \in P} W_p \log W_p I / I_p.$$

As we will see, the additive decomposability of Theil's entropy index is a convenient feature. The index can range from 0 (complete equality) to a maximum of log$I$.

Next we discuss the measurement of inequality among n welfare attributes. A Cobb-Douglas function $C_i = \Pi_{n\in N} W_{in}^{\beta_n}$ is used to aggregate the attributes, since it can be shown that this functional form minimizes the divergence of the distribution of the attribute aggregate $C_i^* = C_i/\Sigma_{j\in I} C_j$ from the n underlying distributions of the individual attributes (Maasoumi, 1986). Following Maasoumi (1989; see also Judge *et al.*), we use the Kullback-Leibler information criterion,

$$(4) \qquad D_o(C^*, w; \beta) = \Sigma_{n\in N}\beta_n\{\Sigma_{i\in I} C_i^*\log(C_i^*/w_{in})\}.$$

Clearly, some subjectivity is implicit in the choice of β_n, but a sensitivity analysis using different vectors β will reveal the robustness of the inequality index under alternative weighting schemes. An advantage of this functional form (as opposed to the linear form, for example) is that it permits substitution among welfare attributes.

Matters are simplified when we assume that the W_n are jointly distributed as log-normal variates. Further following Maasoumi (1989), let $\mu = E\log W_n$ be the vector of means of the attributes and $\Psi = (\sigma_{nm})$ the covariance matrix. Then, using familiar results in Aitchison and Brown as well as Kendall and Stuart, it can be verified that

$$(5) \qquad C \sim \Lambda(\beta'\mu, \beta'\Psi\beta),$$

and that the multivariate equivalent of Theil's index is

$$(6) \qquad \mathbb{I}_v = \tfrac{1}{2}\beta'\Psi\beta = \Sigma_{n\in N}\beta_n^2(\tfrac{1}{2}\sigma_{nn}) + \Sigma_{n\in N}\beta_n\Sigma_{n<m\in N}\beta_n\sigma_{nm}.$$

The first summation is the sum of the direct contributions of the attributes, while the second represents the weighted covariances of the welfare attributes. The individual contribution of attribute W_m is

$$(7) \qquad \partial \mathbb{I}_v / \partial W_m = \mathbb{I}_{vm} = \tfrac{1}{2}\beta_m^2 \sigma_{mm} + \beta_m \Sigma_{n \neq m \in N} \beta_n \sigma_{nm}.$$

The contribution of a given attribute to aggregate inequality can be reduced if the attribute is negatively correlated with another attribute (i.e., $\sigma_{nm} < 0$).

III. Data

The data used in this application were collected in a mail survey of a random sample of Kentucky farm households in 1986, with a follow-up survey in 1990 (using the methods described in Dillman). Data represent activities during the 1985 and 1989 calendar years. Due to the retirement of some respondents from farming and other forms of sample attrition that arise in panel studies, the useable sample size fell from (I=) 859 to 542 over the five-year period.[2] Follow-up interviews of non-respondents in the first year revealed no systematic bias from analyzing the remaining sample. In the two years, 61.8% and 62.0% of the farm households had income from non-farm sources. These percentages are very similar to those reported in Ahearn and El-Osta for a nationally representative sample of farmers (using data from the 1988 "Farm Costs and Returns Survey"). Further details and results of the Kentucky surveys are discussed in Collins *et al*. Attributes of inequality W used in this paper include net household income from all sources, gross sales of crops and livestock, net worth (farm household assets minus debts) and human capital as measured by years of formal education of the household head (so that N=4).

IV. Results and Discussion

Theil measures of single-attribute inequality declined in the two years investigated (Table 1). This suggests the exodus from farming is leading to an increasingly equal distribution of the four attributes among families remaining in farming. This result is at variance with the general trend in the U.S. economy as a whole (Slottje, footnote 3 on p. 268) and the distribution of income among Kentucky counties (Goetz *et al.*). The greatest amount of inequality is found in gross farm sales of households, which also exhibited the greatest degree of stability over the period examined. Years of education exhibit the lowest extent of inequality. Household net worth was also unequally distributed, but the degree of inequality in that measure declined by the largest proportion of all measures shown during the two years. In comparison, Theil (1989) reports a value of 0.6407 for the degree of international income inequality in 1985, and 0.0859 for 21 more-developed nations for the same year. To examine how these trends in Kentucky compare with trends in more traditional agricultural areas such as the midwestern corn belt, a separate analysis for total inequality of net household income was carried out for farm households located in five counties which border on and are representative of midwest agriculture (the counties are Daviess, Henderson, McLean, Ohio and Union). The inequality indices for the two years are 0.2865 and 0.1540; this suggests the results apply to a wider range of U.S. agriculture, and not only to Kentucky.

Table 1 also suggests that the smallest discrepancy in inequality among those with ($\mathcal{F}_1$) and without ($\mathcal{F}_2$) off-farm income occurs in the case of gross sales (see the within-group columns; in our case $P=2$). For net household income and education the degree of inequality is lower among farm households with off-farm income, while the opposite is true for the net worth of farmers. This may suggest that off-farm income equalizes net household income, a result also reported in Ahearn and El-Osta (who use a Gini index for one year) and elsewhere, while increasing the disparities of net household worth. Both average net household income and years of education are higher for farms

Table 1: Univariate Inequality Measures for Various Attributes by Off-Farm Income Status, Kentucky Farm Households, 1985-1989

Attribute & Year	Off-Farm Income No	Off-Farm Income Yes	Theil Inequality Total	Theil Inequality Within Off-Farm Income No	Theil Inequality Within Off-Farm Income Yes	Theil Inequality Between
Net Household Income†	----------$----------					
1985	26,545	35,313	0.2653	0.4001	0.1848	0.0093
1989	30,533	44,446	0.1886	0.2658	0.1307	0.0158
Off-Farm Income						
1985		27,731			0.3213	
1989		30,664			0.2838	
Gross Sales						
1985	81,011	30,095	1.1209	1.0100	0.9859	0.1184
1989	60,386	45,300	1.0257	1.0257	1.1238	0.0101
Net Worth‡						
1985	235,155	160,704	0.8918	0.7104	0.9904	0.0171
1989	309,925	205,120	0.6546	0.5355	0.7260	0.0210
Education	-------Years-------					
1985	10.95	12.40	0.0353	0.0686	0.0312	0.0035
1989	10.86	12.62	0.0346	0.0403	0.0277	0.0026

Data Source: Kentucky Farm Change Surveys (Various Years), University of Kentucky.

†. Income data were solicited in eight categories owing to the sensitivity of this type of data. A value of $84,000 is assumed for the open-ended upper interval average in each of the two years.

‡. Some households had non-positive net worths (assets minus debts) which were recoded as "1". Indices consequently are lower bounds. Corrections were made for the following number of households (note: the first column is for households with off-farm income, the second for those without; Average $ is the average net worth of the group, N is the number of households with a negative net worth):

Year	Average $	N	Average $	N
1985	−$166,207	11	−$40,443	27
1989	−$4,100	1	−$17,700	15

The inability of Theil's measure to include negative values can be a serious disadvantage. The Gini index can be calculated if there are negative numbers, but it can then exceed a value of 1, making it difficult to interpret.

with off-farm income, but they have lower gross sales and net worths. The higher average years of education may help farmers obtain off-farm employment by acting as a recognizable signal for potential employers. Further, the between-group inequality (i.e., between those with and without off-farm income) increased in the case of net household income and worth, but fell drastically for gross sales. The latter result is not surprising since the average gross farm sales of the two groups are converging during the period analyzed, suggesting those with off-farm income are "catching-up" with those not holding jobs off-the-farm in terms of gross farm sales.

Multivariate or aggregate measures of inequality for the same set of households are shown in Table 2.[3] The total measure is largely insensitive to the choice of the weighting vector β in the two years, but the same is obviously not true of the marginal or individual contributions of the single attributes. When the attributes are weighted equally (β_2), gross sales contribute the most to total inequality. The covariances among the single attributes are all positive. Not surprisingly, therefore, the total inequality declined along with the individual measures over the period analyzed. However, that decline was slight. Also, in cases where an attribute is negatively correlated with another, it is conceivable that the overall effect of an attribute is to reduce aggregate inequality. In farming, an example may be general vs. specific human capital (i.e., formal schooling vs. on-farm experience). For illustrative purposes the multivariate inequality measure is decomposed for those with and without off-farm income in 1989. As expected from the earlier analysis, the total inequality among those without off-farm income is about twice that of those who have such an activity; this may suggest that off-farm income also contributes to a reduction in aggregate inequality.

V. Summary

This paper applies single- and multi-dimensional inequality Theil measures to examine changes in inequality over time for a panel of Kentucky farm households. One convenient feature of

Table 2: Multivariate Inequality Measures for Kentucky Farm Households, 1985 and 89

Year & Weights β_s†	Total	Individual Contribution: Net HH Income	Individual Contribution: Gross Sales	Individual Contribution: Education
1985 Sample				
β_1	0.3526	0.1550	0.1800	0.0128
β_2	0.3347	0.0880	0.1652	0.0231
β_3	0.3386	0.0364	0.1505	0.0310
1989 Sample				
β_1	0.3215	0.1334	0.1713	0.0128
β_2	0.3152	0.0793	0.1555	0.0227
β_3	0.3260	0.0348	0.1398	0.0299
1989: No Off-Farm Inc.				
β_1	0.5032	0.2278	0.2507	0.0174
β_2	0.4924	0.1359	0.2177	0.0316
β_3	0.5060	0.0600	0.1847	0.0427
1989: With Off-Farm Inc.				
β_1	0.2222	0.0742	0.1364	0.0083
β_2	0.2236	0.0445	0.1280	0.0156
β_3	0.2345	0.0197	0.1195	0.0220

† Weights are $\beta_1 = (3/6, 2/6, 1/6)$; $\beta_2 = (2/6, 2/6, 2/6)$; $\beta_3 = (1/6, 2/6, /3/6)$ for income, gross sales and years of education.

Theil's index is that it can easily be decomposed into a total, within- and between-group inequality measure. This is useful for assessing inequality among and between different subpopulations of farms, such as those which do and do not participate in federal farm programs, or do and do not work off the farm; similarly, regional inequality differences among U.S. farm households can easily be calculated. The index can also be used for calculating a multi-dimensional or aggregate inequality measure.

Panel data from a sample of Kentucky farm households suggest that overall inequality in a variety of welfare measures declined from 1985 to 1989. Further, the inequality of net household income and years of education is higher for those without off-farm income, while the inequality is about the same for gross farm sales, and lower in the case of net household worth. At the same time, the analysis showed that the degree of inequality in gross farm sales between those who do and do not work off the farm declined sharply during the period investigated. Due to strictly positive covariances among the welfare measures considered here, aggregate or total inequality followed the same trend as the individual measures.

References

Ahearn, M.C. and H.S. El-Osta, "Income Inequality Among Farm People: The Role of Off-Farm Work," in *Proceedings*, Rural Planning and Development: Visions of the 21st Century, Feb. 13-15, 1991, Orlando, Florida (Vol. 2).

Ahearn, M., J. Johnson, and R. Strickland, "The Distribution of Income and Wealth: An Analysis of Farm Family 'Well-Being'," *American Journal of Agricultural Economics*, 67 (1985) 1087-94.

Aitchison, J. and J.A.C. Brown, *The Log-Normal Distribution*, Monograph No. 5, Cambridge University Press, Cambridge, 1957.

Carlin, T. and E. Reinsel, "Combining Income and Wealth: An Analysis of Farm Family Well-Being," *American Journal of Agricultural Economics*, 55 (1973) 38-44.

Collins, T., K. Stephenson, J.R. Skees and L.E. Swanson, "The Farmer Speaks: Kentucky Farm Change, 1986-1988," Departments of Agricultural Economics and Sociology, College of Agriculture, University of Kentucky, Lexington, KY, Special Extension Publication, April, 1990.

Cowell, F.A. *Measuring Inequality: Techniques for the Social Sciences*, Halsted Press, John Wiley and Sons, New York, 1977.

Dillman, D.A. *Mail and Telephone Surveys: The Total Design Method*, New York, John Wiley and Sons, 1987.

Goetz, S.J. and D.L. Debertin, "Income Growth, Stability and Inequality in Kentucky: A Regional Analysis for 1969-87," University of Kentucky, Department of Agricultural Economics, Staff Paper No. [draft; 1991-forthcoming]

Gould, W.B. and W.E. Saupe, "Changes in the Distribution of Income and Wealth of Farm Households: Evidence from Wisconsin Panel Data," *North Central Jrnl. Agricultural Economics*, 12 (1990) 31-46.

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lütkepohl and T.-C. Lee, *The Theory and Practice of Econometrics*, 2nd edition, Wiley Series in Probability and Mathematical Statistics, New York, 1985.

Kendall, M. and A. Stuart, *The Advanced Theory of Statistics*, Vol. 1, 3rd Edition, C. Griffin and Company, London, 1969.

Lin, W., G. Coffman and J.B. Penn, "U.S. Farm Numbers, Sizes and Related Structural Dimensions: Projections to Year 2000," USDA, Economics, Statistics and Cooperative Service, Technical Bulletin No. 1625,1980.

Maasoumi, E., "Measurement and Decomposition of Multi-Dimensional Inequality, *Econometrica*, 54 (1986) 991-7.

Maasoumi, E., "Continuously Distributed Attributes and Measures of Multivariate Inequality," *Journal of Econometrics*, 42 (1989) 131-44.

U.S. Congress, Office of Technology Assessment, *Technology, Public Policy, and the Changing Structure of Agriculture*, OTA-F-285, Washington, D.C.: U.S. Government Printing Office, March 1986.

Slottje, D.J., "Change in the U.S. Income Distribution from 1970 to 1980: Assessing State and Regional Impacts," *Journal of Human Resources*, XXV 2 (1990) 267-74.

Theil, H. *Economics and Information Theory*, North-Holland, Amsterdam, 1967.

Theil, H. "The Development of International Inequality, 1960-1985," *Journal of Econometrics*, 42 (1989) 145-55.

Notes

1. All logs are natural logs (base *e*).

2. An important caveat is in order here: in the second and subsequent administrations of a panel survey, the retiring population is lost but no "newcomers" are added to the sample. Consequently, younger farmers are under-represented. If younger farmers have lower or higher incomes than older farmers, the inequality index would represent a lower bound of the true index in 1989. However, the qualitative results are likely to remain valid, since there is a net exodus *from* agriculture.

3. Net worth is excluded because of the data problem discussed in the second footnote to Table 1.