Explaining Changes in the Distribution of Annual Dairy Farm Income over Time

Hung-Hao Chang, Richard N. Boisvert, and Loren W. Tauer*

Selected paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Orlando, Florida, July 28-29, 2008

Abstract

This paper identifies factors affecting the distribution of farm income among dairy producers over time. Using data from participants in Cornell's Dairy Farm Record Program, we decompose differences of farm income distributions into those due to: differences in means of observable farm characteristics, differential marginal effects of characteristics and unobserved random error. The distribution of farm income is affected by factors reflecting the operators' experience and investment in human capital and indicators of management efficiency and level of capital investment. The marked changes in marginal contributions of these factors explain most of the total change in the distribution of income.

Key words: Dairy farm income distribution, decomposition.

---

*The authors are, respectively, Assistant Professor, Department of Agricultural Economics, National Taiwan University; Professor, Department of Applied Economics and Management, Cornell University; and Professor, Department of Applied Economics and Management, Cornell University. The research on which this bulletin is based is funded by Cornell University Hatch Project 121-7419, Integrated Risk Management Decision Strategies for Dairy Farmers.

Introduction

Due in large measure to the continual growth in income from non-farm sources, the gap between incomes of farm and non-farm households began to narrow about 1960, and for nearly three decades now, average farm household income has been above that of the non-farm sector (Mishra, *et al*. 2002). The particularly rapid growth at the lower income levels was accompanied by reductions in poverty and income inequality (Gardner, 2000). These trends have also reduced substantially the variability in annual farm household income, but variability remains high relative to non-farm household income, due mostly to the persistent fluctuations in income from farming (Mishra, *et al*. 2002).

While this is perhaps not surprising, it is also true the share of the remaining variation in farm household income that stems from the variability in farm income differs by region. And, in contrast to other regions, farm income's contribution to household income variability is on the rise in the West and Northeast. Mishra, *et al*. (2002) document that these regional differences in farm income's contribution to household income variability are partially explained by the commodities produced. Because of the labor intensive nature of dairy production, dairy farmers, in particular, typically have less time for off-farm work, and the variability of income from farming is the largest contributor to variability in household income as well, although during recent years, the rather large fluctuations in milk prices have contributed to the increased variability in dairy farm income. However, this price volatility fails to explain the increased inequality of dairy farm income among farms in states such as New York that seems to persist in real dollar terms, even after controlling for farm size. Moreover, it appears that over time the performance of those farmers at the top end of the income distribution has increased relative to those at the middle and low end of the income distribution.

In a sector such as dairy where the development of new inputs and production processes and techniques continues, this result may well be consistent with Cochrane's (1958) early notion of an agricultural treadmill where the primary beneficiaries of the introduction of new technology are the farmers who are early adopters. This is also consistent with Welch's (1970) early work on the value of education that helped lead to successful adoption of new technology during the green revolution. It is perhaps also consistent with Coxhead's (1992) model whereby the best farm managers can effectively assess and adopt new technologies appropriate to different production environments and in the process ensure a high return for their efforts.

The purpose of this paper is to document the changing inequality of farm income among dairy producers in New York, and to identify the contributing factors. To identify these factors, we decompose farm income inequality by applying two methods heretofore used exclusively (or nearly so) by labor economists to decompose wage inequality. Using the method attributed to both Blinder (1973) and Oaxaca (1973), we decompose differences in farm income between any two years into two separate effects: the differences in farm income due to the differences in the mean levels of observable characteristics (or endowments) of the farm operations between the two years (i.e. the endowment effect), and the differences in farm income due to the differential marginal effects between the two years of the observable farm characteristics or endowments (i.e. the coefficient effect).[1] By also applying a method proposed by Juhn, Murphy and Pierce (1993) we are able to account for differences due to the unobserved random error, in addition to differences accounted by the endowment and coefficient effects. More important, this strategy allows for the examination of how the entire farm income distribution has changed (e.g. how the effects have been different for inequality above or below the mean, or for any particular percentile). Being able to distinguish the change in the performance of those at the upper extremes of the distribution from that of those in the middle and lower end of the income distribution in this way is critically important to the realization of our research objectives.

Using data from individual dairy producers who have participated in Cornell's annual Dairy Farm Record Program, we apply these methods to identify those factors that have accounted for the recent changes in the distribution farm income in New York. We proceed by first describing the methods of decomposition. Then, we describe the data in detail and discuss the empirical specification. After a discussion of the empirical results, we offer some concluding comments.

---

[1] Primarily because of the differences in their objectives, these methods stand in sharp contrast with efforts by Schmit *et al*. (2001) to decompose the contributions of both input and output quantities and prices to the variance in farm income by farm by adapting the linear approximation method for decomposing the variance of the product of random variables proposed by Bohrnstedt and Goldberger (1969). The methods also contrast with efforts to decompose the Gini measure of inequality of net farm household income by source (e.g. Boisvert and Ranney, 1990; Findeis and Reddy, 1987).

Methods of Decomposition

To identify the factors that have led to changes in the distribution of dairy farm income in New York in recent years, we apply the method proposed by Juhn, Murphy and Pierce (1993). By starting with a short discussion of the method originally proposed by Blinder (1973) and Oaxaca (1973), we can make transparent the way in which Juhn, Murphy and Pierce (1993) deal with its several limitations.

Decomposition Based on the Means

Perhaps the most straightforward way to examine differences in dairy farm income across years is to focus on the differences in the mean level of the data between two years. This can be done using the method proposed originally by Blinder (1973) and Oaxaca (1973) to decompose wage inequality between different socio-economic groups. In their applications, the decomposition accounts for differences in wage levels due to the differences in the mean levels of observable characteristics or endowments of individuals in each of the groups, and to those differences accounted for by the differential marginal effects of the characteristics between the groups.

In our application, we would use the Blinder-Oaxaca methodology to decompose the differences in dairy farm income between any two years. In so doing we quantify the separate contributions of differences in mean levels of observable farm and household characteristics and input use between the two years and the differential marginal effects of the characteristics between the two years.

We begin this decomposition by specifying a linear regression model of farm income as:

(1) $Y_{ji} = \beta_j X_{ji} + \varepsilon_{ji}$ $i = 1,\ldots,$ N; $j = 1,\ldots,$ J

where Y_{ji} is the income for farm i in year j; β_j are the parameters of interest; the vector X_{ji} contains covariates that determine farm income; and ε_{ij} is the random error. Drawing a comparison between any two years, for example, the standard Blinder-Oaxaca decomposition of the differences in farm income can be expressed as:

(2) $Y_{1i} - Y_{0i} = [\hat{\beta}_1(\overline{X}_1 - \overline{X}_0)] + [\overline{X}_0(\hat{\beta}_1 - \hat{\beta}_0)]$

where Y_{1i} - Y_{0i} gives the differences in predicted farm incomes between the reference year (year 0) and alternative year (year 1). The terms $\hat{\beta}_1, \hat{\beta}_0$ are vectors of estimated coefficients for separate regressions of the form in equation (1) for year 1 and year 0, respectively. The terms $\overline{\mathrm{X}}_1, \overline{\mathrm{X}}_0$ are vectors of average values of the characteristics (endowments) for the individual farms in the samples for year 1 and year 0, respectively.

By examining both of these equations, it is apparent that this Blinder-Oaxaca method decomposes the differences in farm income into two components. The first term in brackets in equation (2) is usually referred to as the *endowment* effect; it represents that part of the differences in farm income that is accounted for by the differences in the average levels of observed characteristics $(\overline{X}_1 - \overline{X}_0)$ between two years. The second term is usually referred to as the *coefficient* effect; it accounts for the part of the differences in farm income accounted for by differences between years in the processes by which income is determined. The differences in these processes reflect the differences in marginal effects of the endowments on income, and they are captured in the differences in the estimated regression coefficients between years. In our application, for example, these differences in coefficients could reflect changes in the productivity of inputs, or changes in the marginal payoff to human capital or other characteristics of the farmer, farm, or farm household.

Although this rather simple method of decomposition provides important insights into explaining changes in farm income over time, it focuses only on differences in the mean. It fails to identify those factors that also explain changes in the overall shape of the distribution. Thus, it is unable to distinguish the differential effects of these important farm characteristics on the incomes of farms at the lower end of the income distribution relative to the effects on those at the other extreme of the distribution. Furthermore, this simple method ignores any possible contribution to the differences in farm income due to the unobserved random error. These limitations are addressed directly by Juhn, Murphy and Pierce (1993).

Decomposition Based on the Entire Distribution

To introduce the method developed by Juhn, Murphy and Pierce, we repeat equation (1) for year 1 and year 0, respectively, as:

(3) $Y_1 = \beta_1 X_1 + \varepsilon_1$

$$Y_0 = \beta_0 X_0 + \varepsilon_0$$

Again, Y_1 and Y_0 are the vectors of farm income in year 1 and year 0, respectively; X_1, X_0 are the vectors of observable quantities; β_1 and β_0 are the vectors of parameters; and ε_1 and ε_0 are the residuals.

To isolate the residual effects, we distinguish two components: an individual's percentile in the distributions for each of the years (θ_1,θ_0), and the cumulative distribution functions of the residuals in farm income, denoted, F_1(.) and F_0(.) for year 1 and year 0, respectively. By the definition of the cumulative density function, the residuals can be rewritten as:

$$\varepsilon_1 = F_1^{-1}(\theta_1 \mid X_1)$$

$$\varepsilon_0 = F_0^{-1}(\theta_0 \mid X_0)$$

where F_1^{-1}(.) and F_0^{-1}(.) are the inverses of the cumulative distribution functions for year 1 and year 0, respectively.

Within this framework, changes in the inequality of farm income now come from three sources: changes in the distribution of the *X*'s (e.g. changes in the distribution of observable farm endowments and other characteristics, etc.); changes in the contribution of the observable farm characteristics, etc. to farm income (e.g. changes in the *β*'s); and changes in the distributions of the residuals. Accordingly, we wish to decompose the inequality into these three components. This is accomplished using the following relationship:

(4) $$Y_1 - Y_0 = [Y_1^1 - Y_0^1] + [(Y_1^2 - Y_0^2) - (Y_1^1 - Y_0^1)] + [Y_1^3 - Y_0^3) - (Y_1^2 - Y_0^2)],$$

where each of the three component is isolated by the terms in brackets [ ]. Each of these components is discussed in detail below. In so doing, we also make transparent the strategy by which they are estimated.

Component 1. Based on these definitions, the first term in brackets [ ] in equation (4) captures the effects of changing the distribution of observable farm characteristics, *X*, while holding their effects on farm income constant. It can be expressed as the difference between the following two equations:

(5) $Y_1^1 = \beta_0 X_1 + F_0^{-1}(\theta_1 \mid X_1)$

$$Y_0^1 = \beta_0 X_0 + F_0^{-1}(\theta_0 \mid X_0)$$

As suggested by Blau and Kahn (1996), (Y_1^1, Y_0^1) can be viewed as hypothetical farm income distributions based on the different endowments of measurable factors (X_1, X_0) between these two years, but with the coefficients and residual distribution fixed according to the base year regression (e.g. fixed at β_0 and $F_0(.)$, respectively). Based on the estimated coefficients from the base year model, β_0, we can use equation (5) to predict farm income for each observation, and then compute a residual for each observation based on the actual percentile in that year's residual distribution, and the reference year cumulative distribution, $F_0(.)$. Thus, the difference between the two terms $Y_0{}^1$ and $Y_1{}^1$ in these hypothetical distributions isolates the difference in the income distribution due to the change in the observable characteristics between the two years.

As underscored by Juhn, Murphy and Pierce (1993), and in contrast to the more standard variance accounting framework, this strategy for dealing with the residual distribution function allows us to examine how the entire farm income distribution has changed (e.g. how the effects are different for inequality above or below the mean, or within any particular percentile).

Component 2. In a similar fashion, the second term in brackets [ ] in equation (4) captures the change in the distribution of farm income for fixed values of the observable farm characteristics, but by letting their effects differ. To recover these effects, we assume both the observable characteristics and the coefficients differ over time, but that the residual distribution remains the same—that for the base year 0. Then, we estimate the following two equations:

$$(6)\quad Y_1^2 = \beta_1 X_1 + F_0^{-1}(\theta_1 \mid X_1)$$

$$Y_0^2 = \beta_0 X_0 + F_0^{-1}(\theta_0 \mid X_0).$$

We proceed to predict farm income for each of the two hypothetical distributions (Y_1^2, Y_0^2) from equation (6), but we again assign the residual based on the cumulative distribution from equation (5). By subtracting Y_1^2 from Y_0^2, we isolate the combined effects from the changes in the observable characteristics and the coefficients. Then, by subtracting the first component's effects, we isolate the effects due only to the differences in coefficients. This is precisely the calculation in the second term in brackets [ ] in equation (4).

Component 3. The final component of the decomposition (the third term in brackets [ ] in equation (4)) captures the effects due to the change in the residuals for farm income. We isolate this effect by estimating regressions in which the observable characteristics, the coefficients, and the residual distributions are allowed to differ between the two years. These regressions are:

$$(7)\quad Y_1^3 = \beta_1 X_1 + F_1^{-1}(\theta_1 \mid X_1)$$

$$Y_0^3 = \beta_0 X_0 + F_0^{-1}(\theta_0 \mid X_0).$$

Each of these regressions (Y_1^3, Y_0^3) replicates the cumulative farm income distributions in the two respective years, and their difference represents the overall change in the distribution between the two years when all three components are allowed to differ. Thus, to isolate the change in the farm income distribution due to the change in the cumulative error structure, we subtract from this difference the combined effects of the first two components, as calculated by the difference in the two equations from equation (6). This is exactly what appears in the third term in brackets [   ] in equation (4).

The Data and Some Descriptive Statistics

As mentioned above, the data used in the empirical analysis are from individual dairy produces who have participated in Cornell's Dairy Farm Record Program. Through this program, New York dairy producers voluntarily provide annual data related primarily to the farm business. Summary data and some limited analysis of the data are reported in the annual New York State Dairy Farm Business Summary (DFBS) (e.g. Knoblauch, *et al*. 2003).[2] Data from all farms included in the annual summary report for the years 1994 through 2002 are used in this analysis. The number of participants differs by year, ranging from a low of 219 farms in year 2002 to a high of 328 farms in 1995. The sizes of these farms differ substantially as well, ranging from 50 cows to over 2000 cows.

In conducting the analysis, we can apply the methods to decompose the change in farm income inequality between all pairs of years. To illustrate the methods and to capture the

[2]Since participation in this survey is voluntary, the participants in any given year are not a random sample of dairy producers throughout New York. Despite this fact, participants are diverse, both in terms of size and in production per cow, but in any year, they represent farms that are larger and have greater production per cow than the state averages for that year. For this reason, one might also expect that the inequality in farm income among participants would likely be understated somewhat relative to the state.

diversity of the results, we focus on three years, years near the beginning and the end of the period, and a year in the middle. Therefore, in the discussion that follows, we focus on the results for 1995, 1998 and 2001.

For purposes of the analysis, we define net farm operating income (NFOI) as farm receipts minus operating expenses. There are two primary sources of income: milk sales, cull cow sales. The operating expenses include paid labor expenses, and purchased and grown feed expenditures. Fixed costs are not deducted from expenses, primarily because year-to-year variations in fixed costs on these farms are generally small, and typically reflect changes in long term investments rather than annual changes in input and output prices or quantities.

Measures of revenue and expenditures are calculated on an accrual basis to reflect what was actually produced or used in the farming operation within the calendar year. To put the items on a comparable basis they are also converted into constant (1993) dollars. Farm revenues are deflated by the U.S. Index of Farm Prices Received,[3] while farm expenses are deflated by the U.S. Index of Farm Prices Paid. To abstract from the effects of farm size so that the decomposition of inequality would not merely reflect differences in the sizes of the operations, NFOI is measured on a per cow basis.

Over the study period, average NFOI per cow increased in real terms, from a low of $1,200 per cow to a high of more than $1,630 per cow. To gain some perspective on the nature the changes in the distribution of NFOI per cow, we plot its value for selected percentiles for

[3] To calculate milk revenues, the trend in milk production per cow in New York is removed from production levels.

each of the years in the study period in Figure 1. From this figure, it is evident that, with the exception of 2002, the distributions have shifted outward and widened over time, somewhat more at the upper end of the distribution. Through the decomposition of the entire farm income distribution, we hope to identify some of the factors that have led to any differential increases in income across the distribution.

For the three years on which this analysis is focused, the distributions of NFOI per cow are shown in Figure 2.[4] This figure also highlights the fact that the distributions have widened over time, primarily through shifts to the right in the upper half of the distribution. Visually, changes in the lower half of the distributions have been less dramatic.

The data for the explanatory variables used in the regressions needed to perform the decomposition of NFOI are summarized in Table 1 for the three years on which the study is focused. The definitions of most of these variables are straightforward, but others require a bit of explanation.

Some of the variables, such as age and educational level, are thought by many to capture the experience and ability of the farm operator. The variable "age" is the age in years of the principal manager of the farm operation, and "education" is the years of formal education of that manager. We also include a categorical variable for whether or not the farmer reported any income from off-farm work.

---

[4] These distributions are based on the annual farm-level data, but for purposes of presentation the distributions are smoothed using a kernel estimator.

Several variables may also capture some measure of input productivity. The variable "asset value per cow" measures the total farm investment per cow and includes land, buildings and equipment. The variable "acres per cow" measures the total acres divided by the average number of cows in the herd during each year. The variable "cows per worker" measures the average number of cows in the herd divided by the number of full time worker equivalents. The variable "operating cost per hundredweight" is the same as is computed in the annual dairy farm business summary, but for our purposes it is deflated to 1993 dollars using the prices paid by farmers index.

Two of the variables account for the adoption of certain technologies. For example, the categorical variable "milk system" takes the value of one if any type of parlor milking system is used, and it is zero otherwise—e.g. if milking is in stanchions. The variable "rBST used on farm" takes the value of one if recombinant bovine Somatotropin is used on any of the cows during a production year.

A final group of variables reflects differences in important management decisions. The "proportion of land owned" is simply acres owned divided by total acres, and it is designed to reflect differences in asset control management strategies between farmers that own much of their land from those who do not. The variable "grown to total feed expense ratio" is designed to reflect differences in management strategies between those farmers who raise much of their own feed and those who do not. This distinction perhaps isolates those farmers with considerable managerial ability for both crop and animal agriculture from those who view themselves with greater capacity for dairy production. It could also reflect the manager's expectations about the future costs of feed production relative to the cost to purchase feed. The categorical variable "production record" assumes a value of one if the farmer uses some system to monitor or estimate milk production, and it is zero otherwise. In New York, many of the operations that rely on such a system use DHI (Dairy Herd Improvement). The categorical variable "cow purchases" takes the value of one if any cows were purchased during the year. This variable indicates whether or not the farm operator maintains a closed or open herd.

Before moving to a discussion of the empirical results, it is worth commenting on the important changes in some of these explanatory variables over the period. After all, it is the nature of these changes that can account for an important proportion of the change in the entire distribution of farm income. Since many farmers participate in the Dairy Farm Record Program year after year, it is not surprising that on average, some of these variables have not changed a great deal over the three years for which the decomposition is conducted. However, there are some notable exceptions, and in these cases the distributions of these factors has changed as well. As is obvious from the discussion above and Table 1, average NFOI has increased over these three years, but so has the relative dispersion, as measured by the coefficient of variation (CV). The proportion of farms milking in parlors has increased, but this has been accompanied by a smaller CV (Table 1). The numbers of acres per cow has fallen, but its relative dispersion, as measured by the CV, has risen dramatically (Table 1). In contrast, the proportion of farms that use rBST has risen, as has the average number of cows per worker. Finally, the proportion of farmers purchasing cows increased between 1995 and 1998, but is lower in 2001, but its CV increases thereafter (Table 1). Finally, the proportion of farmer reporting some off-farm income is quite different across the years.

Empirical Results

We discuss the empirical results in two separate sections. In the first, we discuss the several regression equations that are needed to accomplish the decomposition. In the second section, we report the results of the decomposition, and discuss them particularly in terms of how the changes in the distributions are explained by differences in observed characteristics over the years and the differences in their marginal effects, as measured by the appropriate estimated regression coefficients.

The Regression Equations

The three regressions of the explanatory variables on NFOI per cow across all farms for 1995, 1998, and 2001 are shown in Table 2. It is these equations that are essential to all three components of the decomposition of the distribution of NFOI per cow between any of the two years.

On balance, the equations perform quite well. Although the values for the adjusted R^2‘s, ranging from 0.30 to 0.44, are perhaps somewhat disappointing, the coefficients on most of the explanatory variables have the expected sign, and many are large relative to their standard errors. For most of the variables, their effects on NFOI per cow are consistent across these years, but there are a couple of important exceptions. As is discussed in the next section, the fact that the coefficients differ substantially in magnitude is one indication that component 2 of the decomposition (which captures the change in the distribution of farm income by letting their effects differ, but for fixed values of the observable farm characteristics) may account for an important share of the total change in the distribution.

One major difference in the technology that distinguishes these farms is the adoption of rBST, and, its use is perhaps one of the best examples of the differential effects across the three years. As one might well expect, its use leads to increased NFOI. The effects of its use decreases over the period—leading to increases in NFOI per cow of 182, 144, and 129 in 1995, 1998, and 2001, respectively, and again the coefficients are statistically significant (Table 2). Since rBST only became available for use in 1994, one might well expect its effects to be somewhat more dramatic during the early years as it is likely that early adopters would have been among the better managers. However, rBST was well publicized before it’s commercial release, allowing all farmers to assess the potential use of rBST in their operation before it’s release. It is also true that the use of production record systems, such as DHIA, lead to higher NFOI in all three years. The effects are statistically significant, and while the magnitudes of the effects differ by year, there is no clear trend, and the differences are not as dramatic as for rBST (Table 2). As one would expect, the farms that use milking parlors tend to have higher farm incomes in both 1998 and 2001, although the reverse is true in 1995. However, the coefficients are not statistically significant in any of the three years (Table 2).

The effects of the two variables often thought to be associated with the operators' experience and the level of human capital are perhaps a bit more difficult to interpret. As expected, the operator's years of education is positively related to NFOI in all years, but the magnitude of the effect in 2001 is between four to six times larger than in the other two years, and it is only in 2001 that the coefficient is statistically significant. While years of education is likely to be highly correlated with investment in human capital and management ability, it is certainly an imperfect measure, and this may explain the statistically insignificance results in 1995 and 1998. In all three years, age, as measured by the logarithm of the farm operator's age, is negatively related to the level of NFOI, although the effect is statistically significant only in the last two years. This result is consistent with the notion that young farmers may well be among the most dynamic and innovative managers. Since the magnitude of this negative effect declines with age, this result could also reflect the fact that the advantage afforded the innovative younger managers is, *ceteris paribus*, partially offset by the experience of older farmers.

Three of the variables that reflect input productivity are also directly related to higher NFOI, and the coefficients are statistically significant. As expected, as acres per cow fall, NFOI rises, but the size of the effect in 2001 is over twice that in 1995, and the size of the effect in 1998 is about mid way between these extremes. The effect of operating cost per hundredweight of milk on NFOI is very large and negative in 2001, and is about two-thirds as large in the early two years. In addition, *ceteris paribus*, farms realize increased NFOI as the value of farm assets per cow rises. The effects are statistically significant, and there is little difference between the magnitudes of the effects across the years. In contrast, as the number of cows per worker rises, there are rather modest decreases in NFOI, and the effect is statistically significant in only 1998. While one might expect NFOI to increase with the number of cows per worker, these results may suggest that labor on some farms is now over extended. It is only in 2001 that raising the proportion of grown feed expenses leads to a decrease in NFOI, and the effect is statistically significant. In the other two years, the signs on the coefficient are positive, but both are statistically insignificant.

The final explanatory variable in the regression equations is a zero-one variable that takes on a value of unity if there are earnings from off-farm work. The coefficients are negative; thus, they are as expected because even though off-farm work may well contribute to overall farm household income, this activity removes labor from the farm, and it may affect the productivity of the operator's remaining labor in production. Despite the fact that the signs on the coefficients are as expected, none is statistically significant. This may be explained by the fact that off-farm earnings are rather modest on most of these farms, and it is not specified in the farm records as to whether it is the operator or a family member working off the farm.

The Results of the Decomposition

By using the results from these three regression equations in the procedures described above, we are able to decompose the changes in the distributions of NFOI per cow between selected years. These decompositions are: between 1995 and 2001 (Table 3); between 1995 and 1998 (Table 4); and between 1998 and 2001 (Table 5). Using the methods by Blinder and

Oaxaca (equation (2)), we decompose these changes at the mean of the distribution of NFOI per cow into those due to differences in the endowment of observed characteristics and those due to changes in the coefficients. We contrast these results with those based on methods by Juhn, Murphy and Pierce (equations (2) through (7)) to underscore the nature of the additional information about the entire distribution of NFOI per cow and the error structure that can be derived through the decomposition over percentiles of the distribution.[5]

It is evident from Tables 3 through 5 that the total effects differ substantially based on the years being compared, and across percentiles of the distribution in NFOI, as do the percentages of these total effects accounted for by the endowment, coefficient, and residual components of the decomposition. To facilitate discussion, it is important to underscore the fact that the decomposition over the entire period is internally consistent with the decompositions over the two inclusive sub-periods. That is, the total effects over the entire period are equal to the sum of the changes between 1995 and 1998 and between 1998 and 2001. These percentages of the total effects accounted for by these two sub-periods are reported in Tables 4 and 5, but for discussion purposes, they are also reported in Figure 3.

It is also evident from Figure 3 that, with the exception of the 5th percentile, over 70 percent of the total change in the distribution of NFOI per cow occurs during the period

---

[5] The total effects of the decomposition at the mean of NFOI are the same regardless of the method of decomposition. Since the method by Juhn, Murphy and Pierce (1993) accounts for differences in the error structure, the contributions of the endowment and coefficients to the total effect differ only slightly. Therefore, to enable consistent comparisons across other percentiles of the distribution, we limit most further discussion of the results at the mean to those based on methods developed by Juhn, Murphy and Pierce (1993).

1995-1998.[6] Furthermore, for this time period, and the entire period (1995-2001), the changes in NFOI at each point in the distribution for which the decomposition results are reported are statistically significant. However, because the largest share of the changes occur between 1995 and 1998, it is not surprising that half of the total changes over the period 1998-2001 are not statistically significant. Thus, to understand the importance of the components on the decomposition of NFOI, we could focus either on the 1995-1998 or the entire period (1995-2001). Primarily for convenience, we focus the discussion on the decomposition results over the entire period, 1995-2001. The individual contributions of the endowment, coefficient and error components to decomposition over this period are provided in detail in Table 3, but are also depicted visually in Figure 4.

Over this period, it is also perhaps not surprising that the results at the 50th percentile are similar to those at the mean (294 at the mean vs. 333 at the 50th percentile). The total effects at percentiles less than the 50th percentile are substantially smaller than at the 50th percentile and above (Table 3).

However, the total effects at percentiles above the 50th percentile offer, at best, only the most modest support of one of our initial hypotheses. Earlier in the paper, we argued that as the distribution of NFOI shifted to the right and widened, we would have also expected that the sizes of the total effects would have increased dramatically relative to those at the mean or the

---

[6] Since the total effects over the period 1998-2001 are negative for the 5th and 10th percentiles, the total effects over the entire period for these two percentiles are smaller than those over the period 1995-1998 (Tables 3 and 4).

50th percentile. This would be consistent with the hypothesis that the "better" managers are able to out perform the others in relative as well as absolute terms. While the total effects at the 75th, 90th and 95th percentiles are somewhat larger than at the mean, it is only at the 75th and 95th percentiles that the total effects are larger than at the 50th percentile.[7] Lack of support for this initial hypothesis is probably explained by the fact that over this period there was only a modest shift to the right in the distribution (e.g. Figures 1 and 2).

Tables 3 through 5 also contain information about the proportions of the total effects accounted for by the endowment, coefficient, and residual effects. As mentioned above, the total effect for mean of the distribution is 294 for the period 1995 to 2001, regardless of the method of decomposition. Based on the 2-component method of decomposition (i.e. Blander-Oaxaca method), 121 percent of this total is due to the coefficient effect, while the endowment effect is negative, accounting for -21 percent of the total. By accounting for the residual effect, the percentage due to the endowment effect increases slightly—to 128 percent of the total. The endowment effect accounts for about a negative 28 percent of the total; there is a negative 1 percent due to the residual component (Tables 3 and Figure 4).

For the period 1995-2001, the relative importance of the coefficient and endowment effects at the 50th percentile and above are similar to those at the mean, as they are for the three percentiles reported that are below the 50th percentile. The only major distinction between these

---

[7] The total effects at these higher percentiles over the period 1995-1998 exceed those both at the mean and the 50th percentile (Table 4), but still not in relation to the differences in the levels of NFOI at these various points on the distributions in any given year.

effects at percentiles above and below the 50th percentile is the contribution of the residual effect. For each of the lowest three percentiles reported, the residual effect is negative, although the effect is statistically significant only for the 5th percentile (Table 3 and Figure 4). These same patterns are also evident for the period 1995-1998, primarily because the "lion's" share of the change over the entire period occurs during the first part of it. These patterns also appear for most of the percentiles for the period 1998-2001, but for this latter period, the sizes of the total effects are substantially smaller than between 1995 and 1998. The total effects between 1998 and 2001 are actually negative for 5th and 10th percentiles.

Some Concluding Observations

In this paper we document the changing distribution of farm income among dairy producers in New York, and identify the contributing factors. Through methods heretofore used exclusively (or nearly so) by labor economists to decompose wage inequality, we decompose a measure of net farm operating income per cow for those farmers participating in Cornell Dairy Farm Records Program. It is evident from the results that not only the mean, but the entire distribution of farm income is affected by factors that account for the experience and investment in human capital of the operator, as well as indicators of management efficiency and the level of capital investment. The marginal contributions of these factors to farm income have changed markedly over our study period, 1995-2001, and it is these changes in the effectiveness of these factors that explain most of the total change in the distribution of income. However, some of this change is also due to differences in the characteristics of the farms and farm operators participating in the farm records program. The magnitudes of these effects were somewhat unanticipated, because many of the same farmers participate from year to year. Moreover, the effect of changes in the average level of these characteristics is negative, thus offsetting some of the gains due to the increased effectiveness of these factors.

Although our present application of these methods of decomposition is limited to a sample of dairy farms in New York, we believe that the methods offer a promising approach to improving our understanding of the changing distribution of farm income over time and across the country. If applied regionally or nationally, we anticipate that estimates of what is called the “coefficient” effect will document the importance or “productivity” of factors related to management ability, the adoption of technology, and the increased propensity for off-farm work, etc. the level and distribution of farm income. Equally important, estimates of what is called the “endowment” effect would seem to document over time how the distribution of farm income is influenced by changes in the structure of agriculture.

References

Blau, F. and L. Kahn. 1996. ”International Differences in Male Wage Inequality: Institutions versus Market Forces,” *Journal of Political Economy* 104(4):791-837.

Blinder, A. 1973. “Wage Discrimination: Reduced Form and Structural Estimates,” *Journal of Human Resources* 8: 436-455.

Bohrnstedt, G. W. and A. S. Goldberger. 1969. “On the Exact Covariance of Products of Random Variables,” *Journal of the American Statistical Association* 64:1439-1442.

Burt, O. R. and R. M. Finley. 1968. “Statistical Analysis of Identities in Random Variables,” *American Journal of Agricultural Economics* 50:734-744.

Boisvert, R. N. and C. Ranney. 1990. "Accounting for the Importance of Non-Farm Income on Farm Family Income Inequality in New York," *Northeastern Journal of Agricultural and Resource Economics* 19:1-11.

Cochrane, W. W. 1958. *Farm Prices: Myth and Reality*. Minneapolis, MN: University of Minnesota Press.

Coxhead, I. A. 1992. “Environment-Specific Rates and Biases of Technical Change in Agriculture,” *American Journal of Agricultural Economics* 74: 592-604.

Findeis, J. L. and V. K. Reddy. 1987. “Decomposition of Income Distribution Among Farm Families,” *Northeastern Journal of Agricultural Economics* 16:165-173.

Gardner, B. 2000. “Economic Growth and Low Incomes in Agriculture,” *American Journal of Agricultural Economics* 82(5): 1059-1074.

Juhn, C., K. M. Murphy, and B. Pierce. 1993. “Wage Inequality and the Rise in Returns to Skill,” *The Journal of Political Economy* 101:410-442.

Knoblauch, W. A., L. D. Putnam, and J. Karszes. 2003. ” Business Summary New York State 2002,” Research Bulletin 2003-03. Department of Applied Economics and Management, Cornell University.

Mishra, A. K., H. S. El-Osta., M. J. Morehart., J. D. Johnson., and J. W. Hopkins. 2002. “Income, Wealth, and the Economic Well-Being of Farm Households.” Agricultural Economic Report 812, Economic Research Service, United State Department of Agriculture, Washington D.C.

Oaxaca, R. 1973. “Male-Female Wage Differentials in Urban Labor Markets,” *International Economic Review* 14: 693-709.

Schmit, T., R. N. Boisvert, and L. Tauer. 2001. “Measuring the Financial Risks of New York Dairy Producers,” *Journal of Dairy Science* 84:411-420.

Welch, F. 1970. “Education in Production,” *Journal of Political Economy* 78:35-5

Figure 1. Distributions of Net Farm Operating Income per Cow, by Percentile

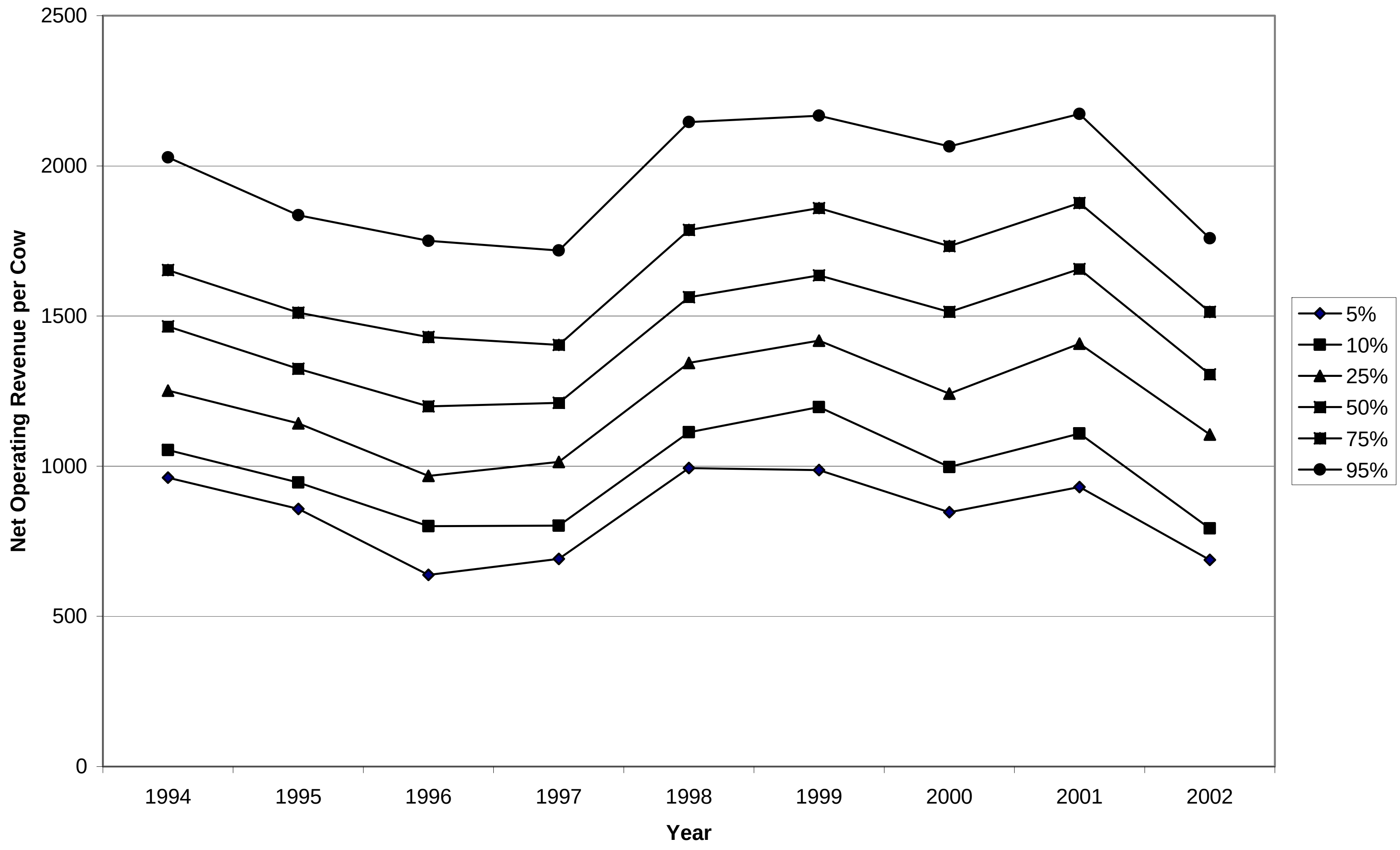

Figure 2. Distributions of Net Farm Operating Income Per Cow (constant 1993 dollars) for 1995, 1998, and 2001
(for purposes of presentation the distributions are smoothed using a kernel estimator)

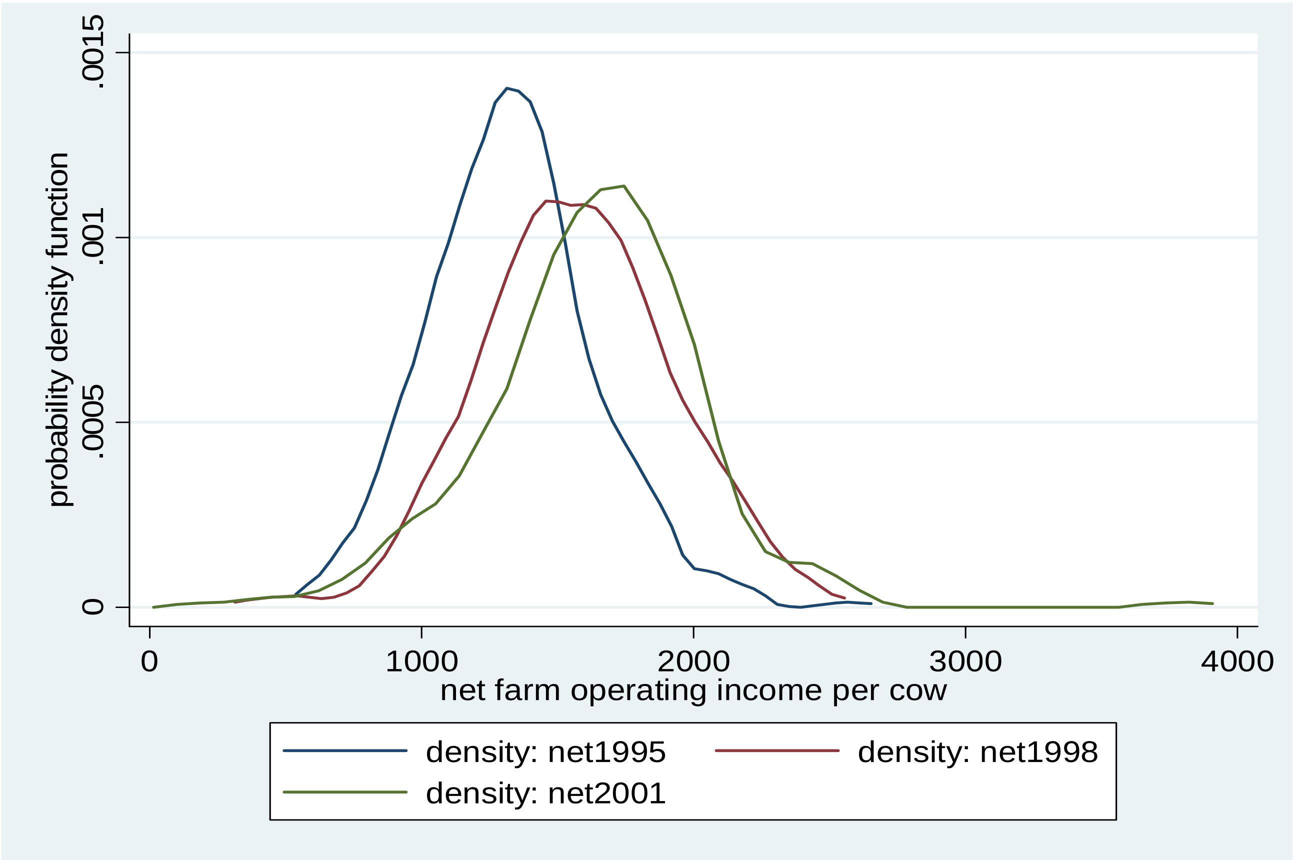

Table 1: Descriptive Statistics for the Samples of New York State Dairy Farms

	Years								
	1995			1998			2001		
Definitions of Variables	Mean	Std. Dev.	CV*	Mean	Std. Dev.	CV*	Mean	Std. Dev.	CV*
Net farm operating income per cow ($)	1331	302	0.23	1566	353	0.23	1625	418	0.26
Operator's age (years)	47	10	0.21	47	10	0.21	49	10	0.20
Operator's dducation (years)	14	2	0.14	13	2	0.13	14	2	0.14
Milking parlor used (1 = yes, 0 = no)	0.52	0.50	0.97	0.61	0.49	0.80	0.63	0.48	0.77
Proportion of land owned	0.69	0.22	0.31	0.67	0.23	0.34	0.65	0.22	0.34
Grown to total feed expense ratio	0.35	0.26	0.73	0.37	0.24	0.65	0.43	0.61	1.43
Production record (1 = yes, 0 = no)	0.90	0.30	0.33	0.88	0.33	0.37	0.85	0.35	0.42
Cow purchases (1 = yes, 0 = no)	0.40	0.49	1.22	0.45	0.50	1.11	0.33	0.47	1.43
Asset value per cow ($10,000)	6899	2094	0.30	6814	2135	0.31	7660	2552	0.33
Acres per cow	4.75	2.17	0.46	4.44	2.22	0.50	4.44	2.37	0.53
Cows per worker	33.54	11.10	0.33	35.77	11.62	0.33	36.26	12.73	0.35
Operating cost per hundredweight ($)	10.41	1.75	0.17	11.28	2.01	0.18	11.97	2.37	0.20
rBST used on farm (1 = yes, 0 = no)	0.46	0.50	1.08	0.52	0.50	0.96	0.51	0.50	0.99
Work off the farm (1 = yes, 0 = no)	0.59	0.49	0.83	0.52	0.50	0.97	0.43	0.50	1.15
Numbers of farms in sample	328			320			225		

*CV is the coefficient of variation.

Table 2: The OLS Regressions of NFOI per cow, Selected Years

	1995		1998		2001	
Variable Definitions	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error
Logarithm of Operator's age	-74.32	69.29	**-129.54**	78.88	**-293.65**	110.38
Operator education (years)	5.60	8.02	8.79	10.01	**31.99**	12.46
Milking parlor used (1=yes, 0=no)	-19.24	36.33	52.10	44.18	49.00	60.58
Proportion of land owned by the operator	**-112.75**	69.18	-30.30	77.01	**191.26**	101.47
Grown to total feed expense ratio	8.24	57.68	20.08	70.50	**-90.05**	35.63
Production record (1=yes, 0=no)	**170.49**	50.08	**163.34**	54.08	**184.78**	66.45
Purchase per cow	28.04	30.17	15.32	36.64	52.43	48.36
Asset value per cow ($10,000)	**0.04**	0.01	**0.05**	0.01	**0.06**	0.01
Acres per cow	**-18.35**	8.56	**-32.08**	9.87	**-41.47**	13.71
Cows per worker	-0.40	1.57	**-3.86**	1.83	-3.36	2.14
Operating cost per hundredweight ($)	**-63.13**	8.44	**-64.50**	8.62	**-82.68**	9.66
rBST used on farm (1=yes, 0=no)	**182.30**	32.64	**144.27**	38.91	**129.31**	50.73
Work off the farm (1 = yes, 0 = no)	-42.36	29.77	-13.96	34.32	-35.08	44.38
Constant	**1907.28**	311.66	**2355.06**	363.53	**2842.37**	489.40
Adjusted *R2*	0.30		0.31		0.44	

The coefficients in bold are significant at least at the 10% level.

Table 3: Decomposition of the Differences in Net Farm Operating Income, 2001 with 1995

	Total Effects		Due to Endowments~		Due to Coefficients~		Due to Residuals~	
	Coefficient	Std. Error#	Coefficient	Std. Error#	Coefficient	Std. Error#	Coefficient	Std. Error#
	Decomposition based on the mean levels (2 components)*							
mean	**294**	34	**-61**	25	**355**	33		
			-21		*121*			
	Decomposition of the Distributions (3 components)**							
mean	**294**	33	**-82**	27	**377**	31	-2	2
			-28		*128*		*-1*	
5 percentile	72	73	**-146**	64	**339**	68	**-121**	55
			-202		*468*		*-166*	
10 percentile	**163**	75	-82	60	**289**	56	-43	38
			-50		*177*		*-26*	
25 percentile	**265**	42	-50	42	**327**	45	-11	17
			-19		*123*		*-4*	
50 percentile	**333**	33	**-82**	32	**402**	38	12	12
			-25		*121*		*4*	
75 percentile	**365**	34	**-63**	33	**411**	40	17	16
			-17		*113*		*5*	
90 percentile	**327**	55	**-151**	45	**459**	52	18	28
			-46		*140*		*6*	
95 percentile	**337**	109	-63	78	**343**	73	57	45
			-19		*102*		*17*	
standard deviation	**115**	38	21	18	**50**	23	**45**	24
			18		*43*		*39*	

~Numbers in italics in the components columns are proportions of the total effect accounted for by the component.

[#] Standard errors calculated with bootstrap methods, 1,000 replications, coefficients in bold significant at least at the 10% level.

* Based on the method by Oaxaca (1973), using equation (2).

[##]Numbers in italics here and elsewhere represent the proportion of the total effect accounted for by the component.

** Based primarily on methods by Juhn, Murphy, and Pierce (1993) and Blau and Kahn (1996), using equations (3) through (7).

Table 4. Decomposition of the Differences in Net Farm Operating Income, 1998 with 1995

	Total Effects[&]		Due to Endowments~		Due to Coefficients~		Due to Residuals~	
	Coefficient	Std. Error#	Coefficient	Std. Error#	Coefficient	Std. Error#	Coefficient	Std. Error#
	Decomposition based on the mean levels (2 components)*							
mean	**235**	27	**42**	17	**276**	24		
	80		*18*		*118*			
	Decomposition of the Distributions (3 components)**							
mean	**235**	25	**-45**	17	**280**	23	-1	1
	80		*-19*		*119*		*0*	
5 percentile	**136**	55	-44	44	**263**	45	**-82**	37
	188		*-32*		*193*		*-61*	
10 percentile	**168**	37	-52	35	**282**	38	**-62**	24
	103		*-31*		*168*		*-37*	
25 percentile	**201**	29	**-48**	27	**263**	28	-14	16
	76		*-24*		*131*		*-7*	
50 percentile	**239**	35	**-53**	28	**280**	30	12	12
	72		*-22*		*117*		*5*	
75 percentile	**276**	34	-33	27	**290**	30	18	14
	75		*-12*		*105*		*7*	
90 percentile	**325**	58	-41	40	**322**	44	**44**	26
	99		*-13*		*99*		*14*	
95 percentile	**310**	48	-44	47	**280**	47	**74**	38
	92		*-14*		*90*		*24*	
standard deviation	**50**	19	9	9	9	12	**33**	15
	44		*17*		*18*		*65*	

[&]The numbers in italic in the total effects column is the proportion of the total effect over the entire period is due to 1995-1998.
~Numbers in italics in the components columns are proportions of the total effect accounted for by the component.
[#] Standard errors calculated with bootstrap methods, 1,000 replications, coefficients in bold significant at least at the 10% level.
* Based on the method by Oaxaca (1973), using equation (2).
** Based primarily on methods by Juhn, Murphy, and Pierce (1993) and Blau and Kahn (1996), using equations (3) through (7).

Table 5. Decomposition of the Differences in Net Farm Operating Income, 2001 with 1998

	Total Effects[&]		Due to Endowments~		Due to Coefficients~		Due to Residuals~	
	Coefficient	Std. Error#	Coefficient	Std. Error#	Coefficient	Std. Error#	Coefficient	Std. Error#
	Decomposition based on the mean levels (2 components)*							
mean	**59**	35	-10	24	**69**	31		
	20		*-17*		*117*			
	Decomposition of the Distributions (3 components)**							
mean	**59**	35	-22	26	**82**	29	-2	1
	20		*-37*		*139*		*-3*	
5 percentile	-64	74	-85	68	50	72	-29	54
	-88		*133*		*-79*		*45*	
10 percentile	-5	77	-76	67	44	59	28	35
	-3		*1671*		*-964*		*-607*	
25 percentile	64	44	0	43	44	42	20	17
	24		*0*		*69*		*30*	
50 percentile	**94**	38	-8	33	**97**	33	5	13
	28		*-9*		*104*		*5*	
75 percentile	**90**	34	-33	34	**130**	34	-7	18
	25		*-37*		*145*		*-8*	
90 percentile	2	53	**-128**	46	**128**	50	2	25
	1		*-6197*		*6202*		*95*	
95 percentile	27	113	-83	84	106	70	4	43
	8		*-308*		*393*		*15*	
standard deviation	**65**	37	10	17	**39**	23	16	22
	56		*15*		*60*		*25*	

[&]The numbers in italic in the total effects column is the proportion of the total effect over the entire period is due to 1998-2001.

~Numbers in italics in the components columns are proportions of the total effect accounted for by the component.

[#] Standard errors calculated with bootstrap methods, 1,000 replications, coefficients in bold significant at least at the 10% level.

* Based on the method by Oaxaca (1973), using equation (2).

** Based primarily on methods by Juhn, Murphy, and Pierce (1993) and Blau and Kahn (1996), using equations (3) through (7).

Figure 3. Percent of Total Change in Distribution of NFOI per Cow, 1995-1998 & 1998-2001

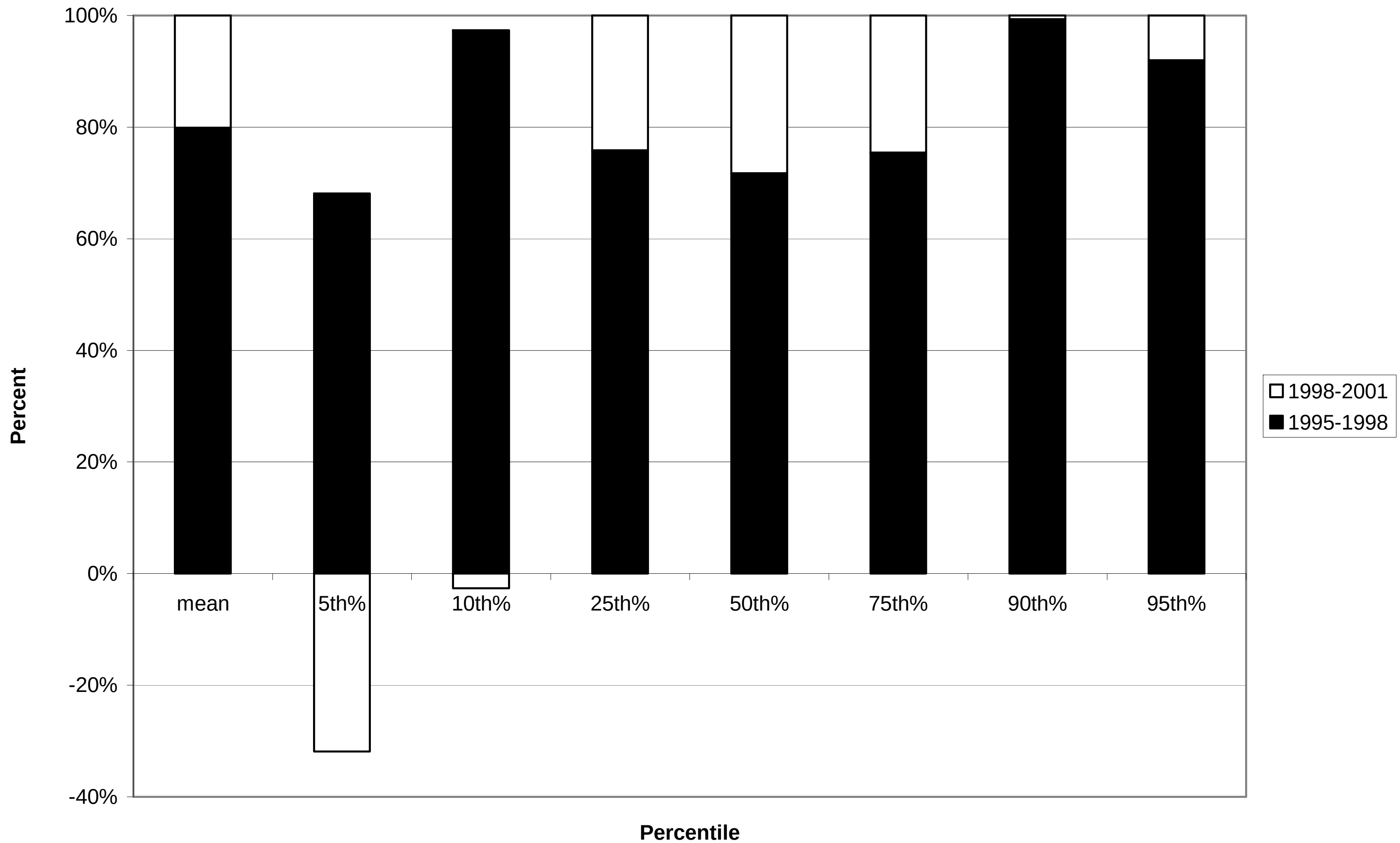

Figure 4. Percent of Total Change in the Distribution of NFOI by Component, 1995-2001