EVALUATION OF INTEGER WEIGHTING FOR THE 1997 CENSUS OF AGRICULTURE, by Wendy Scholetzky, Sampling and Estimation Research Section, Research Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, DC 20250-2000, January 2000, Report No. RD-00-01.

ABSTRACT

The census of agriculture is an important source of statistics about the Nation's agricultural production and provides consistent, comparable data at the county, state, and national levels. Each census record has weights which are used to produce totals for the entire population. The process of rounding weights to integer values has been in place for the last several censuses. When a record's weight is rounded to an integer value, the totals represented by that record may or may not change dramatically. These changes may or may not become negligible when producing totals at the state or county level. This report compares totals calculated with the noninteger weights to the published totals (calculated with the integer weights) for the 1997 Census of Agriculture, evaluates how different these totals are, and examines how the differences relate to the standard error. The analysis examines a number of characteristics at both the state and county levels. The report also examines another weighting approach where the noninteger weights are applied to the record and the weighted data values are rounded at the record level. The difference between totals produced with these values and the noninteger weights is calculated and compared to the above differences.

The reasons for rounding weights, to ease data review procedures and to ensure that publication totals add, are legitimate concerns. The author asserts that it is possible to address these two concerns and improve the totals produced when NASS revamps the census processing system for the 2002 census.

KEY WORDS

1997 Census of Agriculture; Integer weight; Noninteger weight; t-value; Percent difference.

ACKNOWLEDGMENTS

The author would like to thank Dale Atkinson for his guidance in the development of this report, Gail Wade for the Mapinfo tutorial, and Phil Kott for providing the standard error methodology. A special thanks to Chadd Crouse for his tremendous assistance throughout the entire project.

TABLE OF CONTENTS

SUMMARY

The census of agriculture is an important source of statistics about the Nation's agricultural production and provides consistent, comparable data at the county, state, and national levels. Each census record has two weights which are used to produce totals for the entire population. The first weight is the nonresponse weight which accounts for farm operators who did not respond to the census despite numerous attempts to contact them. The second weight is the sample item weight which accounts for both nonresponse and subsampling for data items that are only asked on the long form.

The process of rounding weights to integer values has been in place for the last several censuses. When a record's weight is rounded to an integer value, the totals represented by that record may or may not change dramatically. These changes may or may not become negligible when producing totals at the state or county level. The definition of what a person considers negligible is open to debate. Rather than round the noninteger weights to integer values, another possible approach is to apply the noninteger weights to the record and round the weighted data values at the record level. Thus, the operation's rounded-weighted data values are used to produce the totals for the entire population. With this approach, alternative methodology would have to be used for characteristics such as number of farms and demographic data.

This report compares totals calculated with the noninteger weights to the published totals (calculated with the integer weights) for the 1997 Census of Agriculture, evaluates how different these totals are, and examines how the differences relate to the standard error. The report also compares these differences to the difference between totals produced with the noninteger weights and the record's rounded-weighted data values. The analysis examines a number of characteristics at both the state and county levels.

The reasons for rounding weights, to ease data review procedures and to ensure that publication totals add, are legitimate concerns. The author asserts that it is possible to address these two concerns and improve the totals produced when NASS revamps the census processing system for the 2002 census. In reference to data review procedures, one argument for using the integer weights is that one can "easily" obtain weighted totals for a record of interest during the data review phase. In the 1997 system, this would be accomplished by multiplying the integer weight by each record's data value of interest. Thus, this manual computation is easier if the nonresponse weight is 2 rather than 1.7. However, for the 2002 system, the computer can and should perform this calculation. With this improvement, the value of the weight is no longer relevant. In reference to the concern that the publication totals add, a combination of the integer weights and the record's rounded-weighted data values should be used. The integer weights should be used to produce totals for indicator and categorical variables and for any question where the data values are small for most farms. The record's rounded-weighted data values should be used to produce totals for all other characteristics. With this improvement, the integer sample item weight is no longer necessary unless the 2002 long form contains a question that requires it. The 1997 long form contained no such item. Once the 2002 long form is finalized, the questions will need to be evaluated to determine whether the integer sample item weight is needed.

INTRODUCTION

The census of agriculture is an important source of statistics about the Nation's agricultural production and provides consistent, comparable data at the county, state, and national levels. Census statistics are used by Congress to develop and change farm programs, study historical trends, assess current conditions, and plan for the future. Many national and state programs use census data to design and allocate funding for extension service projects, agricultural research, soil conservation programs, and land-grant colleges and universities. Private industry uses census statistics to provide a more effective production and distribution system for the agricultural community.

The primary purpose of the census of agriculture is to collect information from every farm operation in the U.S. However, weighting adjustments are still necessary to produce totals for the entire population. A nonresponse adjustment was applied because some farm operators did not respond despite numerous attempts to contact them and a sampling adjustment was applied because additional questions were asked to a subset of the population. After the values of the weighting adjustments were determined, each in-scope record was assigned a noninteger nonresponse weight and a noninteger sample item weight (which takes into account both nonresponse and sampling). Two weights are assigned to each record due to the design of the census. The nonresponse weight is used to produce totals for data items collected from all respondents and the sample item weight is used to produce totals for data items collected only from the respondents in the subset. The noninteger nonresponse weight ranged between one and two, inclusively; the noninteger sample item weight ranged between one and twenty-four, inclusively, for respondents in the subset and was equal to zero otherwise.

To simplify certain census processes, a systematic sampling of records was performed to round the noninteger weights to integer values. These integer weights were then used to produce totals for the entire population. To assign integer weights, a sample of records (separate from the long form sampling) was chosen within a group of records with the same noninteger weight, and the integer weight was the next largest integer value for sample records and the truncated value for nonsample records. For example, suppose that ten records had the same noninteger nonresponse weight of 1.2. Then, two records (20% of the records) were selected to receive an integer nonresponse weight of 2 and the remaining eight records were assigned an integer nonresponse weight of 1.

There were two main reasons for using integer weights. First, integer weights make the data review process much easier. For example, during analytical review (a census process where an analyst can examine weighted record-level data), the analyst can easily determine the totals represented by the record by a quick multiplication of each data item by the integer weight. When the record's weight is one (which is true for a majority of records), this calculation is not even needed. If a noninteger weight was used during this process, the analyst would need to hand-calculate the totals represented by the record for each data item, which would slow down the process considerably. Second, integer weights make the publication process simpler. Record-level integer weights ensure that the cells within a column or row add to the total for the column or row. This is true both within and across all publication tables. Additivity of cells is extremely important because the tables are broken into extremely detailed cells for various characteristics and a tremendous number of tables are produced during the publication process (one volume for each state plus one for the United States). If noninteger weights were used during this process, each cell would need to be rounded which does not guarantee additivity of cells. For example, noninteger weights could result in the following scenario...there are 100 farms with less than 1,000 acres and 50 farms with 1,000 acres or more, but the total number of farms is equal to 151.

Totals produced using the integer weights are unbiased estimates of totals produced using the noninteger weights. Therefore, in a perfect world, these two values are equal. However, various factors may affect the results. One possibility is that the records chosen to receive an integer nonresponse weight of 2 may not be a representative sample. Returning to the previous example, suppose that the two sample records had 100 and 200 acres of cropland and the average of the eight nonsample records was 400 acres. This would result in 3,800 total acres using the integer weights compared to 4,200 total acres using the noninteger weights. Another possibility is that the distribution of the data is skewed, which is common with agricultural data. For many characteristics, the distribution is skewed to the right since there are a lot of farms with smaller values and a few farms with larger values. Therefore, the probability of assigning an integer weight of 2 to farms below the average is greater than ½ and the probability of assigning an integer weight of 2 to farms above the average is less than ½. (If the data were normally distributed, both probabilities would equal ½.) In this case, there is a tendency for the integer weights to underestimate the total, especially for small geographic areas (i.e., county-level data).

Note that the process of rounding weights affects the sample item weights and the nonresponse weights differently. First, since the sample item weights are larger in magnitude than the nonresponse weights, there is less of an effect on the totals. Second, since the integer sample item weights are equal to one less frequently than the integer nonresponse weights, more records contribute to the adjustment. (The average noninteger sample item weight is 3.819 and the average noninteger nonresponse weight is 1.139.) In other words, records with an integer nonresponse weight of one do not represent a nonrespondent.

This report compares totals calculated with the noninteger weights to the published totals (calculated with the integer weights) for the 1997 Census of Agriculture, evaluates how different these totals are, and examines how the differences relate to the standard error. The author discusses totals for a number of characteristics at both the state and county levels. The author also examines another possible approach where the noninteger weights are applied to the record and the weighted data values are rounded at the record level (referred to as record's rounded-weighted data values). The difference between totals produced with the noninteger weights and the record's rounded-weighted data values are calculated and compared to the above differences.

BACKGROUND

For more than 150 years, there has been a census of agriculture. The first agriculture census was taken in 1840. From 1840 to 1950, the agriculture census was taken as part of the decennial census. A separate mid-decade census of agriculture was conducted in 1925, 1935, and 1945. From 1954 to 1974, a census of agriculture was taken for the years ending in 4 and 9. In 1976, Congress authorized the census of agriculture to be taken for 1978 and 1982 to adjust the data reference year so that it coincided with other economic censuses. This adjustment in timing established the agriculture census on a 5-year cycle collecting data for years ending in 2 and 7. Beginning with the 1997 Census of Agriculture, responsibility was transferred from the U.S. Department of Commerce, Bureau of the Census to the U.S. Department of Agriculture, National Agricultural Statistics Service.

To reduce data collection costs for the 1997 Census of Agriculture, a screening operation was conducted by mail and telephone to survey about 500,000 records identified as having a low probability of being a farm. Records with no agricultural activity were removed from the census mail list. In December 1997, approximately 3,155,000 records were mailed a census questionnaire and about 34,000 tagged records (farms which were abnormal, multi-unit, in the ARMS survey, or identified by a State Statistical Office (SSO) for special handling) were given directly to the SSOs for data collection. A thank you/reminder card was mailed to everyone in early January 1998. Nonrespondents, except for 1992 census nonrespondents and large farms, were then sent two follow-up mailings in mid-February and late March. Telephone calls were made in early February to nonrespondents who were also 1992 census nonrespondents and in early March to nonrespondents classified as large farms. From early April until late May, telephone calls were made to all remaining nonrespondents to encourage them to respond to the census and to ensure at least a 75% response rate within each county.

To reduce respondent burden, the census used two forms - a long form and a short form. The long form is the same as the short form but contains additional questions on the usage of fertilizers and chemicals, farm production expenditures, value of machinery and equipment, value of land and buildings, and farm related income. All records classified as certainty (tagged records; farms greater than a state-specified level for acreage and total value of products sold (TVP); special insert cases such as Christmas trees and maple sap farms; farms in Rhode Island, Alaska, and Hawaii; or records in a county that contained fewer than 100 farms in the 1992 census) received a long form. A systematic sample was then taken at the county level from the remaining noncertainty records to also receive the long form. The county's sampling rate was based on 1992 farm counts and was either 1-in-1, 1-in-2, 1-in-4, or 1-in-6.

Each census record has two weights which are used to produce totals for the entire population. The first weight is the nonresponse weight which accounts for farm operators who did not respond to the census despite numerous attempts to contact them. Information on farming status for the nonrespondents was obtained from the 1997 Nonresponse Survey. At the end of the census follow-up operations, a stratified systematic sample of nonresponse records was selected for this survey within each state (except for Alaska and Rhode Island which required a 100% response rate). The strata were formed based on screener status, TVP, and number of sources from which the record was obtained during the development of the mail list. From the survey, the number of census nonrespondents that operated farms was estimated at the stratum level in each state, and these estimates were allocated to the county level. Within each county/stratum, the noninteger nonresponse weight was calculated as the total number of farms (respondents that operated farms plus the estimated number of nonrespondents that operated farms) divided by the number of respondents that operated farms. Strata were collapsed if the noninteger nonresponse weight was greater than two, to prevent an individual record from representing more than one nonrespondent. The noninteger nonresponse weight for each farm was then rounded to either one or two. All farms classified as large based on total acreage, TVP, commodity production, or certain characteristics (i.e., value of land and buildings, value of machinery and equipment, farm-related income, number of workers, etc.); all tagged records; and all farms with uncommon commodities were assigned a nonresponse weight of one. This was because these records either required a 100% response rate or could not represent a nonrespondent. From the remaining records, a systematic sample of records was selected within each county/stratum for the integer rounding; the integer nonresponse weight was two for sample records and one for nonsample records.

The second weight is the sample item weight which accounts for both nonresponse and subsampling for data items that are only asked on the long form. Operationally, this weight was referred to as the sample weight; the author has expanded the name to avoid confusion with the standard definition of a sample weight. The sample item weight was calculated by multiplying the noninteger nonresponse weight by the "adjusted" sampling weight (the "adjusted" sampling weight uses only respondents that operated farms). Certainty farms were always assigned an adjusted sampling weight of one. The noncertainty farms were stratified based on TVP, total acreage, and Standard Industrial Classification code. Within each stratum, the adjusted sampling weight was calculated as the sum of the noninteger nonresponse weights for noncertainty farms (long-form noncertainty farms plus short-form noncertainty farms) divided by the sum of the noninteger nonresponse weights for long-form noncertainty farms. Strata were collapsed if the adjusted sampling weight was greater than twice the original sampling weight or if the sum of the noninteger nonresponse weights for long-form noncertainty records was less than 10. The noninteger sample item weight for each noncertainty farm was then rounded to an integer value. A systematic sample of records was selected within each county/nonresponse stratum/sample stratum for the integer rounding; the integer sample item weight was the next largest integer value for sample records and the truncated value for nonsample records.

As previously stated, the purpose of the nonresponse and sample item weights is to produce totals for the entire population (for more details on the weighting adjustments, refer to the memoranda written by Swan and Scholetzky). The nonresponse weight is used when generating totals for data items collected from all respondents (the question appears on both the long and short forms) and the sample item weight is used when generating totals for data items collected from a subset of all respondents (the question appears only on the long form).

The process of rounding weights to integer values has been in place for the last several censuses. In 1993, research was done to examine how this process affects the aggregates by analyzing county-level totals for three long-form characteristics in two states (Kraus-Winters, 1993). Although that paper concentrated more on examining the effect on the variance of the totals, the authors concluded "the results indicate that minor discrepancies at the strata level due to systematic rounding accumulate to sizable differences at the published level for some farm characteristics."

This paper represents the first assessment of the rounding process on totals computed for data items on both the long and short forms.

RESULTS AND DISCUSSION

Before proceeding, the author will return to the previous example to illustrate how rounding a record's weight to an integer value affects the data for the individual record. Suppose that the sample farm with 200 acres of cropland reported 100 acres of soybeans and 20 cattle (recall that the noninteger nonresponse weight is 1.2 and the integer nonresponse weight is 2). Using the noninteger nonresponse weight, the record represents 1.2 farms totaling 240 acres of cropland, 120 acres of soybeans, and 24 cattle. Using the integer nonresponse weight, which is used in producing the publication tables, the record represents 2 farms totaling 400 acres of cropland, 200 acres of soybeans, and 40 cattle. For this record, the difference between the totals calculated with the noninteger weight and the published totals are 0.8 farms, 160 acres of cropland, 80 acres of soybeans, and 16 cattle.

When a record's weight is rounded to an integer value, the totals represented by that record may or may not change dramatically. These changes may or may not become negligible when producing totals at the state or county level. The definition of what a person considers negligible is open to debate. One thing is clear - the difference must be considered in relation to the size of the total and the standard error of the total.

The following characteristics were considered in the analysis:
Value of Land and Buildings
Farm Production Expenditures
Total Value of Products Sold
Land in Farms
Acres of Harvested Cropland
Acres of Corn for Grain Harvested
Bushels of Corn for Grain Harvested
Acres of Wheat for Grain Harvested
Bushels of Wheat for Grain Harvested
Acres of Soybeans Harvested
Bushels of Soybeans Harvested
Acres of Cotton Harvested
Bales of Cotton Harvested
Total Cattle and Calves
Dollars Received for Dairy Products
Total Hogs and Pigs
Layers 20 Weeks Old and Older
Broilers and Other Meat-Type Chickens Sold
Farms Operated by Blacks and Other Races
Farms Operated by Females
Farms Operated by Persons of Hispanic Origin

The first two characteristics are long-form data items and the remaining characteristics are data items on both the long and short forms. In addition to these characteristics, the number of farms was examined. Number of farms is a unique variable since the data value is equal to one for every record. Therefore, in this case, the theory behind integer weighting dictates that totals produced using the noninteger weights are equal to totals produced using the integer weights. As expected, there was no difference between the number of farms calculated with the noninteger and integer nonresponse weights at either the state or county levels.

Note that the number of farms was also published using the integer sample item weight. The totals produced using the noninteger and integer nonresponse weights were equal to the number of farms calculated with the noninteger sample item weight (this is true because the "adjusted" sampling weight is used when calculating this weight). However, there was a difference between these totals and the number of farms calculated with the integer sample item weight. The author is not sure why this resulted; it could have occurred because of the small number of farms within a county/nonresponse stratum/sample stratum or because the rounding methodology wasn't consistent within the weighting program. Whatever the reason, the differences are not large compared to the number of farms in the state. The three largest differences were 53 farms in New York, 35 farms in West Virginia, and 25 farms in Maryland. At the county level, the differences were small. But, both the state and county differences could be troublesome to a data user. One example of this appears on page C-16 of Table F in the United States publication, where the two totals are adjacent to each other. A data user might improperly conclude that these differences were an error.

In this report, the author refers to the totals calculated with the integer weights as the published totals. This terminology is not literal; totals for the above characteristics do not appear in the census publication when the total was suppressed due to confidentiality concerns or the commodity was not published for that state. The analysis done for this report concentrated more on comparing the county-level totals than the state-level totals. This is because the differences are not as predominant at the state-level and because the primary purpose of the census is to produce county-level totals. The report first presents a brief discussion of the state-level totals and then a more in-depth analysis of the county-level totals.

Differences in Totals at the State Level

The t-value was used to evaluate the difference between the published and noninteger totals at the state level. A t-value was determined by subtracting the noninteger total from the published total and dividing this quantity by the standard error of the difference. The standard error was calculated using the formula described in Appendix A and the t-value was compared to a two-tailed Student t distribution. When the t-value for a characteristic exceeded the threshold for 90% significance, the difference in the state-level totals was defined to be "large." Table 1 lists the states that were determined to have large differences for each characteristic.

Table 1: State Totals with Statistically Significant Differences	
Characteristic	States
Value of Land and Buildings	CO, KY*, LA, MD*, MI*, MT, NC*, OK**, SC, VT
Farm Production Expenditures	ND
Total Value of Products Sold	AZ**
Land in Farms	CA*, GA*, OK
Acres of Harvested Cropland	AZ*, AR, CT, HI*, ME, NC*, SD*
Acres of Corn for Grain Harvested	CO, DE, KS, LA, NH, ND, VT**
Bushels of Corn for Grain Harvested	DE*, KS, NH, ND, VT**, WY*
Acres of Wheat for Grain Harvested	IA, MS, NV*, NY**
Bushels of Wheat for Grain Harvested	FL, IA, MS, NV**, NY**
Acres of Soybeans Harvested	AL, OK
Bushels of Soybeans Harvested	None
Acres of Cotton Harvested	MO
Bales of Cotton Harvested	MO*, NM
Total Cattle and Calves	KY, VT*, WI
Dollars Received for Dairy Products	DE, IL*, LA**, MS*, UT, WY*
Total Hogs and Pigs	LA**, NM*, OH, TN**
Layers 20 Weeks Old and Older	MD, NH*, PA*, VT*, WI
Broilers and Other Meat-Type Chickens Sold	KS**, ME, WI*
Farms Operated by Blacks and Other Races	AZ*, KY*, NM
Farms Operated by Females	CO*, GA, HI*, ME, MN*, ND**, SD*
Farms Operated by Persons of Hispanic Origin	IN, KY*, MD*, MI**, MN*, NM, OH**, OR*, WI*, WY*

* Greater than or equal to a 95% significance level.
** Greater than or equal to a 99% significance level.

There were several situations where the t-value was just less than a 90% significance level. Since state totals for these are large enough to be potentially of practical importance, the author also considers the difference in the state-level totals to be large for: total value of products sold in SD with a t-value of 1.631, acres of corn for grain harvested in AR with a t-value of 1.617, acres of wheat for grain harvested in GA and NE with t-values of 1.627 and 1.611, bushels of wheat for grain harvested in MI with a t-value of 1.626, acres of soybeans harvested in PA with a t-value of 1.618, acres of soybeans harvested in PA with a t-value of 1.625, total cattle and calves in FL with a t-value of 1.632, and total hogs and pigs in AZ with a t-value of 1.644.

Interpretation of the results presented in Table 1 depends on the reader. For example, a person may not be concerned with a significant difference in a state when the characteristic is a small percentage of the national total. On the other hand, another person may be concerned that the process of rounding weights results in a significant difference for any characteristic at the state level. The difference and percent difference can be used to help evaluate the importance of the statistical significance. The percent difference is computed by subtracting the noninteger total from the published total, dividing by the noninteger total, and converting this to a percentage. To illustrate this methodology, the author will discuss two states which were significant at the 99% level.

For acres of corn for grain harvested in Vermont, the difference was -147 acres and the percent difference was -1.75 percent. For broilers and other meat-type chickens sold in Kansas, the difference was -3,507 chickens and the percent difference was -9.10 percent. A listing of all state/characteristic combinations that are statistically significant at the 90% level is provided in Appendix B.

Although not specifically presented in a table, it should be noted that it is possible for a state to have a large percent difference between the published and noninteger totals that is not statistically significant. For example, for acres of corn for grain in Nevada, the difference was 50 acres and the percent difference was 14.45 but the t-value was 0.901. For broilers and other meat-type chickens sold in Colorado, the difference was 2,082 and the percent difference was 21.13 but the t-value was 0.917.

Differences in Totals at the County Level

As stated earlier, a more comprehensive analysis was performed on totals at the county level. The analysis was performed on all 21 characteristics for every county in the 48 contiguous states. "Small" counties were identified for each characteristic and excluded from the analysis. Originally, a county was defined as "small" when the total for the characteristic was less than 1% of the state total. However, this definition resulted in eliminating too many counties in states with a large number of counties (i.e., Texas). In order to account for the variation in the number of counties among states, the concept of "expected county contribution" was adopted. Expected county contribution is defined as the multiplicative inverse of the state's number of counties times the state total for the characteristic. So, a county was defined as "small" when it contained less than the expected county contribution. For example, if a county in Texas contained 20,000 acres of cotton, then it was considered to be small since 20,000 is less than the inverse of the number of counties (254^{-1}) times the total number of acres of cotton in Texas (5,221,561). Using this procedure, the definition of a small county for number of acres of cotton in Texas was 20,557 rather than the original definition of 52,215. In addition to the concept of expected county contribution, for the three demographic variables, a county was determined to be "small" when the county's sum of the integer nonresponse weights was less than three.

After eliminating small counties from the analysis, two statistics were used to evaluate the difference between the published and noninteger totals. First, a t-value was calculated by subtracting the noninteger total from the published total and dividing this quantity by the standard error of the difference. Like the state-level analysis, the standard error was calculated using the formula described in Appendix A, and the t-value was compared to a two-tailed Student t distribution. When the t-value for a characteristic exceeded the threshold for 90% significance, the difference in the county-level totals was defined to be "large." Second, a percent difference was computed by subtracting the noninteger total from the published total, dividing by the noninteger total, and converting this to a percentage. Unlike the t-value, this method did not use a specific cutoff to define a "large" difference in the county-level totals. Instead, the analysis focused on counties with the largest percent differences for each characteristic.

Various tables and maps were created in an effort to show the key results of the analysis. Results from each method are presented as well as results from combining the methods. While interpretation of these results depends on the reader, the author has attempted to present an unbiased summarization of the analysis.

Table 2 presents the number of nonsmall counties with statistically significant differences. For each characteristic, the table shows the number of nonsmall counties that were significant at the three most commonly-used levels and the percentage of statistically significant nonsmall counties at the 90% level.

The three significance levels are shown to give a measure of how different the published total is from the noninteger total. The number of counties at the 95% and 99% significance levels are subsets of the previous significance level. For example, 100 counties are significant at the 90% level for value of land and buildings; 44 of these 100 counties are significant at the 95% level. The last column, the percentage with significant differences, is shown to give a measure of the number of differences. This percentage was calculated by dividing the number of nonsmall counties that were significant at the 90% level (shown in column three) by the total number of nonsmall counties for the characteristic (shown in column two) and multiplying by 100. Farms operated by persons of Hispanic origin had the largest percentage of significant counties and farm production expenditures had the smallest. The two long-form characteristics performed much better than the characteristics collected on both the long and short forms. This is not surprising since the process of rounding weights has less of an effect on the totals as the weight increases, and the sample item weights

Table 2: Number of Nonsmall Counties with Statistically Significant Differences

Characteristic	Number of Nonsmall Counties	Significance Level			% with Significant Differences
		$ 90%	$ 95%	$ 99%	

Value of Land and Buildings	1285	100	44	13	7.78
Farm Production Expenditures	1033	23	15	5	2.23
Total Value of Products Sold	1073	125	77	19	11.65
Land in Farms	1307	133	71	11	10.18
Acres of Harvested Cropland	1160	126	65	15	10.86
Acres of Corn for Grain Harvested	932	131	76	29	14.06
Bushels of Corn for Grain Harvested	885	129	76	23	14.58
Acres of Wheat for Grain Harvested	853	111	56	8	13.01
Bushels of Wheat for Grain Harvested	824	111	52	10	13.47
Acres of Soybeans Harvested	849	91	51	13	10.72
Bushels of Soybeans Harvested	825	96	53	11	11.64
Acres of Cotton Harvested	282	37	17	4	13.12
Bales of Cotton Harvested	276	35	13	2	12.68
Total Cattle and Calves	1170	114	66	23	9.74
Dollars Received for Dairy Products	713	113	56	19	15.85
Total Hogs and Pigs	706	111	55	18	15.72
Layers 20 Weeks Old and Older	362	57	30	6	15.75
Broilers and Other Meat-Type Chickens Sold	296	25	9	0	8.45
Farms Operated by Blacks and Other Races	644	130	87	28	20.19
Farms Operated by Females	1073	148	83	20	13.79
Farms Operated by Persons of Hispanic Origin	643	199	103	18	30.95

are larger in magnitude than the nonresponse weights.

An interesting subset of the counties tallied in Table 2 is the collection of those counties that represented a substantial percentage of the state total. Table 3 provides the number of nonsmall counties with statistically significant differences, where the county represented more than 5 percent of the state total. Again, for each characteristic, the table shows the number of nonsmall counties that were significant at the three most commonly-used levels and the percentage of statistically significant nonsmall counties at the 90% level.

The explanation of the columns in this table is the same as for Table 2, with the exception of the last column. The percentage of significant counties was calculated by dividing the number of nonsmall counties that were significant at the 90% level and contributed more than 5 percent of the state total (shown in column three) by the total number of nonsmall counties for the characteristic that represent more than 5 percent of the state total (shown in column two) and multiplying by 100. These percentages can be compared to the percentages in Table 2 to evaluate whether or not the significant differences occurred in more prominent counties. For example, acres of corn for grain harvested increased from

Table 3: Number of Nonsmall Counties with Statistically Significant Differences that Represent More Than 5 Percent of the State Total

Characteristic	Number of Counties (nonsmall and $5%)	Significance Level			% with Significant Differences
		$ 90%	$ 95%	$ 99%	
Value of Land and Buildings	115	2	0	0	1.74
Farm Production Expenditures	127	0	0	0	0.00
Total Value of Products Sold	136	15	9	2	11.03
Land in Farms	106	12	5	2	11.32
Acres of Harvested Cropland	126	11	7	2	8.73
Acres of Corn for Grain Harvested	159	33	19	8	20.75
Bushels of Corn for Grain Harvested	176	27	16	4	15.34
Acres of Wheat for Grain Harvested	187	16	9	2	8.56
Bushels of Wheat for Grain Harvested	180	15	10	2	8.33
Acres of Soybeans Harvested	121	11	7	1	9.09
Bushels of Soybeans Harvested	121	18	11	2	14.88
Acres of Cotton Harvested	94	12	7	2	12.77
Bales of Cotton Harvested	90	12	6	1	13.33
Total Cattle and Calves	126	15	10	4	11.90
Dollars Received for Dairy Products	182	36	18	3	19.78
Total Hogs and Pigs	178	26	12	5	14.61
Layers 20 Weeks Old and Older	151	24	13	2	15.89
Broilers and Other Meat-Type Chickens Sold	137	14	5	0	10.22
Farms Operated by Blacks and Other Races	102	18	11	4	17.65
Farms Operated by Females	101	9	4	1	8.91
Farms Operated by Persons of Hispanic Origin	91	16	10	6	17.58

14.06 to 20.75 percent while farms operated by persons of Hispanic origin decreased from 30.95 to 17.58 percent.

County-level maps of the absolute percent differences for the 48 contiguous states were produced to analyze the differences for each characteristic. The absolute percent difference was used to decrease the number of ranges. Maps were used as the data analysis tool since map presentation was the most efficient method for displaying a lot of information (i.e., 3,069 counties and 21 characteristics). The maps are included in Appendix C. The ranges are the same for characteristics with a common theme (i.e., corn, wheat, soybeans, and cotton acreages and production have the same ranges). Two maps containing different ranges are displayed for acres of harvested cropland so this characteristic can be compared to individual crop acreages as well as other key characteristics.

The county-level percent differences were the smallest for the two long-form characteristics and total value of products sold. Land in farms and acres of harvested cropland had more counties with a percent difference in the top two ranges than did these three characteristics. The maps looked reasonable for the crop characteristics, but the ranges were also higher than those for the other characteristics. The maps for the livestock characteristics were extremely different from the crop characteristics. With the exception of cattle and calves, there were more counties in the top two ranges and fewer counties in the bottom two ranges. With the exception of farms operated by females, the maps for the demographic characteristics showed the most counties in the top two ranges, even with the increase in the ranges.

To supplement the maps, Table 4 contains the number of nonsmall counties that had a percent difference in the top two ranges. For each characteristic, the table shows the total number of nonsmall counties and the number of nonsmall counties that were statistically significant at each level.

Table 4: Number of Nonsmall Counties with a Percent Difference in the Top Two Ranges

Characteristic	Total	Significance Level			
		< 90%	$ 90%	$ 95%	$ 99%
Value of Land and Buildings	6	5	1	1	1
Farm Production Expenditures	0	0	0	0	0
Total Value of Products Sold	1	0	1	1	0
Land in Farms	21	4	17	12	4
Acres of Harvested Cropland	15	4	11	7	2
Acres of Corn for Grain Harvested	21	11	10	5	1
Bushels of Corn for Grain Harvested	15	11	4	4	1
Acres of Wheat for Grain Harvested	34	25	9	5	1
Bushels of Wheat for Grain Harvested	28	23	5	3	1
Acres of Soybeans Harvested	35	30	5	2	1
Bushels of Soybeans Harvested	31	27	4	1	1
Acres of Cotton Harvested	2	2	0	0	0
Bales of Cotton Harvested	2	2	0	0	0
Total Cattle and Calves	2	1	1	1	1
Dollars Received for Dairy Products	30	17	13	7	1
Total Hogs and Pigs	68	61	7	4	2
Layers 20 Weeks Old and Older	51	44	7	5	0
Broilers and Other Meat-Type Chickens Sold	53	51	2	1	0
Farms Operated by Blacks and Other Races	117	100	17	3	0
Farms Operated by Females	1	0	1	1	0
Farms Operated by Persons of Hispanic Origin	305	228	77	16	0

The reader must be careful when using Table 4 to make comparisons across characteristics, since the top two ranges vary by characteristic. The top two ranges contain percent differences greater than or equal to 2% for the first five characteristics, greater than or equal to 20% for the last three characteristics, and greater than or equal to 5% for the remaining characteristics. Also, the number of nonsmall counties that were not significant at the 90% level is included in column three of this table. This column plus the number of nonsmall counties that were significant at the 90% level (shown in column four) equals the total number of nonsmall counties in column two. Again, the number of counties at the 95% and 99% significance levels are subsets of the previous significance level. A listing of all 192 nonsmall county/characteristic combinations that are statistically significant at the 90% level is provided in Appendix D. The listing contains the percent difference and t-value as well as the percentage of the state total that the county represents.

Table 4 can be compared to Table 2 to evaluate whether or not the significant differences occurred in the top two ranges for the characteristics. For example, Table 2 shows that the percentage of significant counties out of all nonsmall counties for land in farms was 10.18 percent while Table 4 shows the percentage of significant counties out of all nonsmall counties with a percent difference greater than or equal to 2% was 80.95 percent (17 divided by 21). On the other hand, the same comparison for farms operated by persons of Hispanic origin shows that the percentage of significant counties was 30.95 percent in Table 2 and 25.25 percent in Table 4.

A miscellaneous observation for Appendix D is that Screven, GA had seven characteristics with a percent difference in the top two ranges (the next closest county had four). For each characteristic, Screven represented between 1 and 4 percent of the state total. All characteristics except for one were statistically significant at the 95% level.

Again, an interesting subset of the counties tallied in Table 4 is the collection of those counties that represented a substantial percentage of the state total. Table 5 provides a listing of the counties that had a percent difference in the top two ranges, and where the county represented more than 5 percent of the state total. Since this subset is only 16 counties, the county names are given with the associated percent difference, percentage of state total, and t-value (note that these counties are also listed in Appendix D). If the characteristic does not appear in the table, then there were no counties that met these criteria.

Although the same counties appear for acres and bushels of corn for grain harvested, this was not always the case when examining harvested acres and production for a crop. For example, for acres and bushels of wheat for grain harvested, Appendix D shows that Peach, GA was statistically significant for acres but not bushels while Morgan, GA was

Table 5: Nonsmall Counties with a Percent Difference in the Top Two Ranges that Represent More Than 5 Percent of the State Total					
Characteristic	State Name	County Name	Percent Difference	% of State	t-value
Acres of Corn for Grain Harvested	Montana	Rosebud	9.89	5.50	2.203
	Vermont	Rutland	5.91	8.51	2.239
Bushels of Corn for Grain Harvested	Montana	Rosebud	8.77	5.16	2.243
	Vermont	Rutland	5.03	9.07	2.271
Dollars Received for Dairy Products	Kansas	Sedgwick	5.99	5.73	2.138
	Wyoming	Lincoln	5.25	42.20	2.039
Total Hogs and Pigs	New Mexico	Roosevelt	11.55	6.26	2.514
Layers 20 Weeks Old and Older	Mississippi	Leake	6.30	7.76	2.005
	Mississippi	Simpson	6.29	5.23	2.196
	Nevada	Washoe	11.55	12.23	2.071
Broilers and Other Meat-Type Chickens Sold	Colorado	Larimer	8.47	10.32	2.162
	Kansas	Reno	9.84	16.46	1.796
Farms Operated by Blacks and Other Races	Massachusetts	Barnstable	20.00	11.11	1.921
Farms Operated by Persons of Hispanic Origin	Maine	Penobscot	20.00	11.11	1.710
	Maryland	Frederick	27.27	16.47	1.677
	Wyoming	Fremont	21.05	17.56	2.067

statistically significant for bushels but not acres. For the two demographic variables, the percent difference for Barnstable, MA and Penobscot, ME are a little misleading. These counties were not small enough to be eliminated from the analysis, but the weighted number of farms is only 4 (not true for the other two counties). Note that some of these counties contributed substantially to the state-level totals being statistically significant. For example, for dollars received for dairy products, Lincoln, WY represented 42.20 percent of the state total and was statistically significant at the 95% level. Table 1 shows that Wyoming is statistically significant at the 95% level for this characteristic.

An Alternative Weighting Procedure

As previously stated, one of the two main reasons for using integer weights is to make the publication process simpler. Record-level integer weights ensure the table cells within a column or row add to the total for the column or row. However, there is another possible way to address this concern. Rather than round the noninteger weights to integer values, another possible approach is to apply the noninteger weights to the record and round the weighted data values at the record level. Thus, the operation's weighted data values are rounded and these values are used to produce the totals for the entire population.

To illustrate, recall the previous example where the farm had 200 acres of cropland, reported 100 acres of soybeans, and 20 cattle. Using the noninteger nonresponse weight of 1.2, the record represented 1.2 farms totaling 240 acres of cropland, 120 acres of soybeans, and 24 cattle. With the exception of 1.2 farms, no rounding is required because the noninteger weight multiplied by the individual data values are already integer values. This is not a common occurrence in practice. A more applicable example would be that the noninteger nonresponse weight for this record is a noninteger value rounded to six decimal places, say 1.212684. For this example, the record's weighted data values are 1.212684 farms totaling 242.5368 acres of cropland, 121.2684 acres of soybeans, and 24.25368 cattle. After the rounding these data values, the record represents 1 farm totaling 243 acres of cropland, 121 acres of soybeans, and 24 cattle. For this record, the differences between the totals calculated with the noninteger weight and the record's rounded-weighted data value are minimal.

Totals produced using the record's rounded-weighted data values are unbiased estimates of totals produced using the noninteger weights. This procedure performs better than the integer weighting procedure because the two factors mentioned earlier do not affect this procedure (not having a representative sample or the distribution of the data being skewed). However, the downfall of using this procedure is that it does not work well when producing totals for indicator or categorical characteristics. For example, the average noninteger nonresponse weight is approximately 1.139. Therefore, for variables such as number of farms and the demographic characteristics (Blacks and Other Races, Females, Hispanic Origin), the record's rounded-weighted data value will round down to one most of the time and thus underestimate the total.

To examine this methodology on 1997 census data, the author selected three states (Georgia, Texas, and Virginia) which showed fairly large differences between totals calculated with the noninteger weights and the published totals. For each characteristic, the difference between the totals calculated with the noninteger and integer weights is compared to the difference between the totals calculated with the noninteger weights and rounded-weighted data values. This analysis was done at both the state and county levels. An overall summary of these differences is presented in Appendix E. The table shows the differences at the state level along with the minimum and maximum differences at the county level. As expected, these results show that the totals produced using the rounded-weighted data values are more precise than the totals produced using the integer weights except for the demographic characteristics and number of farms.

RECOMMENDATIONS

The reasons for rounding weights, to ease data review procedures and to ensure that publication totals add, are legitimate concerns. The author asserts that it is possible to address these two concerns and improve the totals produced when NASS revamps the census processing system for the 2002 census.

1) *Automate the weights into the data review procedures.*

   One argument for using the integer weights is that one can "easily" obtain weighted totals for a record of interest during the data review phase. In the 1997 system, this would be accomplished by multiplying the integer weight by each record's data value of interest. Thus, this manual computation is easier if the nonresponse weight is 2 rather than 1.7. However, for the 2002 system, the computer can and should perform this calculation. With this improvement, the value of the weight is no longer relevant.

2) *Use a combination of the integer weights and the record's rounded-weighted data values.*

   Another argument for using the integer weights is to ensure that publication totals add. Integer weights should be used to produce totals for indicator and categorical variables. These variables include number of farms as well as variables for the questions concerning type of organization, corporate structure, and characteristics and occupation of the operator (which includes demographic characteristics). In addition, the integer weights should be used for any question where the data values are small for most farms. For the 1997 census, the only example of this is the question concerning injuries or deaths. The record's rounded-weighted data values should be used to produce totals for all other characteristics. Here, the noninteger weight is applied to the record and the weighted data values are rounded at the record level. With this improvement, the integer sample item weight should no longer be necessary since the 1997 long form did not contain a question that requires it.

3) *Evaluate the 2002 questionnaires to determine which weight to use.*

   Once the 2002 long and short forms are finalized, the questions will need to be evaluated to determine which weight to use during the summarization process. In particular, the questions on the 2002 long form will determine whether the integer sample item weight is needed.

REFERENCES

Swan, Carolyn, "1997 Nonresponse Weighting: Nonresponse Survey Sampling Specification," NASS Internal Memo, April 1999.

Scholetzky, Wendy, "1997 Nonresponse Weighting: Creation of Final Nonresponse Weighting Short Record File," NASS Internal Memo, April 1999.

Swan, Carolyn, "Estimation of In-Scope Nonrespondents and Calculation of Initial Noninteger Nonresponse Weights at County/Nonresponse Stratum Level," NASS Internal Memo, April 1999.

Scholetzky, Wendy, "1997 Nonresponse Weighting: Final Nonresponse Weighting Specification," NASS Internal Memo, April 1999.

Scholetzky, Wendy, "1997 Census of Agriculture - Sample Weighting Specifications," NASS Internal Memo, April 1999.

Kraus, Melinda and Winters, Franklin, "An Evaluation of Rounding Techniques Used to Obtain Estimated Totals," American Statistical Association, 1993 Proceedings of the International Conference on Establishment Surveys, pp. 893-898.

APPENDIX A. FORMULA FOR CALCULATING THE STANDARD ERROR OF THE DIFFERENCE

The following formula was used for calculating the standard error of the difference between the totals generated using integer and noninteger weights:

$$\text{Standard Error of Difference} = \sqrt{\sum_{j=1}^{J} \sum_{h=1}^{H} \frac{n_{hj}}{n_{hj} - 1} \sum_{i=1}^{n_{hj}} \left[w_{ihj}(y_{ihj} - \bar{y}_{h(w)}) - a_{ihj}(y_{ihj} - \bar{y}_{h(a)})\right]^2}$$

where

n_{hj} = the number of in-scope, interviewed records in stratum h in county j

w_{ihj} = the noninteger weight associated with record i in stratum h in county j

a_{ihj} = the integer weight associated with record i in stratum h in county j

y_{ihj} = the unweighted data value for record i in stratum h in county j

$\bar{y}_{h(w)}$ = the weighted average total for stratum h in county j using the noninteger weights

$\bar{y}_{h(a)}$ = the weighted average total for stratum h in county j using the integer weights

H = the number of nonresponse strata in the county for characteristics on both the long and short forms
or
= the number of sample strata in the county for long-form characteristics

J = the number of counties in the state for state-level totals
or
= 1 for county-level totals

APPENDIX B. LISTING OF STATES WITH STATISTICALLY SIGNIFICANT DIFFERENCES AT THE 90% LEVEL

Characteristic	State Name	Percent Difference	Difference	t-value
Value of Land and Buildings	Colorado	0.27	53,982,465	1.789
	Kentucky	-0.09	-17,501,972	-1.960
	Louisiana	0.21	19,204,512	1.790
	Maryland	0.25	16,990,601	2.036
	Michigan	0.14	22,722,570	2.024
	Montana	-0.15	-25,874,963	-1.712
	North Carolina	-0.14	-26,145,535	-2.259
	Oklahoma	-0.18	-35,755,465	-2.675
	South Carolina	-0.22	-14,636,941	-1.907
	Vermont	-0.36	-6,858,242	-1.869
Farm Production Expenditures	North Dakota	-0.05	-1,187,380	-1.732
Total Value of Products Sold	Arizona	-0.11	-2,180,052	-3.358
Land in Farms	California	0.40	111,717	2.287
	Georgia	-0.25	-26,798	-2.327
	Oklahoma	0.12	41,009	1.673
Acres of Harvested Cropland	Arizona	-0.26	-2,492	-2.440
	Arkansas	0.13	10,164	1.941
	Connecticut	-0.56	-793	-1.687
	Hawaii	-0.21	-213	-2.091
	Maine	-0.45	-1,804	-1.745
	North Carolina	-0.18	-7,833	-1.964
	South Dakota	-0.15	-21,711	-2.002
Acres of Corn for Grain Harvested	Colorado	-0.37	-3,397	-1.814
	Delaware	0.63	989	1.708
	Kansas	-0.20	-4,976	-1.646
	Louisiana	-0.42	-1,725	-1.898
	New Hampshire	-1.46	-18	-1.800
	North Dakota	-0.35	-2,044	-1.659
	Vermont	-1.75	-147	-2.940

Characteristic	State Name	Percent Difference	Difference	t-value
Bushels of Corn for Grain Harvested	Delaware	0.81	126,614	2.208
	Kansas	-0.21	-748,562	-1.807
	New Hampshire	-1.26	-1,622	-1.816
	North Dakota	-0.37	-203,715	-1.747
	Vermont	-1.58	-15,061	-2.791
	Wyoming	-1.58	-100,573	-1.979
Acres of Wheat for Grain Harvested	Iowa	-2.36	-535	-1.701
	Mississippi	-0.58	-901	-1.818
	Nevada	-1.28	-247	-2.150
	New York	-0.93	-1,137	-3.118
Bushels of Wheat for Grain Harvested	Florida	-2.35	-14,088	-1.716
	Iowa	-2.61	-24,231	-1.766
	Mississippi	-0.49	-31,983	-1.713
	Nevada	-1.11	-21,468	-2.816
	New York	-0.72	-46,268	-2.710
Acres of Soybeans Harvested	Alabama	0.81	2,537	1.731
	Oklahoma	-0.81	-2,654	-1.879
Acres of Cotton Harvested	Missouri	-0.37	-1,457	-1.661
Bales of Cotton Harvested	Missouri	-0.43	-2,401	-2.125
	New Mexico	-0.66	-754	-1.895
Total Cattle and Calves	Kentucky	-0.23	-5,500	-1.736
	Vermont	0.48	1,483	2.292
	Wisconsin	-0.19	-6,492	-1.793
Dollars Received for Dairy Products	Delaware	1.32	250,408	1.666
	Illinois	1.33	3,329,056	2.359
	Louisiana	1.21	1,243,682	2.655
	Mississippi	-0.97	-816,420	-2.549
	Utah	0.25	496,490	1.675
	Wyoming	-2.59	-262,555	-2.140
Total Hogs and Pigs	Louisiana	-2.11	-439	-2.905
	New Mexico	-5.08	-327	-2.399
	Ohio	-0.43	-7,418	-1.773
	Tennessee	-1.21	-3,933	-4.743

Characteristic	State Name	Percent Difference	Difference	t-value
Layers 20 Weeks Old and Older	Maryland	0.81	33,018	1.645
	New Hampshire	-0.39	-721	-2.077
	Pennsylvania	0.24	57,715	2.064
	Vermont	0.34	864	1.965
	Wisconsin	-0.32	-12,011	-1.804
Broilers and Other Meat-Type Chickens Sold	Kansas	-9.10	-3,507	-2.627
	Maine	-0.40	-807	-1.699
	Wisconsin	-0.13	-36,844	-2.098
Farms Operated by Blacks and Other Races	Arizona	-2.63	-14	-2.490
	Kentucky	2.92	20	2.168
	New Mexico	0.99	21	1.681
Farms Operated by Females	Colorado	1.04	33	2.034
	Georgia	-0.75	-32	-1.726
	Hawaii	2.33	21	2.167
	Maine	1.99	16	1.911
	Minnesota	-1.09	-40	-2.173
	North Dakota	2.85	37	2.875
	South Dakota	-1.60	-24	-2.360
Farms Operated by Persons of Hispanic Origin	Indiana	-2.93	-7	-1.853
	Kentucky	3.58	14	2.054
	Maryland	8.97	7	2.195
	Michigan	-3.78	-11	-3.036
	Minnesota	4.84	12	2.166
	New Mexico	-0.69	-24	-1.730
	Ohio	-3.77	-12	-2.750
	Oregon	-2.48	-13	-2.190
	Wisconsin	-3.46	-9	-2.054
	Wyoming	6.50	8	2.122

APPENDIX C. COUNTY-LEVEL MAPS OF PERCENT DIFFERENCES

APPENDIX D. LISTING OF COUNTIES WITH PERCENT DIFFERENCES IN THE TOP TWO RANGES AND STATISTICALLY SIGNIFICANT DIFFERENCES AT THE 90% LEVEL

Characteristic	State Name	County Name	Percent Difference	% of State	t-value
Value of Land and Buildings	Georgia	Oglethorpe	2.01	0.69	2.600
Total Value of Products Sold	Tennessee	Williamson	2.12	1.32	2.386
Land in Farms	Alabama	Coffee	2.29	2.15	2.004
	Alabama	Monroe	3.28	1.55	2.107
	California	Humboldt	5.77	2.11	2.426
	California	Mendocino	3.76	2.31	1.777
	Georgia	Bacon	2.51	0.66	1.831
	Georgia	Franklin	-3.90	0.72	-3.168
	Georgia	Jackson	-3.04	0.73	-1.867
	Georgia	Jeff Davis	-2.65	0.67	-1.925
	Missouri	Laclede	2.17	1.10	2.395
	Montana	Cascade	-2.02	2.46	-3.877
	Nebraska	Buffalo	2.00	1.36	2.336
	Nebraska	Rock	2.02	1.39	1.999
	North Carolina	Franklin	2.15	1.50	2.230
	Ohio	Clinton	2.13	1.58	2.683
	Oregon	Gilliam	2.72	4.26	1.862
	Texas	Live Oak	3.41	0.40	2.257
	West Virginia	Wood	-2.09	1.93	-3.346
Acres of Harvested Cropland	Georgia	Screven	2.65	1.77	2.628
	Kentucky	Barren	2.03	1.83	2.427
	Kentucky	Butler	3.31	0.92	2.198
	New York	Delaware	2.29	1.84	2.161
	North Carolina	Rowan	2.17	1.00	1.690
	Pennsylvania	Centre	3.12	1.94	2.950
	Pennsylvania	Columbia	2.15	1.62	1.920
	Tennessee	Williamson	2.55	1.30	1.786
	Texas	Baylor	2.22	0.51	1.743
	Texas	Haskell	-2.04	1.02	-2.492

Characteristic	State Name	County Name	Percent Difference	% of State	t-value
	West Virginia	Wood	-2.41	2.02	-2.498
Acres of Corn for Grain Harvested	Alabama	Marion	8.52	1.66	1.914
	Florida	Walton	8.57	2.06	1.705
	Georgia	Jeff Davis	-5.87	0.97	-1.668
	Georgia	Screven	6.37	2.56	2.947
	Montana	Rosebud	-9.89	5.50	-2.203
	Texas	Bexar	-5.33	0.66	-2.226
	Texas	Cameron	5.93	0.85	1.976
	Vermont	Rutland	-5.91	8.51	-2.239
	Virginia	Lancaster	6.36	1.31	1.727
	Virginia	Page	5.90	1.07	1.835
Bushels of Corn for Grain Harvested	Georgia	Screven	5.92	2.31	2.673
	Montana	Rosebud	-8.77	5.16	-2.243
	Texas	Cameron	6.51	0.74	2.071
	Vermont	Rutland	-5.03	9.07	-2.271
Acres of Wheat for Grain Harvested	Florida	Santa Rosa	-5.40	3.99	-2.076
	Georgia	Peach	-7.01	0.93	-1.717
	Georgia	Screven	6.08	1.12	2.038
	Georgia	Taylor	-5.14	0.65	-1.727
	Indiana	Randolph	6.61	1.73	2.991
	Iowa	Benton	-10.91	1.11	-1.683
	Iowa	Linn	-14.35	1.86	-1.741
	Iowa	Page	-10.14	1.76	-2.371
	Virginia	Middlesex	6.30	1.60	2.157
Bushels of Wheat for Grain Harvested	Georgia	Morgan	-5.76	0.74	-1.901
	Georgia	Screven	5.24	1.21	2.115
	Indiana	Randolph	6.77	1.81	2.970
	Iowa	Benton	-11.00	1.05	-1.682
	Iowa	Page	-9.62	1.52	-2.414
Acres of Soybeans Harvested	Georgia	Screven	5.15	4.11	2.311
	Georgia	Telfair	-12.72	0.66	-1.807
	Pennsylvania	Snyder	5.15	1.72	1.676
	Texas	Red River	10.27	4.56	1.725
	Virginia	Middlesex	6.53	1.69	2.625

Characteristic	State Name	County Name	Percent Difference	% of State	t-value
Bushels of Soybeans Harvested	Florida	Holmes	-6.98	2.72	-1.694
	Georgia	Telfair	-20.75	0.65	-1.732
	Pennsylvania	Snyder	6.44	1.88	1.821
	Virginia	Middlesex	7.29	1.34	3.267
Total Cattle and Calves	Georgia	Thomas	-5.94	0.73	-2.916
Dollars Received for Dairy Products	Illinois	Douglas	16.99	1.09	1.734
	Illinois	Ogle	10.64	1.95	1.707
	Kansas	Sedgwick	5.99	5.73	2.138
	Kentucky	Edmonson	9.20	0.88	1.683
	Mississippi	Marion	8.57	4.55	1.659
	Mississippi	Pearl River	-5.47	1.39	-1.972
	Nebraska	Boone	-6.05	1.39	-1.706
	Nebraska	Saunders	-10.25	1.17	-2.086
	South Dakota	Bon Homme	-6.42	1.85	-3.146
	South Dakota	Gregory	9.19	2.80	1.788
	Tennessee	Williamson	6.86	1.84	2.005
	West Virginia	Ohio	-9.18	3.17	-2.364
	Wyoming	Lincoln	-5.25	42.20	-2.039
Total Hogs and Pigs	Alabama	Madison	-6.67	1.73	-1.765
	Florida	Hernando	-12.15	1.74	-1.703
	Idaho	Gem	-5.29	2.35	-1.793
	Louisiana	Calcasieu	-7.86	2.88	-2.654
	New Mexico	Roosevelt	-11.55	6.26	-2.514
	Texas	Fayette	-5.00	1.07	-2.102
	Wisconsin	Outagamie	-5.34	1.79	-3.478
Layers 20 Weeks Old and Older	Alabama	Cleburne	-6.78	1.60	-1.970
	Arkansas	Yell	-7.18	2.63	-1.767
	Mississippi	Jefferson Davis	-17.64	2.56	-2.003
	Mississippi	Leake	-6.30	7.76	-2.005
	Mississippi	Simpson	-6.29	5.23	-2.196
	Missouri	Pettis	-8.06	1.88	-1.829
	Nevada	Washoe	-11.55	12.23	-2.071

Characteristic	State Name	County Name	Percent Difference	% of State	t-value
Broilers and Other Meat-Type Chickens Sold	Colorado	Larimer	-8.47	10.32	-2.162
	Kansas	Reno	-9.84	16.46	-1.796
Farms Operated by Blacks and Other Races	Arkansas	Clark	25.00	2.14	1.975
	Georgia	Baldwin	28.57	0.68	1.734
	Georgia	Monroe	37.50	0.83	1.965
	Indiana	Allen	-25.00	1.74	-1.725
	Indiana	Lake	-25.00	1.74	-1.645
	Indiana	Putnam	-20.00	2.33	-1.927
	Iowa	Lee	-20.00	3.25	-1.727
	Massachusetts	Barnstable	-20.00	11.11	-1.921
	Minnesota	Clearwater	-20.00	2.04	-1.863
	Minnesota	Morrison	-20.00	2.04	-1.944
	Minnesota	Otter Tail	40.00	3.57	1.663
	Nebraska	Douglas	-25.00	1.59	-1.649
	Nebraska	Sioux	-25.00	1.59	-1.741
	Nebraska	Thurston	-20.00	4.23	-2.105
	Ohio	Ashtabula	37.50	3.45	1.806
	Wisconsin	Pierce	-20.00	2.17	-1.956
	Wisconsin	Washburn	-25.00	1.63	-1.646
Farms Operated by Females	North Dakota	Pierce	20.83	2.17	2.105
Farms Operated by Persons of Hispanic Origin	Alabama	Jackson	-25.00	1.61	-1.683
	Alabama	Shelby	-20.00	2.15	-1.653
	Alabama	Washington	-20.00	2.15	-1.946
	Georgia	Cobb	-25.00	0.96	-1.971
	Georgia	Turner	-25.00	0.96	-1.717
	Illinois	Bureau	-20.00	1.38	-1.961
	Illinois	Calhoun	-25.00	1.04	-1.870
	Illinois	Champaign	-20.00	1.38	-1.976
	Illinois	Effingham	-20.00	1.38	-1.726
	Illinois	Fayette	-25.00	1.04	-1.726
	Illinois	Macoupin	-20.00	1.38	-1.666
	Illinois	Peoria	-20.00	1.38	-1.984

Characteristic	State Name	County Name	Percent Difference	% of State	t-value
	Illinois	Pike	-20.00	1.38	-1.941
	Illinois	Shelby	-20.00	1.38	-1.777
	Indiana	Allen	-25.00	1.29	-1.719
	Indiana	Morgan	-25.00	1.29	-1.700
	Indiana	Ripley	-20.00	1.72	-1.895
	Indiana	Switzerland	-20.00	1.72	-1.764
	Iowa	Butler	-20.00	1.17	-1.753
	Iowa	Harrison	-20.00	1.17	-1.698
	Iowa	Mills	-20.00	1.17	-1.776
	Iowa	Taylor	-20.00	1.17	-1.968
	Iowa	Union	-20.00	1.17	-1.956
	Iowa	Washington	-20.00	1.17	-1.970
	Kansas	Jackson	-20.00	1.20	-1.871
	Kansas	Republic	-20.00	1.20	-1.926
	Kansas	Washington	-20.00	1.20	-1.969
	Kansas	Wilson	-20.00	1.20	-1.686
	Kentucky	Hart	-20.00	1.98	-2.491
	Louisiana	Morehouse	-20.00	1.87	-1.937
	Louisiana	Plaquemines	-20.00	1.87	-1.672
	Louisiana	Saint Bernard	20.00	2.80	2.115
	Maine	Penobscot	-20.00	11.11	-1.710
	Maryland	Frederick	27.27	16.47	1.677
	Michigan	Ionia	-20.00	1.43	-1.809
	Michigan	Oakland	-20.00	1.43	-1.933
	Minnesota	Le Sueur	75.00	2.69	1.755
	Minnesota	Marshall	-20.00	1.54	-1.901
	Minnesota	Sherburne	-20.00	1.54	-1.838
	Mississippi	Jackson	-25.00	2.01	-1.753
	Missouri	Camden	-20.00	0.90	-1.899
	Missouri	Howard	-20.00	0.90	-1.963
	Missouri	Jefferson	-20.00	0.90	-1.912
	Missouri	Macon	-20.00	0.90	-1.987
	Missouri	Moniteau	-20.00	0.90	-1.863
	Missouri	Reynolds	-20.00	0.90	-1.921

Characteristic	State Name	County Name	Percent Difference	% of State	t-value
	Missouri	Warren	-20.00	0.90	-1.845
	Montana	Missoula	-20.00	2.31	-1.948
	Nebraska	Butler	-25.00	1.18	-1.724
	Nebraska	Cheyenne	-20.00	1.57	-1.857
	Nebraska	Nuckolls	-25.00	1.18	-1.733
	Nebraska	Otoe	-25.00	1.18	-1.736
	Nebraska	Seward	-25.00	1.18	-1.709
	Nebraska	Sioux	-20.00	1.57	-1.945
	North Carolina	Alexander	-20.00	1.25	-1.647
	North Carolina	Davidson	-20.00	1.25	-1.729
	North Carolina	Iredell	-22.22	2.19	-2.127
	North Carolina	Pender	-20.00	1.25	-1.697
	North Dakota	Oliver	-25.00	2.07	-1.733
	North Dakota	Pembina	-20.00	2.76	-1.699
	North Dakota	Williams	-20.00	2.76	-1.731
	Ohio	Clermont	-20.00	1.31	-1.948
	Ohio	Gallia	-20.00	1.31	-1.704
	Ohio	Geauga	-20.00	1.31	-1.917
	Ohio	Madison	-20.00	1.31	-1.940
	Oklahoma	Payne	23.08	2.90	1.825
	South Dakota	Day	60.00	4.76	1.693
	South Dakota	Turner	-20.00	2.38	-1.931
	Tennessee	Cannon	-20.00	1.07	-1.688
	Tennessee	Davidson	-28.57	1.33	-2.209
	Tennessee	Dickson	-20.00	1.07	-1.915
	Tennessee	Mcnairy	-20.00	1.07	-1.957
	Virginia	Frederick	-20.00	1.72	-1.964
	Wisconsin	Clark	-20.00	1.59	-1.931
	Wisconsin	Eau Claire	-20.00	1.59	-1.923
	Wisconsin	Shawano	-20.00	1.59	-1.986
	Wyoming	Fremont	21.05	17.56	2.067

APPENDIX E. SUMMARY OF DIFFERENCES BETWEEN STATE AND COUNTY TOTALS FOR GEORGIA, TEXAS, AND VIRGINIA

Georgia data - 1997 census			
Characteristic	Level	Integer Total - Noninteger Total	Rounded Total - Noninteger Total
Value of Land and Buildings	state	-7,033,641	72
	min cnty	-2,892,379	-5
	max cnty	2,881,849	5
Farm Production Expenditures	state	-207,412	60
	min cnty	-135,120	-8
	max cnty	204,418	6
Total Value of Products Sold	state	-1,410,654	180
	min cnty	-536,106	-9
	max cnty	463,555	11
Land in Farms	state	-26,801	303
	min cnty	-3,529	-13
	max cnty	2,403	18
Acres of Harvested Cropland	state	-4,889	32
	min cnty	-1,393	-13
	max cnty	1,723	9
Acres of Corn for Grain Harvested	state	-885	35
	min cnty	-348	-5
	max cnty	620	4
Bushels of Corn for Grain Harvested	state	-130,369	27
	min cnty	-34,972	-4
	max cnty	52,426	6
Acres of Wheat for Grain Harvested	state	-1,777	5
	min cnty	-354	-4
	max cnty	404	3
Bushels of Wheat for Grain Harvested	state	-56,259	6
	min cnty	-18,331	-3
	max cnty	19,843	3
Acres of Soybeans Harvested	state	-296	15
	min cnty	-340	-2
	max cnty	706	3
Bushels of Soybeans Harvested	state	1,065	27
	min cnty	-12,088	-2
	max cnty	15,117	5
Acres of Cotton Harvested	state	-1,202	9
	min cnty	-723	-3
	max cnty	697	3

Georgia data - 1997 census			
Characteristic	Level	Integer Total - Noninteger Total	Rounded Total - Noninteger Total
Bales of Cotton Harvested	state	-943	18
	min cnty	-995	-4
	max cnty	900	5
Total Cattle and Calves	state	-3,759	182
	min cnty	-630	-15
	max cnty	659	9
Dollars Received for Dairy Products	state	93,501	-3
	min cnty	-120,013	-1
	max cnty	140,939	1
Total Hogs and Pigs	state	2,858	-1
	min cnty	-491	-4
	max cnty	700	4
Layers 20 Weeks Old and Older	state	-79,527	4
	min cnty	-31,671	-4
	max cnty	19,248	4
Broilers and Other Meat-Type Chickens Sold	state	-445,904	1
	min cnty	-935,341	-1
	max cnty	450,302	1
Farms Operated by Blacks and Other Races	state	7	-151
	min cnty	-3	-8
	max cnty	3	1
Farms Operated by Females	state	-33	-443
	min cnty	-5	-10
	max cnty	4	1
Farms Operated by Persons of Hispanic Origin	state	6	-25
	min cnty	-1	-1
	max cnty	1	0
Number of Farms	state	0	-4,014
	min cnty	0	-92
	max cnty	0	6

Texas data - 1997 census			
Characteristic	Level	Integer Total - Noninteger Total	Rounded Total - Noninteger Total
Value of Land and Buildings	state	3,016,549	316
	min cnty	-6,887,964	-6
	max cnty	5,526,046	15
Farm Production Expenditures	state	1,048,512	-1
	min cnty	-222,899	-10
	max cnty	214,559	12

Texas data - 1997 census			
Characteristic	Level	Integer Total - Noninteger Total	Rounded Total - Noninteger Total
Total Value of Products Sold	state	-92,594	558
	min cnty	-501,672	-21
	max cnty	990,152	37
Land in Farms	state	-45,143	285
	min cnty	-10,832	-45
	max cnty	20,212	54
Acres of Harvested Cropland	state	-3,658	181
	min cnty	-4,155	-64
	max cnty	2,402	60
Acres of Corn for Grain Harvested	state	-451	2
	min cnty	-766	-6
	max cnty	979	6
Bushels of Corn for Grain Harvested	state	-44,145	-12
	min cnty	-145,745	-3
	max cnty	99,626	3
Acres of Wheat for Grain Harvested	state	-748	14
	min cnty	-2,314	-6
	max cnty	2,359	8
Bushels of Wheat for Grain Harvested	state	-1,629	44
	min cnty	-36,454	-6
	max cnty	67,443	12
Acres of Soybeans Harvested	state	2,107	7
	min cnty	-384	-2
	max cnty	1,618	3
Bushels of Soybeans Harvested	state	9,886	5
	min cnty	-11,092	-4
	max cnty	24,216	2
Acres of Cotton Harvested	state	2,827	25
	min cnty	-1,488	-4
	max cnty	1,970	7
Bales of Cotton Harvested	state	2,705	-5
	min cnty	-1,103	-4
	max cnty	2,047	6
Total Cattle and Calves	state	1,986	701
	min cnty	-1,834	-41
	max cnty	1,621	45
Dollars Received for Dairy Products	state	325,215	9
	min cnty	-274,514	-2
	max cnty	354,517	2

Texas data - 1997 census			
Characteristic	Level	Integer Total - Noninteger Total	Rounded Total - Noninteger Total
Total Hogs and Pigs	state	-1,281	-147
	min cnty	-357	-7
	max cnty	744	6
Layers 20 Weeks Old and Older	state	-35,673	28
	min cnty	-31,299	-9
	max cnty	31,572	8
Broilers and Other Meat-Type Chickens Sold	state	-324,496	-1
	min cnty	-170,131	-1
	max cnty	118,908	1
Farms Operated by Blacks and Other Races	state	3	-1,223
	min cnty	-5	-47
	max cnty	9	1
Farms Operated by Females	state	-51	-3,053
	min cnty	-14	-53
	max cnty	11	1
Farms Operated by Persons of Hispanic Origin	state	3	-1,164
	min cnty	-8	-95
	max cnty	10	1
Number of Farms	state	0	-26,925
	min cnty	0	-343
	max cnty	0	7

Virginia data - 1997 census			
Characteristic	Level	Integer Total - Noninteger Total	Rounded Total - Noninteger Total
Value of Land and Buildings	state	720,977	128
	min cnty	-2,328,309	-5
	max cnty	3,887,524	15
Farm Production Expenditures	state	-76,832	34
	min cnty	-116,301	-6
	max cnty	196,232	6
Total Value of Products Sold	state	-575,119	170
	min cnty	-346,315	-33
	max cnty	414,753	14
Land in Farms	state	6,261	119
	min cnty	-2,538	-21
	max cnty	1,933	21
Acres of Harvested Cropland	state	604	-206
	min cnty	-582	-53
	max cnty	893	15

Virginia data - 1997 census			
Characteristic	Level	Integer Total - Noninteger Total	Rounded Total - Noninteger Total
Acres of Corn for Grain Harvested	state	717	-27
	min cnty	-171	-9
	max cnty	283	5
Bushels of Corn for Grain Harvested	state	81,713	19
	min cnty	-19,575	-4
	max cnty	26,827	6
Acres of Wheat for Grain Harvested	state	302	7
	min cnty	-175	-6
	max cnty	259	7
Bushels of Wheat for Grain Harvested	state	2,761	11
	min cnty	-13,882	-7
	max cnty	19,456	4
Acres of Soybeans Harvested	state	1,537	27
	min cnty	-338	-2
	max cnty	617	4
Bushels of Soybeans Harvested	state	22,121	3
	min cnty	-9,354	-3
	max cnty	13,408	4
Acres of Cotton Harvested	state	95	6
	min cnty	-86	0
	max cnty	104	2
Bales of Cotton Harvested	state	103	3
	min cnty	-112	-1
	max cnty	104	2
Total Cattle and Calves	state	1,085	-22
	min cnty	-876	-26
	max cnty	786	11
Dollars Received for Dairy Products	state	-292,542	2
	min cnty	-122,914	-3
	max cnty	102,213	2
Total Hogs and Pigs	state	972	-56
	min cnty	-125	-4
	max cnty	420	2
Layers 20 Weeks Old and Older	state	-24,316	0
	min cnty	-25,298	-3
	max cnty	14,642	3
Broilers and Other Meat-Type Chickens Sold	state	-493,666	1
	min cnty	-423,092	-1
	max cnty	65,829	1

Virginia data - 1997 census			
Characteristic	Level	Integer Total - Noninteger Total	Rounded Total - Noninteger Total
Farms Operated by Blacks and Other Races	state	11	-143
	min cnty	-5	-19
	max cnty	3	0
Farms Operated by Females	state	-7	-473
	min cnty	-5	-25
	max cnty	9	0
Farms Operated by Persons of Hispanic Origin	state	7	-12
	min cnty	-1	-1
	max cnty	2	0
Number of Farms	state	0	-4,110
	min cnty	0	-191
	max cnty	0	0