United States Department of Agriculture

National Agricultural Statistics Service

Research and Development Division
Washington DC 20250

RDD Research Report
RDD-08-01

January 2008

Assessing the Effect of Calibration on Nonresponse Bias in the 2005 ARMS Phase III Sample Using 2002 Census of Agriculture Data

Morgan S. Earp
Jaki S. McCarthy
Nick D. Schauer
Phil S. Kott

EXECUTIVE SUMMARY

Phase III of the Agricultural Resource Management Survey (ARMS) is one of the most complex and detailed sample survey data collections conducted by the United States Department of Agriculture's (USDA) National Agricultural Statistics Service (NASS). For this survey, NASS collects highly detailed, calendar year economic data from agricultural producers nationwide. The USDA uses ARMS data to evaluate the financial performance of farms and ranches, which influence agricultural policy decisions. The Department also uses Phase III data for objective evaluation of critical issues related to agriculture and the rural economy.

In September 2006 the Executive Office of the President released the *Office of Management and Budget Standards and Guidelines for Statistical Surveys* based on the recommendations of the Federal Committee on Statistical Methodology's (FCSM) Subcommittee on Standards for Statistical Surveys. The Office of Management and Budget's (OMB) new standards and guidelines for statistical surveys addressed a number of federal statistical agency issues, specifically response rates below 80 percent and analysis of unit nonresponse. The *2005 Agricultural Resource Management Survey (ARMS) Phase III Survey Administration Analysis* (Hopper, 2006) reported a response rate of 70.5 percent, and thus NASS was required by OMB to conduct an analysis of nonresponse bias. This report focuses on fulfilling OMB's directive, by assessing the effect of nonresponse bias as well as the utility of the NASS calibration process in adjusting for such biases.

Records sampled for the 2005 ARMS Phase III were matched with those from the 2002 Census of Agriculture, and means of census data were calculated for matching records which had also provided 2002 expenditure data for the Census. Nonresponse bias in ARMS data was assessed, using census data as a proxy, in terms of the degree to which the mean based on all sample cases versus respondent cases differed. Three means were computed and compared across 20 regions in order to assess relative bias: 1) the mean of all matching cases using base sampling weights, 2) the mean for only matching ARMS respondents using the same base sampling weights, and 3) the mean for matching ARMS respondents using the sampling weights as adjusted through calibration.

Using 17 "study variables," relative bias of the mean was assessed using a variation of the formula provided by OMB in Guideline 3.2.9. Although significant biases were exhibited in 9 of 17 variables using the 2005 ARMS Phase III base sampling weights, the 2005 ARMS Phase III calibration weights were able to reduce the bias so that it was no longer significantly different from zero ($p < .05$) in almost 90 percent (8/9) of the study variables. For this analysis, the calibration process varied slightly from that of the 2005 ARMS Phase III, in that egg and milk production were not included, since there were no comparable variables in the 2002 Census. The inability to replicate the 2005 ARMS III calibration process may in part account for the one variable, fertilizer expenses, still demonstrating a significant level of bias after the use of calibrated weights. This study suggests that the process of calibration is an effective tool in reducing nonresponse bias levels, so they are no longer significantly different from zero.

RECOMMENDATIONS

1. Nonresponse bias of all study variables, especially livestock purchases and fertilizer expenses should be reevaluated when the 2007 Census data are available, since this Census will contain equivalent calibration target variables for egg and milk production, and expenditure data for all Census respondents allowing for an assessment that will be consistent with the calibration targets used for the 2005 ARMS Phase III.

2. Methods should be developed to assess biases not measured in this analysis, especially those that may exist in only a single region.[1]

[1] ARMS III estimate regions include the 15 leading cash receipts states (Arkansas, California, Florida, Georgia, Illinois, Iowa, Kansas, Minnesota, Nebraska, North Carolina, Texas, Washington, and Wisconsin) and the five main production regions (Atlantic, South, Midwest, Plains, and West) comprised of the remaining 35 states.

Assessing the Effect of Calibration on Nonresponse Bias in the 2005 ARMS Phase III Sample Using 2002 Census of Agriculture Data

Morgan Earp, Jaki McCarthy, Nick Schauer, & Phil Kott[2]

Abstract

The United States Department of Agriculture's National Agricultural Statistics Service conducts the annual Agricultural Resource Management Survey (ARMS) in three phases. The third phase of the ARMS collects detailed economic data which is highly sensitive, and thus this phase suffers from lower response rates. With the release of the 2006 *Office of Management and Budget Standards and Guidelines for Statistical Surveys*, response rates lower than 80 percent may not only result in nonresponse bias, but they can jeopardize the future of surveys carried out by federal agencies. Based on the assumption that the use of calibrated weights derived from appropriate targets addresses nonresponse bias, the effect of the 2005 Phase III ARMS calibrated weights on relative bias of the mean was tested for all cases sampled in the 2005 ARMS with Census 2002 expenditure data.

The results showed that calibrated weights decreased bias levels so that they were no longer significantly different from zero for almost 90 percent of the "study variables" ($p < .05$).

Key Words: Nonresponse; response rate; bias; calibration weights.

[2] Morgan Earp is a survey & mathematical statistician in the USDA's National Agricultural Statistics Service (NASS) – Research & Development Division (RDD), located in Room 305, 3251 Old Lee Highway, Fairfax, VA 22030. Jaki McCarthy provided assistance with this research while the Chief Research Methodologist with the USDA's NASS – RDD. Nick Schauer provided assistance with this research while a mathematical statistician in the USDA's NASS - Statistics Division. Phil Kott provided assistance with this research while the Chief Research Statistician with the USDA's NASS – RDD.

1. INTRODUCTION

On September 22, 2006, the Executive Office of the President released the *Office of Management and Budget Standards and Guidelines for Statistical Surveys* based on the recommendations of the Federal Committee on Statistical Methodology's (FCSM) Subcommittee on Standards for Statistical Surveys. The Office of Management and Budget's (OMB) new standards and guidelines for statistical surveys pertain to aspects of surveys conducted by federal statistical agencies.

Federal statistical agencies, such as the United States Department of Agriculture's (USDA) National Agricultural Statistics Service (NASS), are directly affected by OMB's new standards and guidelines for statistical surveys. One of the standards (3.2) issued by OMB addresses response rates and analysis of nonresponse bias. According to Standard 3.2,

> Agencies must appropriately measure, adjust for, report, and analyze unit and item nonresponse to assess their effects on data quality and to inform users. Response rates must be computed using standard formulas to measure the proportion of the eligible sample that is represented by the responding units in each study, as an indicator of potential nonresponse bias. (Office of Management and Budget, 2006, p. 14)

In 2005, the Agricultural Resource Management Survey (ARMS) Phase III response rate was 70.5 percent (n = 34,937), which fell below the OMB response rate cut off of 80 percent listed in Guideline 3.2.9; therefore, NASS was required by OMB to research the effect of nonresponse bias. Since the Phase II response rate of 80.5 percent exceeded OMB's 80 percent threshold, nonresponse bias assessment was only required for Phase III, the "problem" stage. Guideline 3.2.9 states

> Given a survey with an overall unit response rate of less than 80 percent, conduct an analysis of nonresponse bias using unit response rates as defined above, with an assessment of whether the data are missing completely at random. As noted above, the degree of nonresponse bias is a function of not only the response rate but also how much the respondents and nonrespondents differ on the survey variables of interest. For a sample mean, an estimate of the bias of the sample respondent mean is given by:

$$B(\bar{y}_r) = \bar{y}_r - \bar{y}_t = \left(\frac{n_{nr}}{n}\right)(\bar{y}_r - \bar{y}_{nr})$$

Where:

$\bar{y}_t$ = The mean based on all sample cases;

$\bar{y}_r$ = The mean based only on respondent cases;

$\bar{y}_{nr}$ = The mean based only on nonrespondent cases;

n = The number of cases in the sample; and

n_{nr} = The number of nonrespondent cases.

For a multistage (or wave) survey, focus the nonresponse bias analysis on each stage, with

particular attention to the "problem" stages. A variety of methods can be used to examine nonresponse bias, for example, make comparisons between respondents and nonrespondents across subgroups using available sample frame variables. In the analysis of unit nonresponse, consider a multivariate modeling of response using respondent and nonrespondent frame variables to determine if nonrespondent bias exists. (Office of Management & Budget, 2006, p. 16)

Currently NASS calculates the unweighted unit response rates (*RRU*) for the ARMS based on the formula provided under Guideline 3.2.2 of the *Office of Management and Budget Standards and Guidelines for Statistical Surveys*:

$$RRU = \frac{C}{C + R + NC + O + e(U)}$$

Where:

C = The number of completed cases or sufficient partials;

R = The number of refused cases;

NC = The number of noncontacted sample units known to be eligible;

O = The number of eligible sample units not responding for reason other than refusal;

U = The number of sample units of unknown eligibility, not completed; and

e = The estimated proportion of sample units of unknown eligibility that are eligible. (p. 14)

NASS sums the number of positive usables, out of business, and non-farms and calculates the percentage this sum represents of the total number of reports to calculate the response rate for ARMS Phase III.

The ARMS is conducted in three phases. Phase I screens for potential samples for Phases II and III. Phase II collects data on cropping practices and agricultural chemical usage and Phase III collects detailed economic information about the agricultural operation, as well as the operator's household. Phase III is the only phase of the ARMS with response rates lower than 80 percent.

Due to lower response rates with ARMS Phase III, the potential for nonresponse bias is greater there. NASS weights the ARMS Phase III respondent sample in such a way that estimated variable totals for a large set of items match "targets" determined from other sources. This is done through a weighting process called "calibration." Calibration is the process of adjusting survey weights so that certain targets are met. NASS uses official estimates of farm numbers, corn, soybean, wheat, cotton, fruit and vegetable acres as well as cattle, milk production, hogs, broilers, eggs and turkeys as calibration targets. For example, after calibration the sum of weights multiplied by the survey data will equal the NASS estimate for corn acres. In addition to reducing confusion in the user community that might result from NASS releasing alternative estimates for the same totals, calibration

weighting produces 2005 ARMS Phase III estimates with generally lower variances and, hopefully reduced nonresponse biases. This report describes an ongoing research effort aimed at measuring the potential for nonresponse bias in the ARMS Phase III and the success or failure of calibration in removing it.

Nonresponse bias is very difficult to measure directly. Fortunately, an indirect measure of nonresponse bias is available for the 2005 ARMS Phase III, hereafter called simply the "ARMS."

The Census of Agriculture is a mandatory collection of data from all known agricultural operations. NASS has data from the Census on items of interest for many of the ARMS nonrespondents; however, the Census itself is incomplete. An estimated 17.90 percent of all farms are missing from the 2002 Census Mailing List, and 12.26 percent of farms on the List failed to respond to the Census. Moreover, not all ARMS sampled farms could be matched to 2002 Census records. Nevertheless, by comparing the 2002 Census values of ARMS respondents to the full sample of ARMS respondents, we can measure the difference between the average ARMS respondent and the average of the full sample without any nonresponse adjustment. Additionally, we hope to measure the reduction of that difference from using a calibration-weighting process similar to the one used for the 2005 ARMS.

While the 2002 Census data do not perfectly match the 2005 ARMS data, they are correlated (see Appendix A), so the present evaluation will compare respondents on the 2005 ARMS survey to nonrespondents using their 2002 Census data for each.

2. METHODS

Our analytical data set consists of census values for farms sampled for the ARMS that also provided 2002 expenditure data on the Census. In the 2002 Census, only a sample of farms received the long version that asks the expenditure data. [3] 2002 Census data were available for 81.4% of all 2005 ARMS III sampled operations. However, only the 48% of those which had completed the Census long forms with expenditure data were included in this analysis.[4]

The base sample weights (each farm's ARMS sample weight before calibration multiplied by its Census sample weight) for the subset of farms responding to the ARMS were calibrated so that the final weighted totals computed from them equaled the raw weighted total computed from the entire matched set for the following variables: cattle, corn, cotton, pigs, soybeans, wheat, fruit, vegetables, broilers, and turkeys.[5] Each of these target variables plus egg and milk production was used operationally to calibrate the ARMS data.

As in the operational program, the ARMS respondent subset was calibrated independently in 20 regions. These included the 15 leading cash receipts states (Arkansas, California, Florida,

[3] Only "sample farms" were used for this analysis.

[4] The match rate for 2005 ARMS Phase III was significantly higher for nonrespondents (86.5%) than for respondents (79.2%) ($z = 16.04$, $p < .05$).

[5] The sample weight was used to determine which operations were to report expenditure data from (USDA, 2002).

Georgia, Illinois, Indiana, Iowa, Kansas, Minnesota, Missouri, Nebraska, North Carolina, Texas, Washington, and Wisconsin). The remaining 33 states (Alaska and Hawaii are not sampled for ARMS) were grouped using the five production regions: 1) Atlantic, 2) South, 3) Midwest, 4) Plains, and 5) West.

Our analysis focuses on 17 specific variables collected on both the ARMS and the Census: total acres, total sales, acres rented, cropland acres, total production expenses, crop expenses, seed expenses, fertilizer expenses, chemical expenses, livestock purchases, feed purchases, hired labor expenses, machinery and equipment value, government payments, operator's age, operator's race, and farm type.

Letting $\bar{y}_r$ denote the 2002 Census preliminary-sample or calibrated-sample mean among the ARMS respondent subset for a study variable, and $\bar{y}_t$ denote the corresponding preliminary-sample mean among the entire matched sample, it is a simple matter to compute the relative bias of the former with respect to the latter, $\text{relBias} = \dfrac{\bar{y}_r - \bar{y}_t}{\bar{y}_t}$. The statistical significance of this value is much harder to assess since the samples on which $\bar{y}_r$ and $\bar{y}_t$ are based are complex and overlapping.

Fortunately, we can easily test the persistence or absence of a systematic bias across the 20 regions. To this end, we compute the following measure of bias of an ARMS-respondent mean (before or after calibration) with respect to the Census mean in every region:

$$
\begin{aligned}
M &= \log(\bar{y}_r) - \log(\bar{y}_t) \\
&= \log\left(\bar{y}_r \Big/ \bar{y}_t \right) \\
&= \log\left(1 + \frac{\bar{y}_r - \bar{y}_t}{\bar{y}_t} \right) \\
&\approx \frac{\bar{y}_r - \bar{y}_t}{\bar{y}_t}
\end{aligned}
$$

This measure is conveniently symmetric, $\log(\bar{y}_t) - \log(\bar{y}_r) = -[\log(\bar{y}_r) - \log(\bar{y}_t)]$, while retaining the scale-invariance property of the relative bias (i.e., multiplying the reported item value on each farm by a fixed factor does not affect the overall relative bias).

The bias measure *M* for a study variable in a region can be treated as an independent random variable. The null hypothesis of no bias (again, either before or after calibration) can be tested against an alternative hypothesis of a persistent bias (*p*%) across all the regions. The conventional *t* test based on the 20 observations (one per region) is asymptotically normal under both the null and alternative hypotheses. We follow the standard practice of approximating the distribution of this test statistic with a Student's *t* having 19 degrees of freedom. This may lead to liberal inferences (the inappropriate rejection of the null hypothesis when it is true) because the *M*-values for the study variable may not be normally distributed with a common variance across regions. Nevertheless, by taking logs we create a test statistic that is more nearly normal and homoscedastic than relative biases would be.

A sign and a ranked-sign test of the 20 paired observations for a study variable before and after calibration was

conducted. The sign test is not as powerful as the other two tests (i.e., it more often fails to find that M is significantly different from 0 when, in fact, there is a persistent bias across the regions), but it assumes neither that M is normal nor homoscedastic. The signed-rank test assumes the latter but not the former. We include it in our results for completeness.

3. RESULTS AND DISCUSSION

Our results are summarized in Table 1. Chemical expenses, machinery and equipment value, government payments, acres rented, farm type, fuel and oil expenses, operator's age, and cropland acres (Variables 1-8) do not exhibit significant biases using either calibrated or uncalibrated weights. Although chemical expenses (Variable 1) did not exhibit significant bias, significantly less bias was exhibited using the calibrated weights versus the uncalibrated weights.

In almost 90 percent (8/9) of the study variables (9-17) exhibiting persistent biases using the base sample weights, calibration weighting is able to reduce the bias so that it was no longer significantly different from zero (9-16) ($p < .05$) according to the t-test, and in 50% of these variables we saw a significant reduction in bias levels (9-12) ($p < .05$) according to the paired t-test. After calibration, only one study variable, fertilizer expenses has a significant bias ($p < .05$) according the t-test (but not according to the sign test). The bias of livestock purchases is indicated to be the largest of the study variables. Using only the sampling weights, it was highly significant in terms of each of the test statistics. After calibration, while still large in magnitude, the indicated bias was reduced to the point that it was statistically insignificant according to all the test statistics. For this variable, calibration does reduce the bias significantly if not completely.

4. CONCLUSION

ARMS data are used by farm organizations, commodity groups, agribusiness, Congress, State Departments of Agriculture, and the USDA. The USDA uses ARMS data to evaluate the financial performance of farms and ranches, which influence agricultural policy decisions. The Department also uses Phase III data for objective evaluation of critical issues related to agriculture and the rural economy; therefore, it is essential that measures be taken to minimize the effect of nonresponse bias in ARMS, specifically Phase III.

In the research on adjustment for nonresponse bias in the 2005 ARMS Phase III, the 2002 Census mean estimates of feed purchases, total production expenses, total sales, seed expenses, livestock purchases, cropland expenses, total acres operated, hired labor expenses, and fertilizer expenses demonstrated significant bias using just the base sample weights. Although the magnitude of the relative bias of the mean estimate remained high for livestock purchases using the calibrated weights, calibration reduced the magnitude of this bias so that it was no longer significant (see Table 1).

Table 1: Mean Comparisons and Indicated Biases for Matching Records Using Base Sampling Weights versus Calibrated Weights

	National Estimates		Regional Estimates (n = 20)				Testing the Effect of Calibration on Regional Bias Means							
Variable	Mean	Bias	Mean[3]	Bias Mean[3]	Bias Mininmum[3]	Bias Maximum[3]	t[3]	p value[3]	Sign[3]	p value[3]	Signed Ranks[3]	p value[3]	t_{paired}[3]	p value[3]
1. Chemical Expenses (Dollars)														
All Matching Records [1]	4,465.99		5,863.14											
Matching Respondents [2]	4,151.03	-7.05%	5,309.03	-6.66%	-43.91%	27.01%	-1.81	0.09	-4.00	0.12	-47.00	0.08		
Matching Respondents Calibrated [2]	4,520.40	1.22%	5,864.94	1.09%	-18.58%	11.28%	0.61	0.55	3.00	0.26	25.00	0.37	-2.42	0.03
2. Machinery & Equipment Value (Dollars)														
All Matching Records [1]	76,243.46		83,052.61											
Matching Respondents [2]	75,262.16	-1.29%	82,612.76	-0.78%	-21.05%	38.37%	-0.29	0.77	0.00	1.00	-24.00	0.39		
Matching Respondents Calibrated [2]	78,882.89	3.46%	84,958.38	2.27%	-8.79%	17.92%	1.68	0.11	2.00	0.50	40.00	0.14	-1.72	0.10
3. Government Payments (Dollars)														
All Matching Records [1]	3,838.39		4,088.32											
Matching Respondents [2]	3,894.15	1.45%	4,109.51	-0.13%	-26.03%	48.60%	-0.04	0.97	-2.00	0.50	-17.00	0.55		
Matching Respondents Calibrated [2]	3,950.51	2.92%	4,218.76	2.95%	-29.67%	26.88%	1.16	0.26	4.00	0.12	42.00	0.12	-0.81	0.43
4. Acres Rented														
All Matching Records [1]	250.88		240.44											
Matching Respondents [2]	247.33	-1.42%	235.34	-2.26%	-18.19%	16.58%	-1.22	0.24	-5.00	0.04	-39	0.15		
Matching Respondents Calibrated [2]	279.55	11.43%	261.23	2.58%	-38.84%	34.13%	0.77	0.45	2.00	0.50	27.00	0.33	-1.58	0.13
5. Farm Type														
All Matching Records [1]	8.25		7.92											
Matching Respondents [2]	8.44	2.30%	7.90	-0.85%	-25.18%	5.01%	-0.60	0.56	1.00	0.82	9.00	0.76		
Matching Respondents Calibrated [2]	8.21	-0.48%	7.87	-1.66%	-47.98%	5.86%	-0.66	0.52	1.00	0.82	20.00	0.47	0.66	0.52
6. Fuel & Oil Expenses (Dollars)														
All Matching Records [1]	3,715.69		4,301.47											
Matching Respondents [2]	3,488.05	-6.13%	4,018.76	-6.09%	-32.46%	38.58%	-1.96	0.06	-5.00	0.04	-69.00	0.01		
Matching Respondents Calibrated [2]	3,722.81	0.19%	4,260.42	-0.79%	-12.48%	8.84%	-0.65	0.52	0.00	1.00	-7.00	0.81	-1.82	0.08
7. Operator's Age (Years)														
All Matching Records [1]	55.41		55.18											
Matching Respondents [2]	55.73	0.58%	55.52	0.59%	-2.84%	10.28%	0.89	0.38	0.00	1.00	11.00	0.70		
Matching Respondents Calibrated [2]	55.33	-0.14%	55.26	0.12%	-5.37%	5.40%	0.21	0.83	0.00	1.00	1.00	0.99	1.43	0.17
8. Cropland Acres														
All Matching Records [1]	266.11		275.60											
Matching Respondents [2]	262.40	-1.39%	266.98	-3.58%	-30.38%	24.14%	-1.45	0.16	-4.00	0.12	-53	0.05		
Matching Respondents Calibrated [2]	268.45	0.88%	276.16	-0.25%	-15.50%	8.09%	-0.22	0.83	-2.00	0.50	-1.00	0.99	-1.37	0.19

[1] Means computed using the base sampling weights for all matching cases with Census 2002 expenditure data (n = 13,875)

[2] Means computed only for ARMS III respondents with Census 2002 expenditure data (n = 9,258)

[3] Regional estimates are based on the 20 ARMS III estimation regions using only ARMS III respondents

Note: Significant bias and corresponding t scores and p values are identified in **red** font
Significant reduction in bias is identified in **blue** font

Table 1 (Cont.): Mean Comparisons and Indicated Biases for Matching Records Using Base Sampling Weights versus Calibrated Weights

Variable	National Estimates		Regional Estimates (n = 20)				Testing the Effect of Calibration on Regional Bias Means							
	Mean	Bias	Mean[3]	Bias Mean[3]	Bias Mininmum[3]	Bias Maximum[3]	t[3]	p value[3]	Sign[3]	p value[3]	Signed Ranks[3]	p value[3]	t_{paired}[3]	p value[3]
9. Feed Purchases (Dollars)														
All Matching Records [1]	11,366.90		12,288.74											
Matching Respondents [2]	9,614.45	-15.42%	9,704.04	-19.08%	-101.53%	50.78%	-2.44	0.02	-5.00	0.04	-5.00	0.04		
Matching Respondents Calibrated [2]	10,680.33	-6.04%	11,184.01	-6.23%	-49.64%	19.34%	-1.84	0.08	-3.00	0.26	-47.00	0.08	-2.25	0.04
10. Total Production Expenses (Dollars)														
All Matching Records [1]	103,777.07		123,417.43											
Matching Respondents [2]	92,131.68	-11.22%	106,969.35	-11.37%	-60.96%	37.77%	-2.65	0.02	-7.00	0.00	-81.00	0.00		
Matching Respondents Calibrated [2]	100,855.44	-2.82%	117,738.17	-3.20%	-33.61%	7.28%	-1.78	0.09	-5.00	0.04	-57.00	0.03	-2.33	0.03
11. Total Sales (Dollars)														
All Matching Records [1]	124,418.59		150,435.97											
Matching Respondents [2]	109,467.22	-12.02%	142,380.46	-11.57%	-60.73%	46.06%	-2.46	0.02	-4.00	0.12	-70	0.01		
Matching Respondents Calibrated [2]	119,408.37	-4.03%	129,506.04	-3.26%	-30.13%	9.94%	-1.73	0.10	-2.00	0.50	-39.00	0.15	-2.19	0.04
12. Seed Expenses (Dollars)														
All Matching Records [1]	4,663.85		5,577.62											
Matching Respondents [2]	4,226.03	-9.39%	5,246.13	-8.11%	-38.22%	29.40%	-2.71	0.01	-5.00	0.04	-74.00	0.00		
Matching Respondents Calibrated [2]	4,769.68	2.27%	5,655.68	-0.68%	-9.67%	14.32%	-0.42	0.68	-3.00	0.26	-17.00	0.55	-2.33	0.03
13. Livestock Purchases (Dollars)														
All Matching Records [1]	8,204.03		9,376.51											
Matching Respondents [2]	6,298.50	-23.23%	6,560.79	-32.90%	-255.50%	50.53%	-2.45	0.02	-6.00	0.01	-6.00	0.01		
Matching Respondents Calibrated [2]	6,936.76	-15.45%	7,516.65	-20.34%	-234.61%	19.21%	-1.65	0.11	-2.00	0.50	-43.00	0.11	-1.93	0.07
14. Cropland Expenses (Dollars)														
All Matching Records [1]	14,959.54		18,366.38											
Matching Respondents [2]	13,758.92	-8.03%	16,896.49	-7.51%	-26.27%	25.43%	-2.66	0.02	-7.00	0.00	-72.00	0.01		
Matching Respondents Calibrated [2]	14,848.57	-0.74%	18,160.97	-1.46%	-8.23%	6.36%	-1.48	0.16	-3.00	0.26	-38.00	0.17	-2.00	0.06
15. Total Acres Operated														
All Matching Records [1]	572.23		530.13											
Matching Respondents [2]	554.20	-3.15%	503.90	-5.13%	-26.43%	9.64%	-2.80	0.01	-6.00	0.01	-68	0.01		
Matching Respondents Calibrated [2]	557.76	-2.53%	522.66	-2.03%	-21.88%	15.73%	-1.16	0.26	-3.00	0.26	-47.00	0.08	-1.37	0.19
16. Hired Labor Expenses (Dollars)														
All Matching Records [1]	12,320.40		15,595.86											
Matching Respondents [2]	10,557.57	-14.31%	13,471.89	-10.42%	-38.49%	38.87%	-2.67	0.02	-4.00	0.12	-67.00	0.01		
Matching Respondents Calibrated [2]	12,015.11	-2.48%	15,256.50	-2.92%	-36.46%	15.38%	-1.00	0.33	1.00	0.82	-10.00	0.73	-1.87	0.08
17. Fertilizer Expenses (Dollars)														
All Matching Records [1]	5,829.70		6,925.62											
Matching Respondents [2]	5,381.87	-7.68%	6,341.34	-7.84%	-34.00%	20.33%	-2.85	0.01	-5.00	0.04	-68.00	0.01		
Matching Respondents Calibrated [2]	5,558.49	-4.65%	6,640.36	-4.69%	-28.67%	5.15%	-2.58	0.02	-3.00	0.26	-61.00	0.02	-0.90	0.38

[1] Means computed using the base sampling weights for all matching cases with Census 2002 expenditure data (n = 13,875)

[2] Means computed only for ARMS III respondents with Census 2002 expenditure data (n = 9,258)

[3] Regional estimates are based on the 20 ARMS III estimation regions using only ARMS III respondents

Note: Significant bias and corresponding t scores and p values are identified in **red** font
Significant reduction in bias is identified in **blue** font

For this analysis, the calibration process varied slightly from that of the 2005 ARMS Phase III, in that egg and milk production were not included, since these data items were not collected for the 2002 Census. This may help to explain why the magnitude of the relative bias of the mean for livestock estimates in Table 1 remained high even after the data were calibrated. While it was not possible to use these as calibration targets in this analysis, their use in the ARMS III survey may reduce the bias for livestock purchases in published ARMS data.

According to Guideline 3.2.13 of the *Office of Management and Budget Standards and Guidelines for Statistical Surveys*, NASS should

> Base decisions regarding whether or not to adjust or impute data for item nonresponse on how the data will be used, the assessment of nonresponse bias that is likely to be encountered in the review of collections, prior experience with this collection, and the nonresponse analysis discussed in this section. When used, imputation and adjustment procedures should be internally consistent, sampled on theoretical and empirical considerations, appropriate for the analysis, and make use of the most relevant data available. If multivariate analysis is anticipated, care should be taken to use imputations that minimize the attenuation of underlying relationships.

Due to the broadness of the ARMS Phase III data user community and the survey's impact on agricultural policy, it is crucial that the calibration process effectively adjusts for nonresponse bias. Assuming that the adjustment process is even more effective than demonstrated here (particularly for livestock purchases and fertilizer expenses) when all calibration targets (including egg and milk production) are available and used, it appears that NASS is appropriately addressing the issue of nonresponse bias in ARMS Phase III through the calibration process.

Limitations of this analysis include: 1) Inability to replicate the 2005 ARMS Phase III calibration process exactly; 2) Inability to assess farms not covered or responding to the Census of Agriculture or for which expenditure data were not available; 3) Inability to recognize localized biases in the ARMS data (tests were limited to persistent biases across regions); and 4) Assessment of nonresponse bias was conducted using 2002 data as opposed to the 2005 data, since Census data is only available every five years.

Knowing that the analyzed data come from the 2002 Census and not from the 2005 ARMS Phase III Survey does not limit, but strengthens the analysis. It allows us to focus entirely on the impact of the nonresponse *per se*.

5. **RECOMMENDATIONS**

Based on the results of the present study, the following recommendations are offered:

1. Nonresponse bias of all study variables, especially livestock purchases and fertilizer expenses should be reevaluated

when the 2007 Census data are available, since this Census will contain equivalent calibration target variables for egg and milk production, and expenditure data for all Census respondents allowing for an assessment that will be consistent with the calibration targets used for the 2005 ARMS Phase III.

2. Methods should be developed to assess biases not measured in this analysis, especially those that may exist in only a single region.

6. REFERENCES

Hopper, R. (2006). *2005 Agricultural Resource Management Survey (ARMS) Phase III Survey Administration Analysis.* Census and Survey Division Staff Report SAB-06-14, United States Department of Agriculture, National Agricultural Statistics Service.

Kott, P.S. (2005), "Using Calibration Weighting to Adjust for Nonresponse and Coverage Errors" Survey Methodology 32: 133-142.

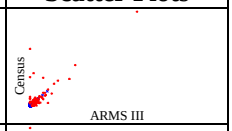
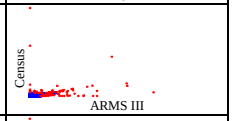
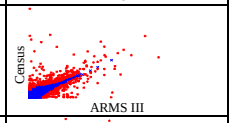
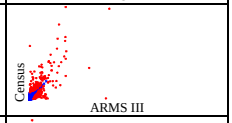
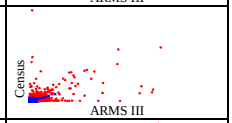
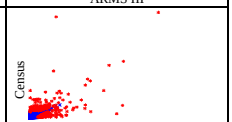
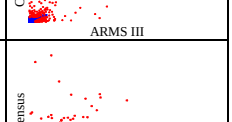
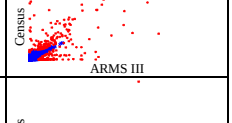
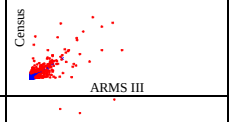
United States Department of Agriculture. (2002). *2002 Census of Agriculture, Vol. 1, Appendix C.* Washington, DC: U.S. Department of Agriculture.

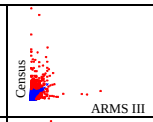
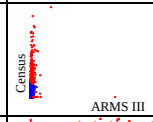
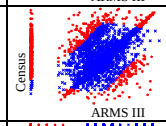
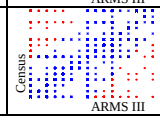
United States Department of Education. (2003). *National Center for Education Statistics Statistical Standards.* Washington, DC: U.S. Department of Education.

United States Executive Office of the President. (2006). *Office of Management and Budget Standards and Guidelines for Statistical Surveys.* Washington, DC: U.S. Executive Office of the President.

APPENDIX A: Census 2002 and ARMS Phase III 2005 Correlations & Scatter Plots

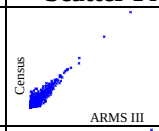
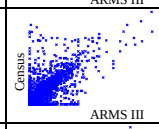
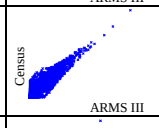
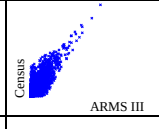
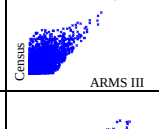
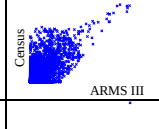
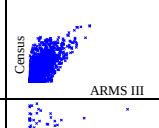
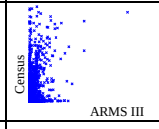
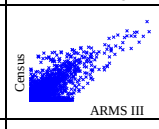
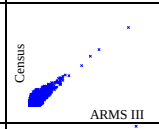
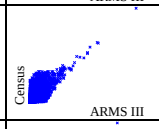
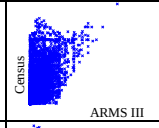
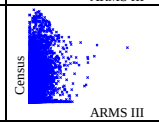
Table A1: Census 2002 and ARMS III 2005 Variable Correlations with Outliers

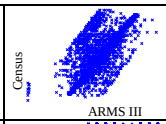
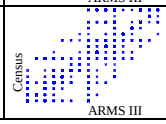
	r	r^2	Scatter Plots
Total Acres Operated	.84805 (*n* = 9,258)	.71919	
Acres Rented	.35984 (*n* = 9,258)	.12948	
Cropland Acres	.83219 (*n* = 9,258)	.69254	
Total Production Expenses	.78981 (*n* = 9,258)	.62380	
Seed Expenses	.57695 (*n* = 9,258)	.33287	
Fertilizer Expenses	.69688 (*n* = 9,258)	.48564	
Chemical Expenses	.60506 (*n* = 9,258)	.36610	
Crop Expenses	.66189 (*n* = 9,258)	.43810	
Livestock Purchases	.43729 (*n* = 9,258)	.19122	
Feed Purchases	.75983 (*n* = 9,258)	.57734	
Hired Labor Expenses	.84178 (*n* = 9,258)	.70859	
Fuel & Oil Expenses	.68970 (*n* = 9,258)	.47569	

Machinery & Equipment	.44797 (*n* = 9,258)	.20068	
Government Payments	.26941 (*n* = 9,258)	.07258	
Operator's Age	.56132 (*n* = 9,258)	.31508	
Farm Type	.80645 (*n* = 9,258)	.65036	

1. All correlations were significant at the .05 level.
2. Correlations were only estimated for ARMS respondents.
3. Outliers were flagged using DFFITS, Cook's D, and studentized residuals and are shown in red.

Table A2: Census 2002 and ARMS III 2005 Variable Correlations without Outliers

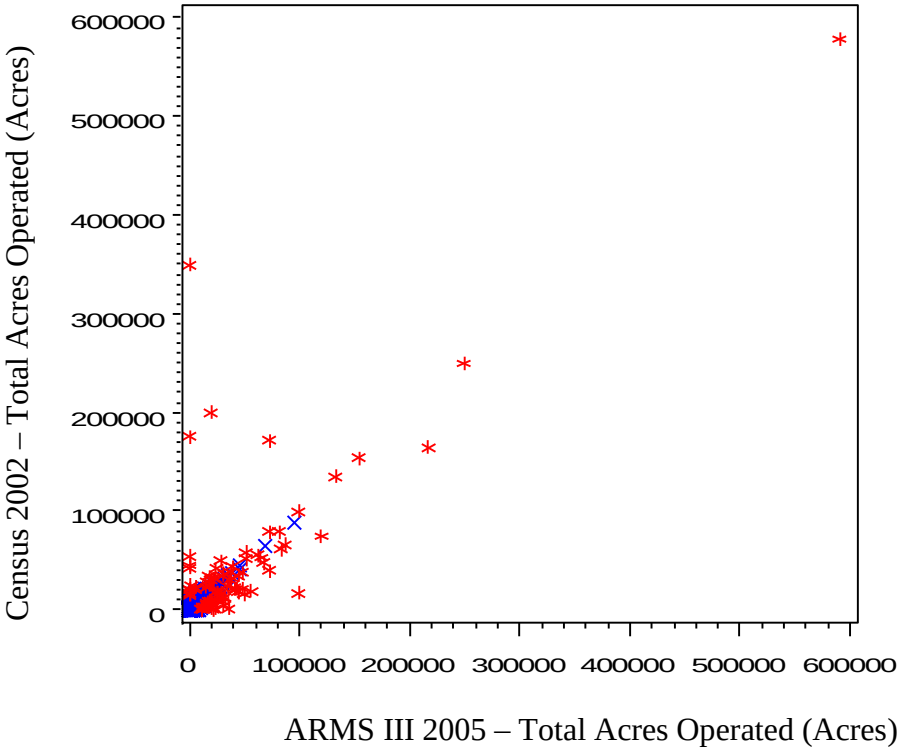
	r	r^2	Scatter Plots
Total Acres Operated	.95246 (*n* = 9,174)	.90718	
Acres Rented	.79452 (*n* = 9,164)	.63126	
Cropland Acres	.94193 (*n* = 8,834)	.88723	
Total Production Expenses	.83579 (*n* = 9,014)	.69854	
Seed Expenses	.68555 (*n* = 9,092)	.46998	
Fertilizer Expenses	.77179 (*n* = 8,946)	.59566	
Chemical Expenses	.75377 (*n* = 9,026)	.56817	
Crop Expenses	.81138 (*n* = 9,036)	.65834	
Livestock Purchases	.41659 (*n* = 9,159)	.17355	
Feed Purchases	.75001 (*n* = 8,977)	.56252	
Hired Labor Expenses	.88404 (*n* = 9,017)	.78152	
Fuel & Oil Expenses	.75420 (*n* = 8,963)	.56882	
Machinery & Equipment	.61434 (*n* = 8,910)	.37741	
Government Payments	.43480 (*n* =9,007)	.18905	

Operator's Age	.84767 (*n* = 8,582)	.71854	
Farm Type	.80645 (*n* = 9,258)	.65036	

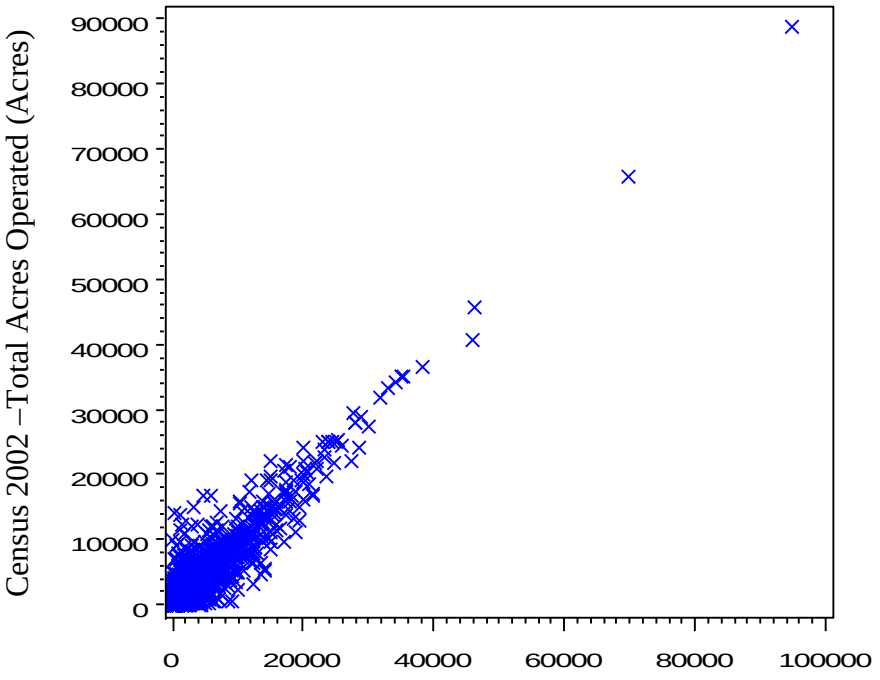
1. All correlations were significant at the .05 level (*n* = 19,483).
2. Correlations were only estimated for ARMS respondents.

Figures A1-A2: Census 2002 versus ARMS III 2005 Total Acres Operated

Scatter Plot of Total Acres Operated with Outliers

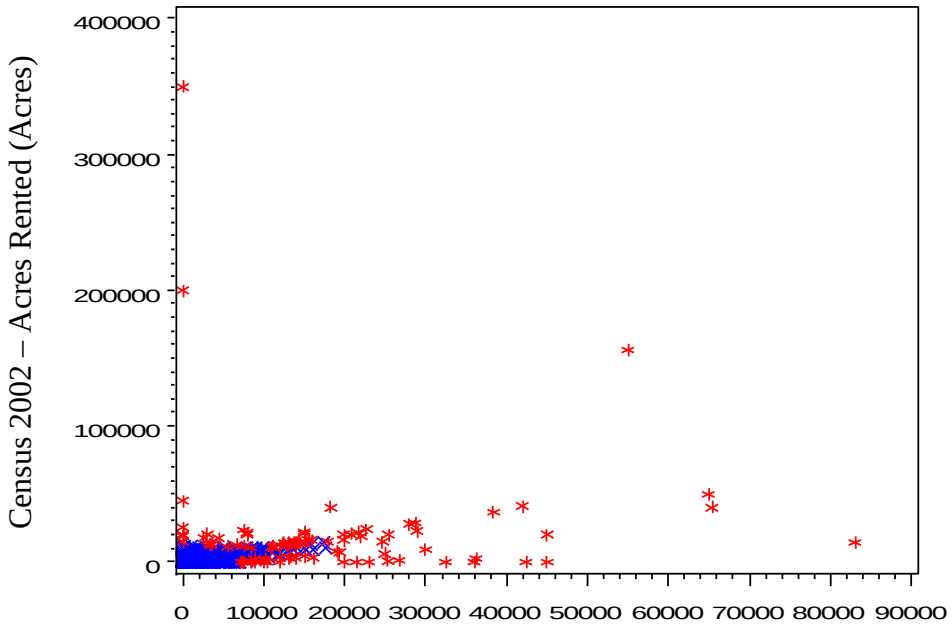


Scatter Plot of Total Acres Operated without Outliers

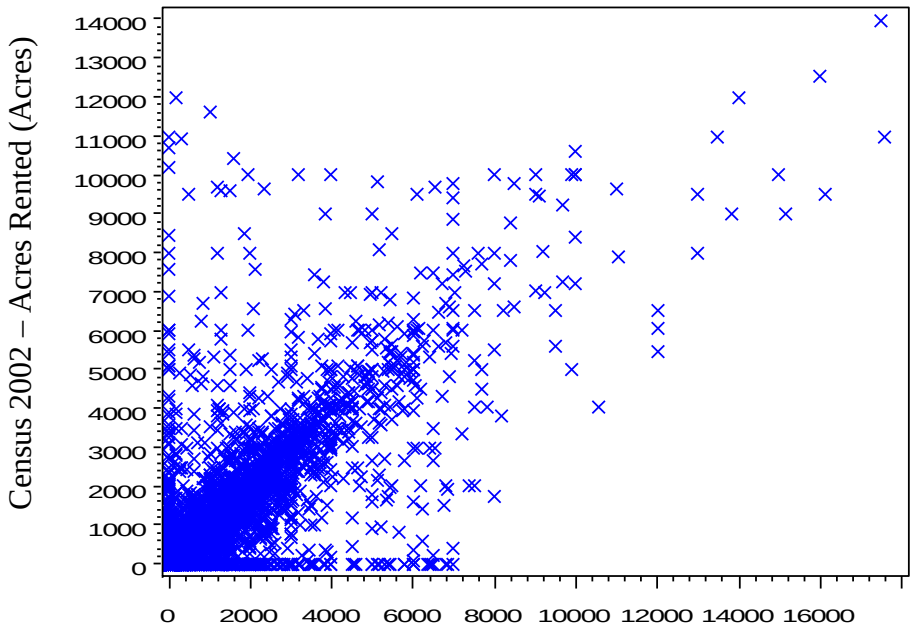


Figures A3-A4: Census 2002 versus ARMS III 2005 Acres Rented

Scatter Plot of Acres Rented with Outliers

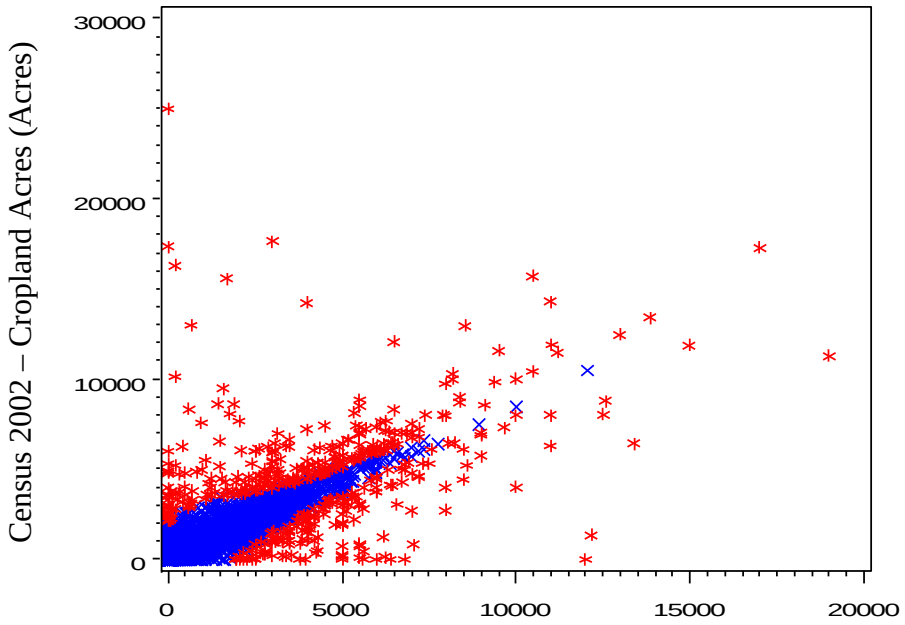


Scatter Plot of Acres Rented without Outliers

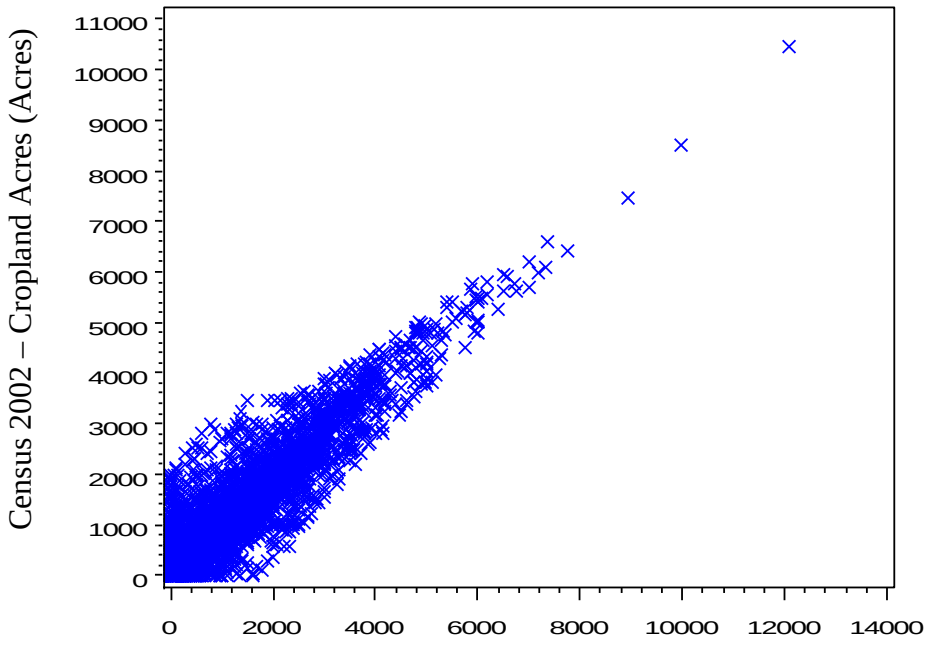


Figures A5-A6: Census 2002 versus ARMS III 2005 Cropland Acres

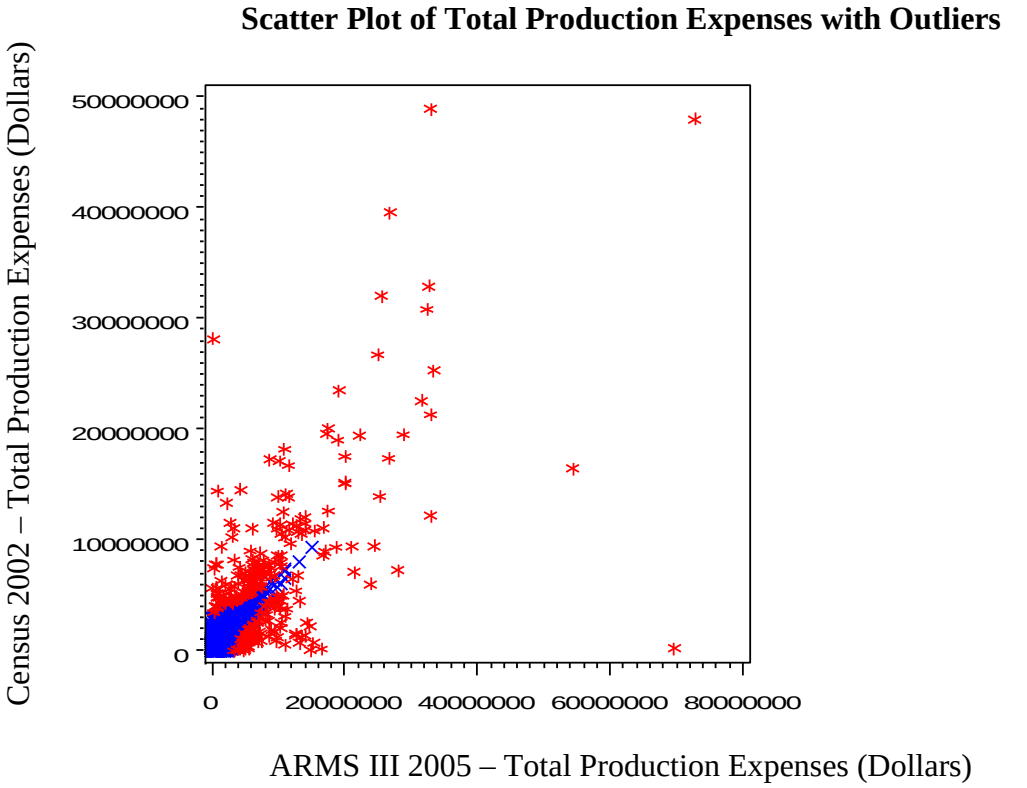
Scatter Plot of Cropland Acres with Outliers

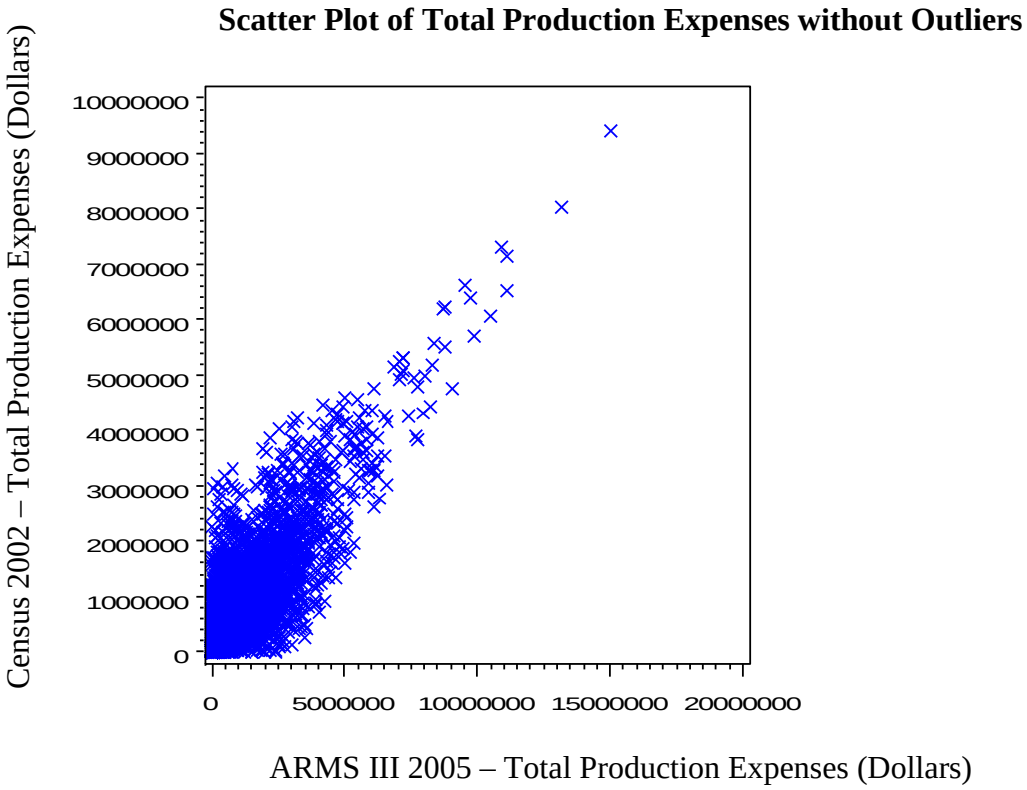


Scatter Plot of Cropland Acres without Outliers



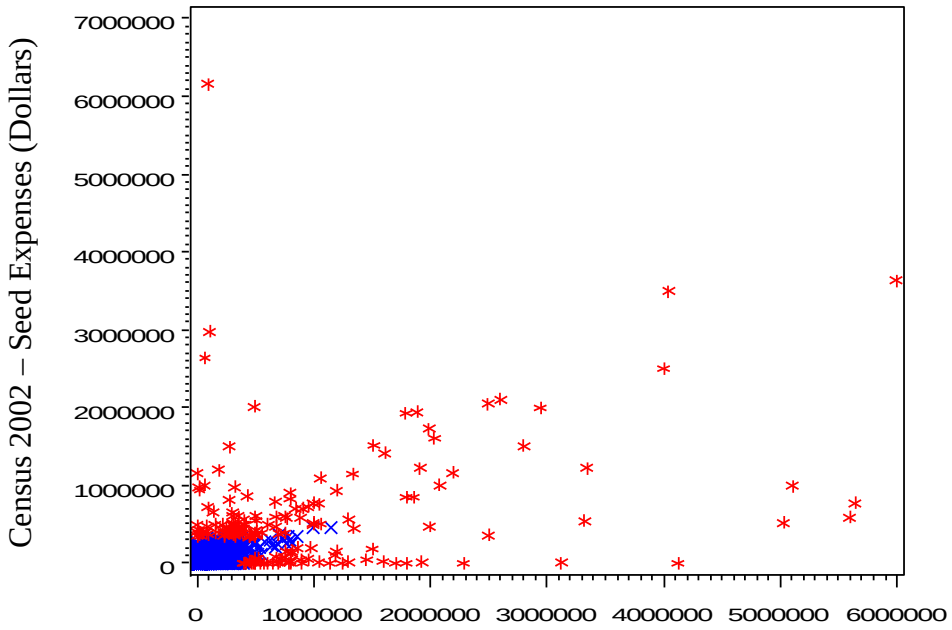
Figures A7-A8: Census 2002 versus ARMS III 2005 Total Production Expenses



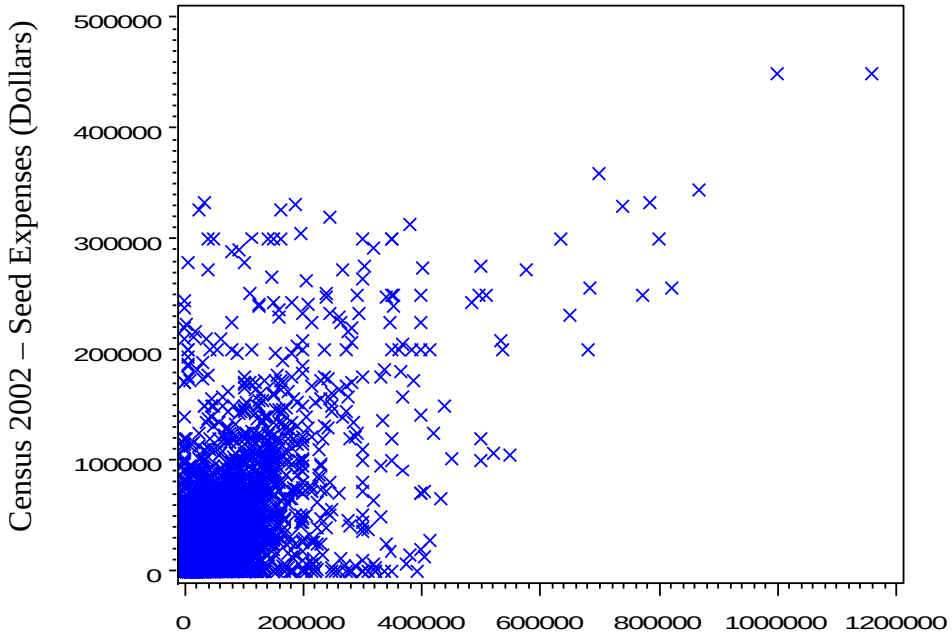


Figures A9-A10: Census 2002 versus ARMS III 2005 Seed Expenses

Scatter Plot of Seed Expenses with Outliers

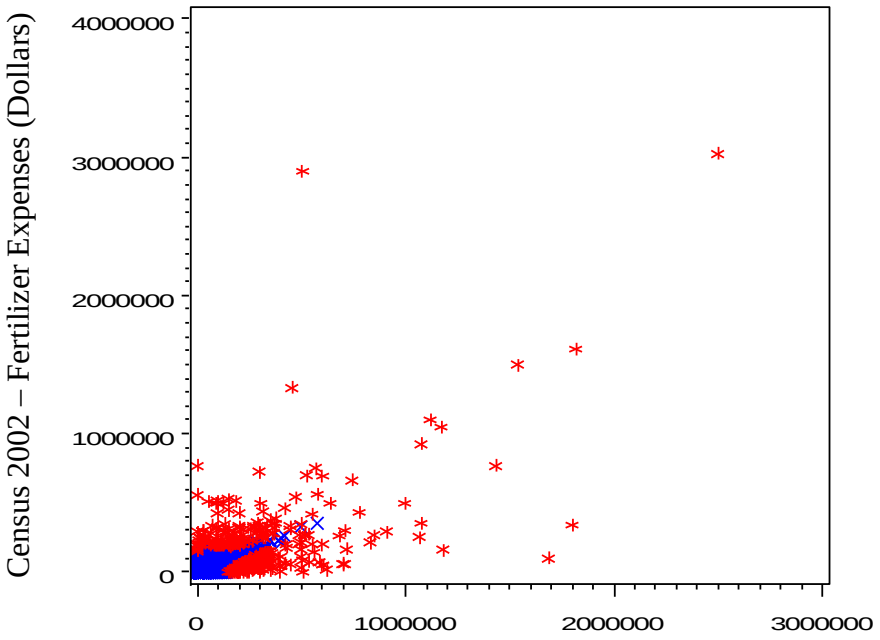


Scatter Plot of Seed Expenses without Outliers

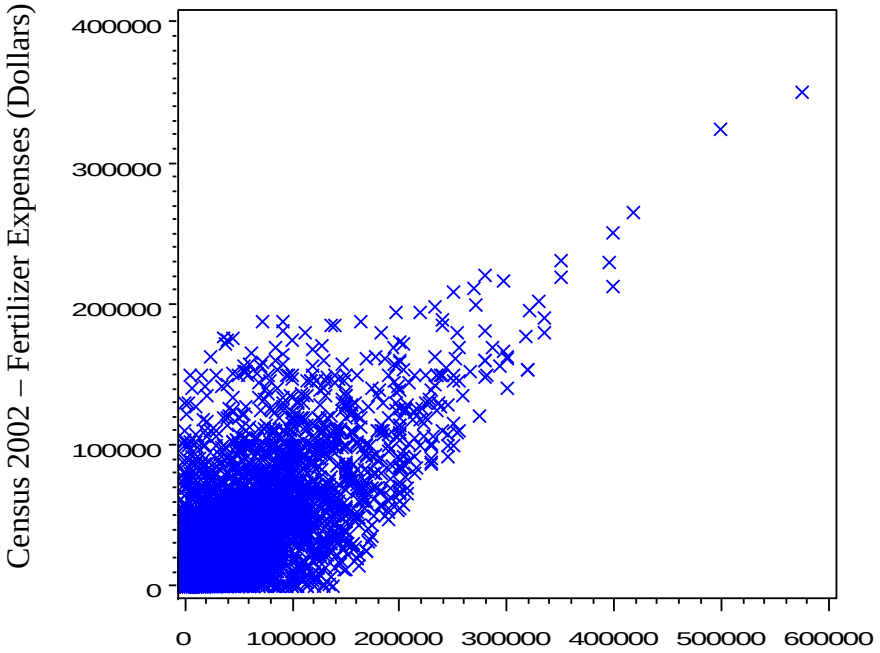


Figures A11-A12: Census 2002 versus ARMS III 2005 Fertilizer Expenses

Scatter Plot of Fertilizer Expenses with Outliers

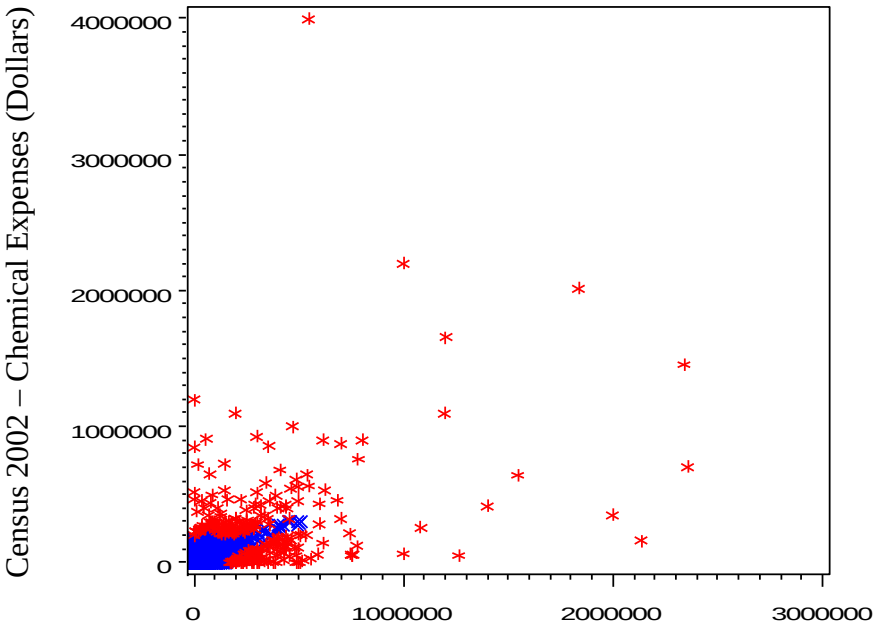


Scatter Plot of Fertilizer Expenses without Outliers

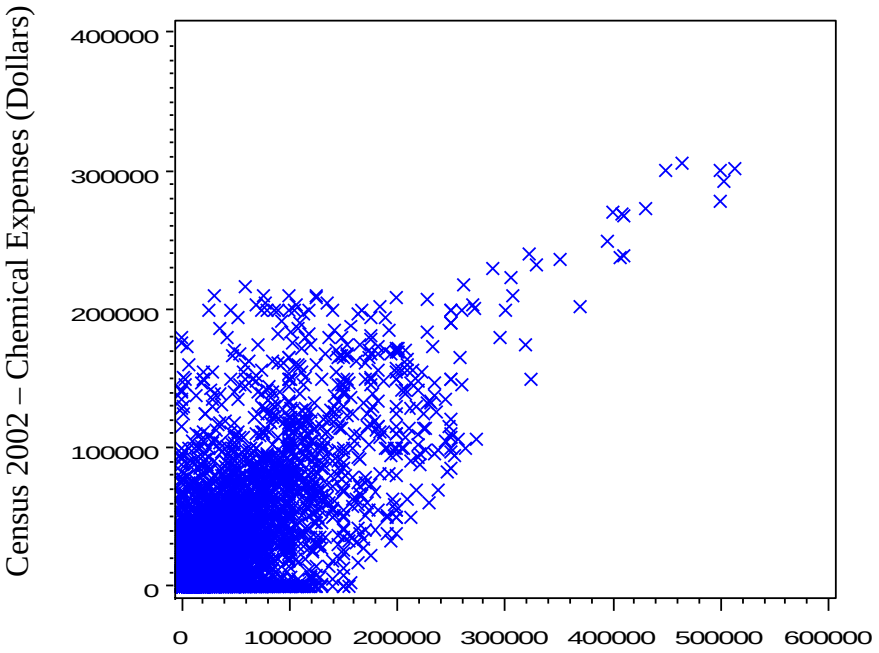


Figures A13-A14: Census 2002 versus ARMS III 2005 Chemical Expenses

Scatter Plot of Chemical Expenses with Outliers

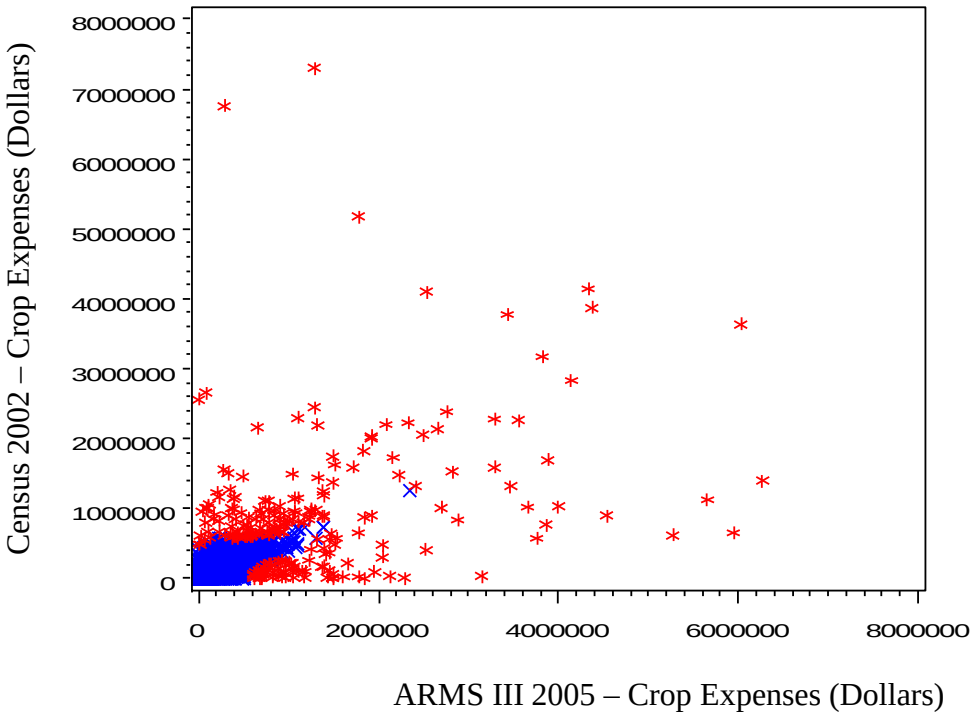


Scatter Plot of Chemical Expenses without Outliers

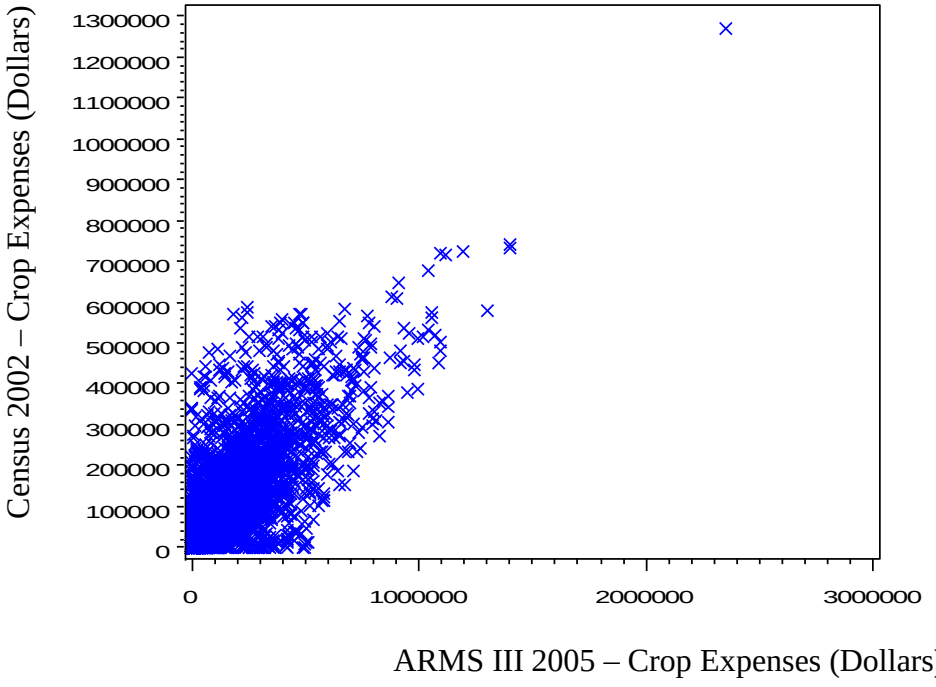


Figures A15-A16: Census 2002 versus ARMS III 2005 Crop Expenses

Scatter Plot of Crop Expenses with Outliers

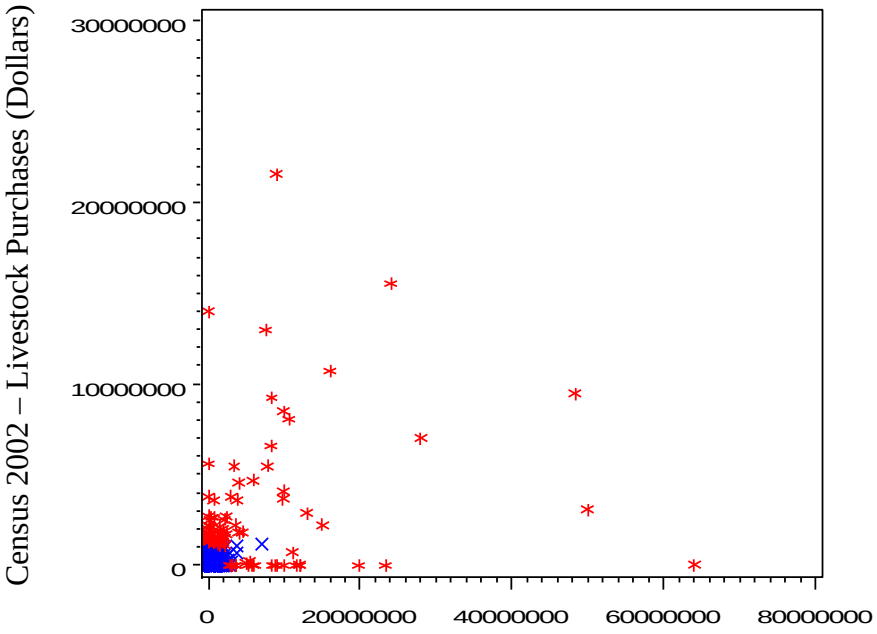


Scatter Plot of Crop Expenses without Outliers

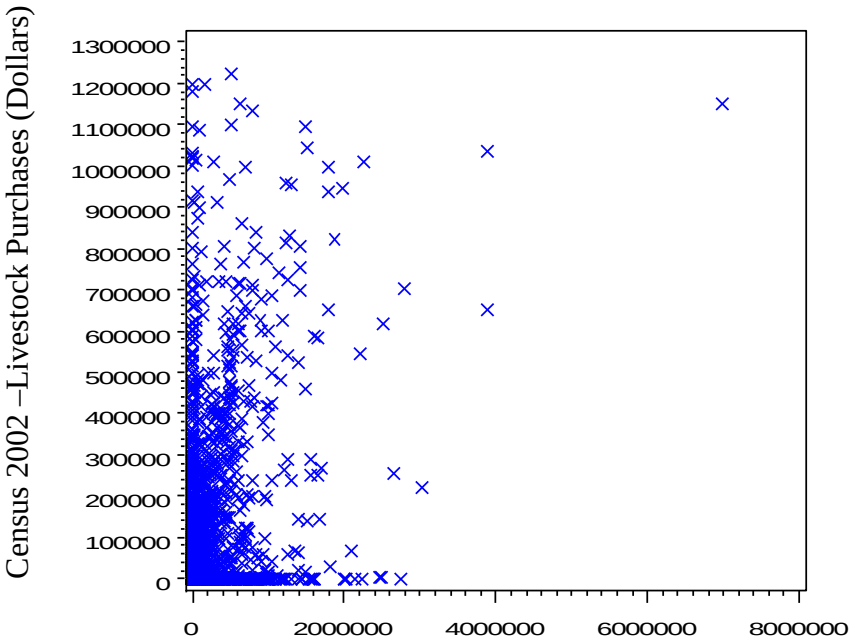


Figures A17-A18: Census 2002 versus ARMS III 2005 Livestock Purchases

Scatter Plot of Livestock Purchases with Outliers

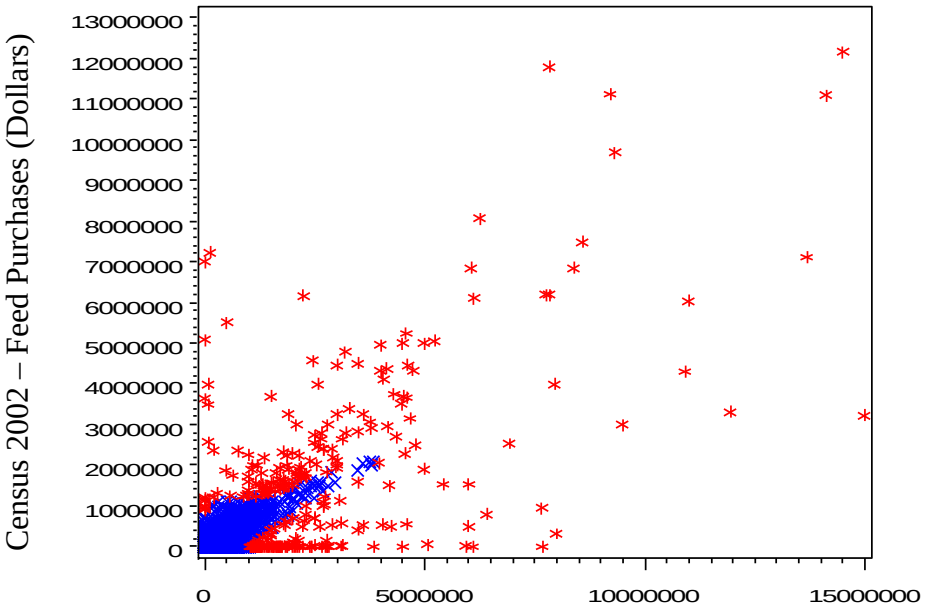


Scatter Plot of Livestock Purchases without Outliers

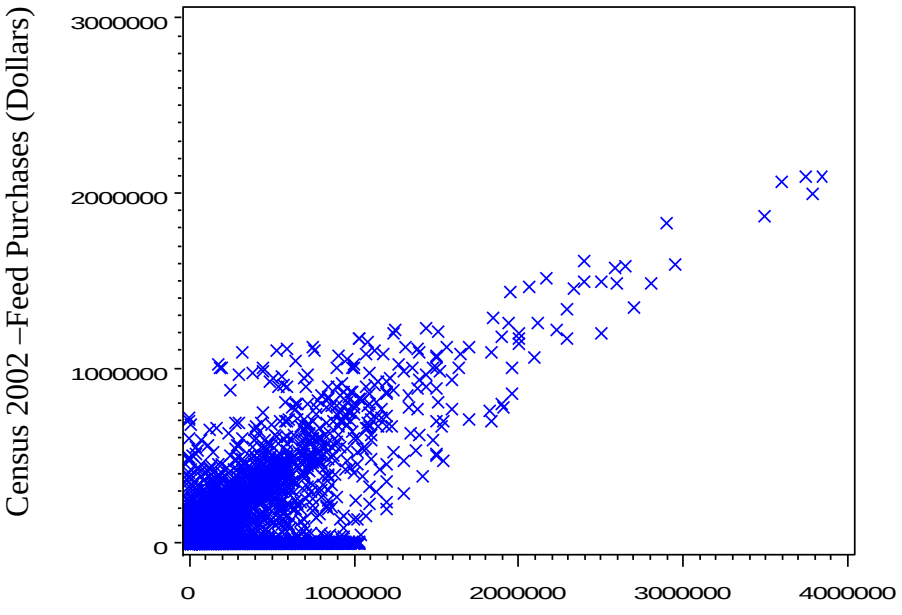


Figures A19-A20: Census 2002 versus ARMS III 2005 Feed Purchases

Scatter Plot of Feed Purchases with Outliers

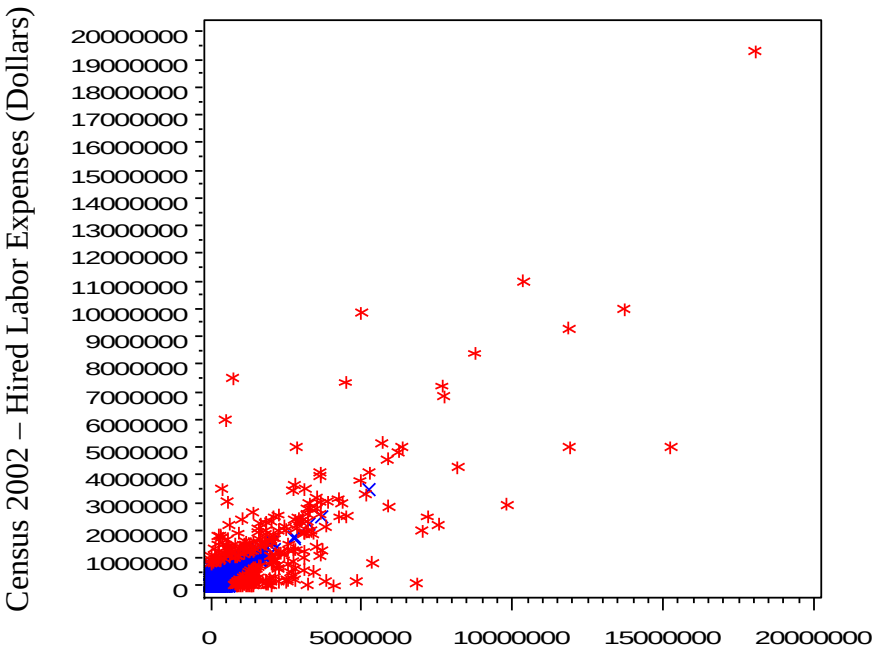


Scatter Plot of Feed Purchases without Outliers

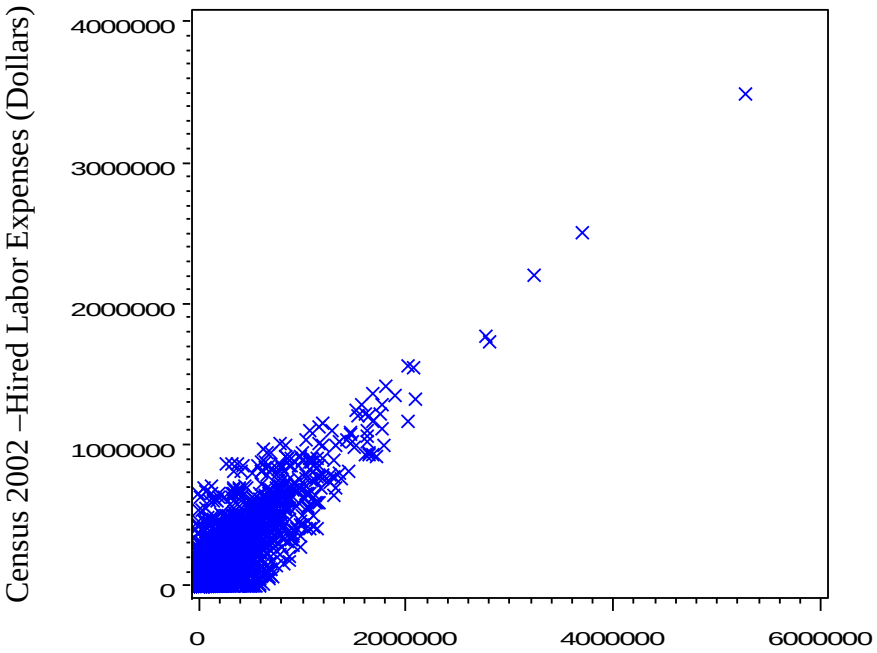


Figures A21-A22: Census 2002 versus ARMS III 2005 Hired Labor Expenses

Scatter Plot of Hired Labor Expenses with Outliers

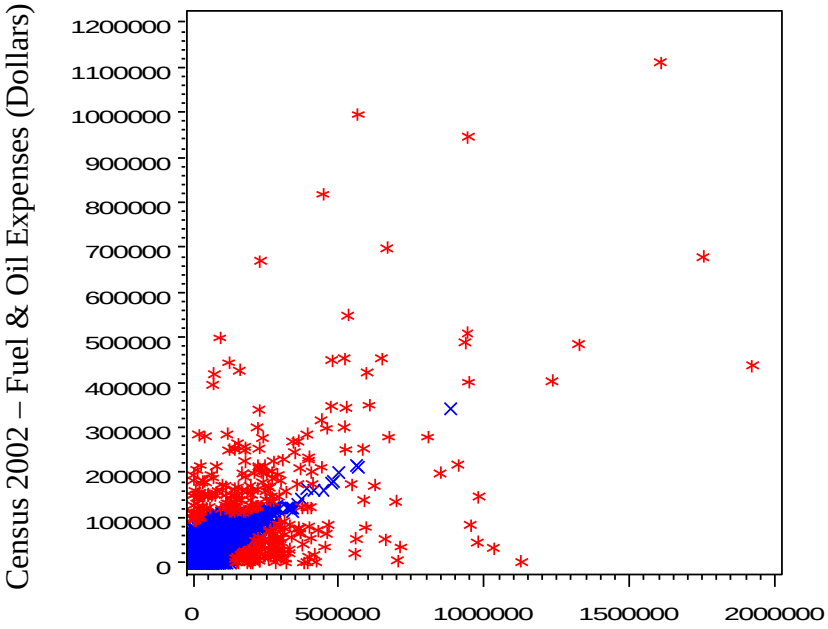


Scatter Plot of Hired Labor Expenses without Outliers

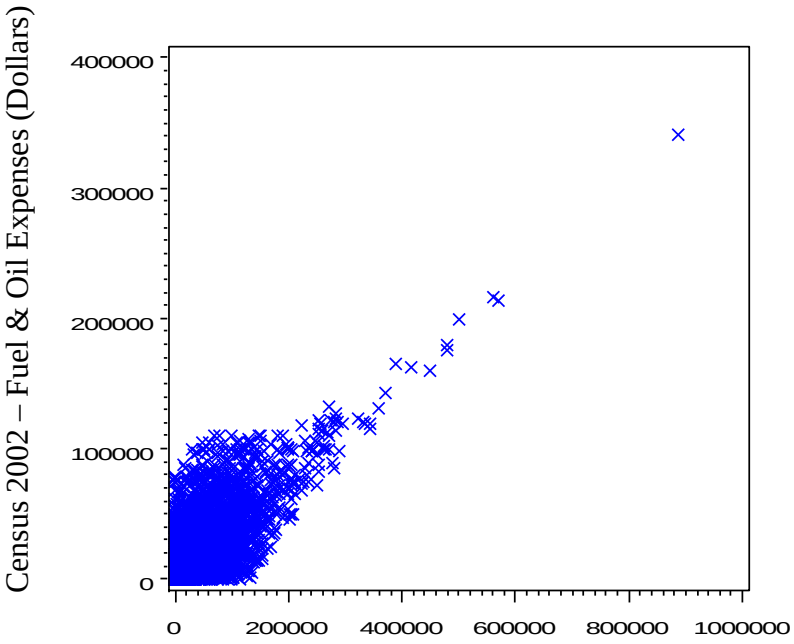


Figures A23-A24: Census 2002 versus ARMS III 2005 Fuel & Oil Expenses

Scatter Plot of Fuel & Oil Expenses with Outliers

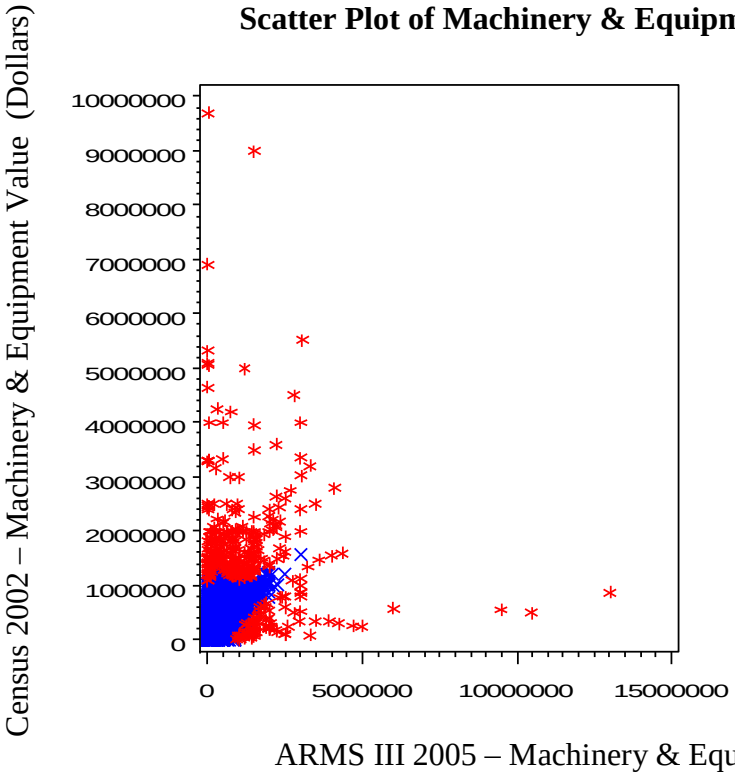


Scatter Plot of Fuel & Oil Expenses without Outliers

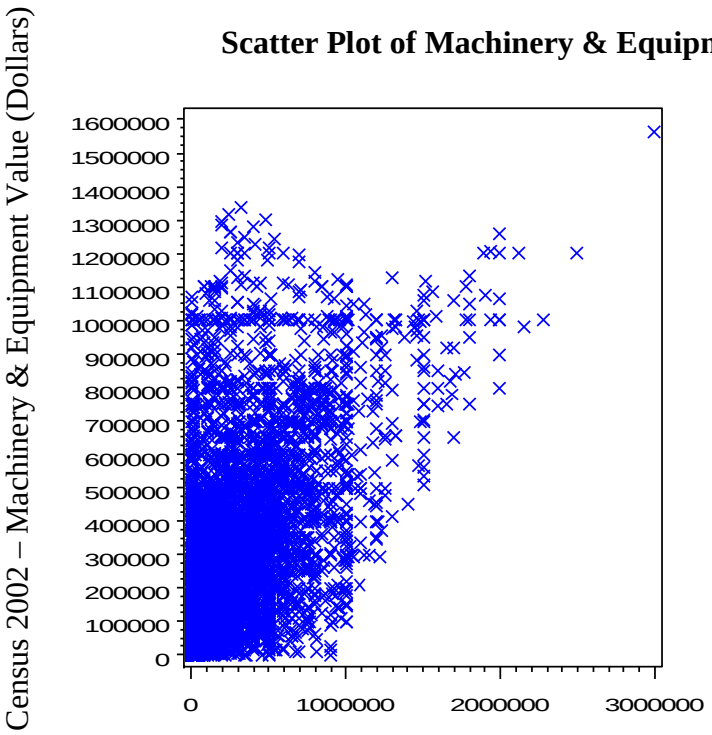


Figures A25-A26: Census 2002 versus ARMS III 2005 Machinery & Equipment Value

Scatter Plot of Machinery & Equipment Value with Outliers

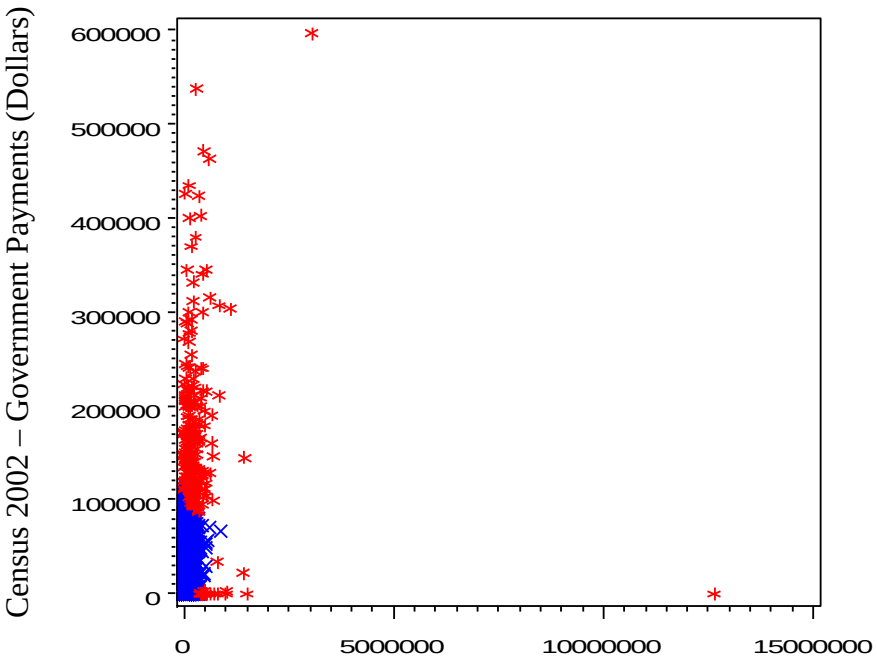


Scatter Plot of Machinery & Equipment Value without Outliers

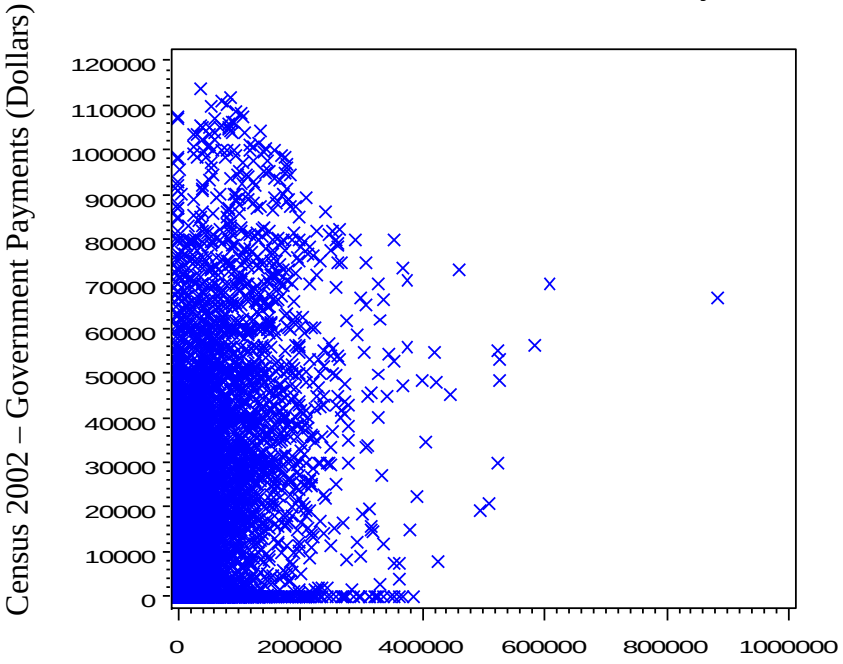


Figures A27-A28: Census 2002 versus ARMS III 2005 Government Payments

Scatter Plot of Government Payments with Outliers

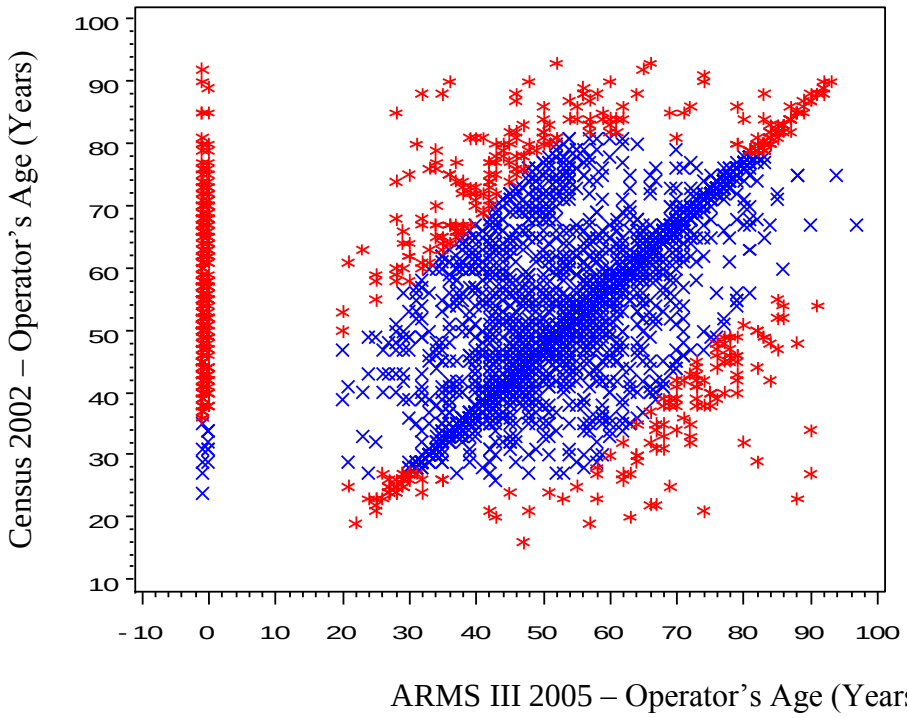


Scatter Plot of Government Payments without Outliers

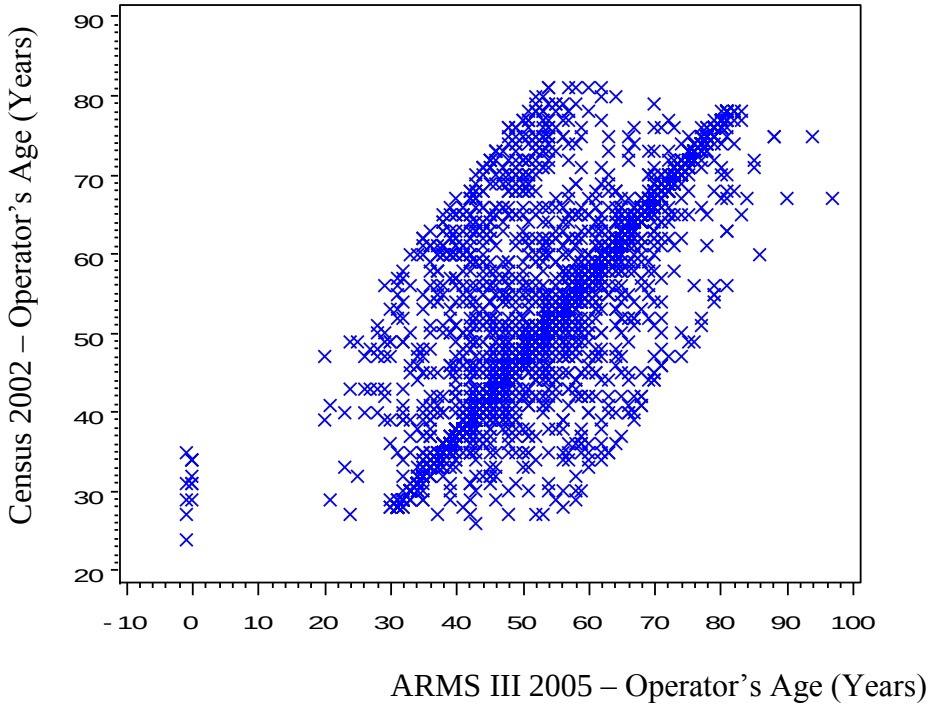


Figures A29-A30: Census 2002 versus ARMS III 2005 Operator's Age

Scatter Plot of Operator's Age with Outliers

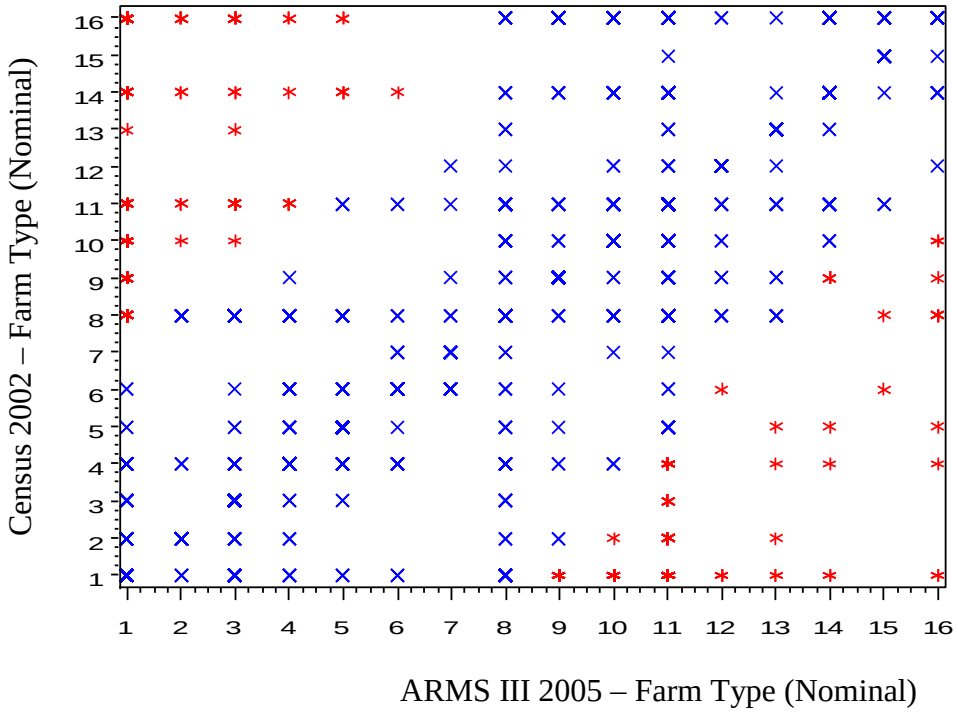


Scatter Plot of Operator's Age without Outliers



Figures A31-A32: Census 2002 versus ARMS III 2005 Farm Type

Scatter Plot of Farm Type with Outliers



Scatter Plot of Farm Type without Outliers