United States
Department of
Agriculture

National
Agricultural
Statistics
Service

Research Division

SRB Research Report
Number SRB-94-11

November 1994

Investigating The Development Of A Peak Worker Index Based on Selected Control Data

Tom Pordugal

INVESTIGATING THE DEVELOPMENT OF A PEAK WORKER INDEX BASED ON SELECTED CONTROL DATA, by Tom Pordugal. Ohio Applications Research Section, Survey Research Branch, Research Division, National Agricultural Statistics Service, United States Department of Agriculture, Washington, DC 20250-2000, August, 1994, Report No. SRB-94-11.

ABSTRACT

Farm value of sales and specialty/nonspecialty crops are used to stratify List Sampling Frame (LSF) records for the Agricultural Labor Survey. With the use of Indiana's and Florida's master file, this project attempts to develop a model and/or models that consider elements of farm labor as it relates to peak number of workers. This peak worker index number would then be assigned to those operations without control data for peak number of workers. A farm labor model could be used in other corn belt States like Ohio and in other highly labor-intensive States like California to assign other farming operations an index for peak number of workers. Analyses were conducted using 1993 control data from Indiana's list master file and 1994 control data from Florida's list master file. Simple correlation analyses examined relationships between peak number of workers and other independent variables. For both States the correlations did not show significance among variables. When these variables were introduced into a regression model, efficient peak worker indices could not be generated due to low R-squares among the independent variables.

KEY WORDS

Agricultural Labor Survey; Correlation; R-square; Regression; Control data; List Sampling Frame; Classification; Stratification.

ACKNOWLEDGEMENTS

The author of this paper would like to thank Lee Brown and Cheryl Turner for reviewing both the early and final drafts of this paper and Jim Davies for suggesting a comparison study with a corn belt State like Indiana and a labor-intensive State like Florida.

TABLE OF CONTENTS

SUMMARY iii

INTRODUCTION 1

METHODS 1

INDIANA ANALYSIS 2

FLORIDA ANALYSIS 6

CONCLUSIONS AND RECOMMENDATIONS 9

REFERENCES 10

APPENDIX A: INDIANA CORRELATION ANALYSES TABLES 11

APPENDIX B: INDIANA SCATTERPLOT ANALYSES 16

APPENDIX C: INDIANA REGRESSION ANALYSES TABLES 19

APPENDIX D: FLORIDA CORRELATION ANALYSES TABLES 26

APPENDIX E: FLORIDA SCATTERPLOT ANALYSES 34

APPENDIX F: FLORIDA REGRESSION ANALYSES TABLES 37

SUMMARY

Currently, farm value of sales and specialty/nonspecialty crops are used to stratify LSF records for the Agricultural Labor Survey in addition to control data for the peak number of workers. This project attempted to develop a model that would consider elements related to farm labor, such as type of farm and farm value of sales, for use in assigning a peak worker index number to those operations without peak worker control data. Hopefully, an improved sampling population would be available for the classification of the Agricultural Labor Survey.

The peak worker index study attempted to use selected control data variables in correlation and regression analyses. The study consisted of two parts: an analysis based on Indiana's 1993 list master file and one based on Florida's 1994 list master file. The modeling was based on a portion of the active records taken from these master files which contained 96,249 active records for Indiana and 72,485 active records for Florida. These master files seemed appropriate because Indiana and Florida both captured zeros for peak number of workers in their control data. By having "zeros" assigned to those operations, the resulting models would not have an upward bias which would skew the possible results of the research. Records containing control data relating to farm labor were extracted and used for the correlation and regression analyses.

The correlation and regression analyses were conducted based on unstratified data and data stratified by farm type and/or farm value of sales. It was found that the low correlation coefficients did not produce reliable relationships. Linear and polynomial regression models were explored. Because of the low correlation coefficients, linear transformations were performed. These transformations did not provide large enough R-squares to make an adequate fit of the Indiana and Florida data.

INTRODUCTION

Farm value of sales and specialty and/or nonspecialty crops are used to stratify LSF records with missing peak number of workers control data. The list frame is stratified into groups of farms that are believed to be similar with respect to farm value of sales or number of hired workers. During spring classify, new labor samples are drawn in all states based on this stratification. To prepare for classify and sample select, States maintain and update control data used for the classify. Important agri-data in the LSF include the following: fruit indicator, vegetable indicator, greenhouse indicator, fruit acreage, nut acreage, vegetable acreage, floriculture sales, sod farms, nurseries, total land, total cropland, and livestock items.

This system of using farm value of sales and specialty/non-specialty crops to stratify LSF records with missing peak number of workers control data does not provide enough information about the size of a labor force to adequately stratify the farm labor sample. Moreover, the current system of assigning farm value of sales does not take into account labor needs for specific labor intensive crops and other factors which influence the size and dimension of farm labor. This research evaluated different types of operations based on different labor characteristics. The research focused on examining and evaluating control data related to farm labor, i.e., examining relationships utilizing stratifications based on crop acreages, number of livestock, type of farming operation, and/or farm value of sales. Therefore, the purpose of this project was the following: to use regression models in the derivation of a peak worker index and to assign this generated peak worker index to those list frame operations with no peak number of workers control data.

METHODS

The SAS procedure PROC CORR was used to find possible relationships that could be used for predicting peak number of workers. This consisted of examining farm acreage control data with peak number of workers. Similarly, number of livestock control data were explored. Other relationships were examined based on the stratification of acres and livestock data by type of farm. Then, a farm value of sales control variable was examined with peak number of workers. Similar analyses examined specific land-related variables (like cropland, all land, vegetable land, fruit land, tobacco, corn, soybeans, and wheat acres) with peak number of workers as well as the number of head control variables for hogs, dairy cows, all cattle, and poultry with peak number of workers. Moreover, tests were run with stratifications by farm acres, number of head, and farm value of sales with the peak number of workers control variable.

Since those initial attempts to find significant correlations proved inconclusive, scatterplots of selected data were constructed and examined using the SAS procedure PROC PLOT to find any non-linear trends and/or outliers. The attempts showed that possible linear or higher-order trends in the data were not evident. Much of the data was replicated at the same location on the plots, which made detection of non-linear trends difficult. Moreover, the appearance of

outliers confused detection.

After these initial attempts, the original extracted data files were examined for apparent inconsistencies by comparing control data within records for both Indiana and Florida. A printout of one hundred records of the extracted file proved inconsistencies exist in approximately twenty percent of the observations. Inconsistencies appeared across records when peak number of workers was compared with the type of farm and farm value of sales control data and when peak number of workers was compared with the land type acres and the livestock control data. These inconsistencies probably introduced outliers and contributed to the low correlations in the previous analyses.

These types of inconsistencies probably appeared throughout the control data because they tended to skew the data with outliers and gave inconclusive results with low correlations. Scatterplots performed on the data supported the fact that these inconsistencies contributed to the low correlations.

The SAS procedure PROC REG /SELECTION = R-SQUARE was used to develop simple linear and/or multiple regression models. Transformations on the independent variables were executed since the correlation results were poor. Various regression models were fitted to the data. Different models tested included a combination of first-order and second-order models without interaction terms, since interaction terms would be hard to interpret. Low R-squares in all models showed that an adequate fit could not be made.

INDIANA ANALYSIS

List frame records were extracted from Indiana's list master file based on zero or positive data for peak number of workers. The extracted file captured 6,263 records. Of these records, 358 had a value of zero for peak number of workers, and 113 contained inconsistent control data for type of farm and farm value of sales. These observations were omitted because the other pertinent control data that determined type of farm and farm value of sales were assigned to zero or inconsistent. A type of farm variable assigned to zero or missing in the control data, or a farm value of sales variable assigned to missing in the control data did not show any promise for analysis. With these record deletions, the resulting data set for Indiana consisted of 6,150 observations.

Modeling the peak number of worker data takes into account many factors that may or may not be present in the control data. Correlation analyses were performed and subsequent regression R-square values were produced based on the control data.

Initial analyses involved specific control data variables that would be expected to show relationships with peak number of workers. The following variables were selected: vegetable land, cropland, all land, tobacco, corn, soybeans, wheat, hogs, dairy cows, all cattle, poultry, and chickens. Maybe these variables would show adequate information about the peak number of workers to stratify the records.

The correlation tests were conducted based on unstratified data and data stratified by farm type and/or farm value of sales, and results from those tests are found in

Appendix A. Table 1 shows the correlation coefficient and the number of records for each of the selected control variables. Table 2 stratifies the selected control variables by farm value of sales, and Table 3a stratifies the variables by farm type and farm value of sales. Table 3b stratifies the all land control variable based on each of the farm types by farm value of sales.

The following can be mentioned about the four tables. Table 1 showed that poultry had the highest correlation with peak number of workers. Table 2 showed that the correlation increased slightly in the farm value of sales stratum greater than $250,000. No other apparent gains were made when stratifying the original selected control data.

Since the acreage and most of the livestock selected control data did not show promise for additional analyses, several acreage variables and livestock number of head variables were summed together. In this way, it was hoped that the number of records would increase and the correlation coefficients would improve. The summing of records was contrived so that these summed records could still be stratified by the farm type variable.

Correlation tests were run between peak number of workers and the following: the summed acreage variables, the number of head variables, the acreage variables and the farm type variable, and the number of head variables and the farm type variable. The analyses were first conducted with unstratified data and then with stratified data using farm type and/or farm value of sales control data.

The different acreage and number of head variables were generated in the following manner. As stated in the "Agricultural Labor Survey: Supervising and Editing Manual" for 1994, the following variables were generated based on the type of operation. A cash grain acre variable was generated across records by summing control data for corn, soybeans, wheat, oats, sorghum, barley, and rye. Likewise, an "other field crop" variable summed all hay, sunflower, canola, popcorn, and all mint variables across records. Fruit acres consisted of apples, apricots, blackberries, blueberries, cherries, nectarines, grapes, plums, peaches, pears, and strawberries. The vegetable crop variable summed asparagus, snapbeans, cabbage, cantalope, honeydew,sweet corn, tomatoes, and watermelon. A poultry variable summed chickens, turkeys, and pullets. Lastly, a livestock variable totaled cattle, horses, hogs, and sheep number of head control data.

Table 4a shows the results of the test with peak number of workers and the acreage variables; Table 4b shows the results of a similar test with peak number of workers and the acreage variables, where the acreage variables are associated with their type of farm variable. Likewise, Table 5a shows the results of the test with peak number of workers and number of head; Table 5b shows the results of a test with peak number of workers and number of head, where number of head is associated with their farm type variable. For example, only those observations that had positive cash grains reported and indicated a cash grains type of farm were included in Table 4b. Refer to Appendix A.

Of the number of records in Tables 4a and

5a, only those records that had positive (indicator) data for the associated farm type variable were analyzed in Tables 4b and 5b. With the motivation of improving correlation, the correlation coefficients in these two tables again showed low to average coefficients. When the acreage variables were analyzed by farm type, the number of records on a whole dropped by approximately forty percent due to missing values in the control data for farm type. Although the control data for Indiana captured data for greenhouse operations, the control data was indicator-type and could not be included in these initial analyses; correlation analyses could only be run based on farm value of sales and peak number of workers, with and without its associated farm type. See Appendix A.

Further stratifications were performed based on farm value of sales. By introducing the farm value of sales variable into the correlation analysis, the data could be further stratified and the coefficients could be scrutinized. Table 6 delineates the results of the test with peak number of workers and farm value of sales. Correlations were then run on the stratification of farm value of sales. In this way, the acreage and number of head variables would be grouped and correlations could be run based on the farm value of sales strata with and without the farm type variable. The first analysis used acreage or number of head variables without the farm type classification. See Tables 7a and 7b. The second analysis considered the acreage or number of head variables and took into consideration their associated farm type variable. Refer to Tables 8a and 8b. See Appendix A.

When comparing the two types of analyses, improvements in correlation were noticed with some of the variables when type of farm was used as a stratification variable. Negligible differences were noted with the cash grain, fruit, livestock, and poultry variables. However, gains were made with the tobacco variable in the farm value of sales strata, i.e., the $20,000 to $100,000 strata. Improvements were made across all strata in the other field crops variable, but the correlations were not high enough to be significant. The vegetable variable improved in the $20,000 to $100,000 farm value of sales strata; there were only ten statistics to consider, however.

The correlation analyses generally produced low coefficients between peak number workers and these generated variables. The stratification by farm value of sales did not improve the correlations with peak number of workers. Again, records were lost in stratification because individual records did not contain all the control data necessary for the analyses. Due to missing control data across many of the records, these records were omitted from the analysis.

Because of the low correlation coefficients, scatterplots of the data were constructed to find any possible non-linear trends. Some of the data were clustered near the intersection of the axes. Many of the individual data points were replicated in the peak number of workers and total farm acres scatterplots, the peak number of workers and the number of head scatterplots, and the peak number of workers and farm value of sales scatterplots. No other type of trends, i.e., quadratic or logarithmic, could be recognized or justified because of this

duplication and\or haphazard scattering of data points in the unstratified and stratified data.

Tables 1, 2, and 3 in Appendix B show scatterplots representative of the correlations in the previous discussion. These scatterplots correspond to the correlations in Tables 4a, 5a, and 6. Tobacco, poultry, and farm value of sales control data were selected because they provided some of the higher correlations. They also exhibited characteristics of scattering detailed in the previous paragraph.

Since these attempts proved inconclusive toward significance, linear regression model generations seemed futile. Moreover, in examining the original extracted data file, many inconsistencies seemed apparent. For example, those records which seemed less labor intensive reported high labor and vice versa. They tended to skew the data with probable outliers and gave inconclusive correlations.

In a last attempt, transformations on the independent variables, i.e., total acres, number of head, and farm value of sales, were performed. Instead of fitting a first-order model, a polynomial model was tried. Transformations included the square, the log, and the square root. The following was the full model used to predict peak number of workers:

$$(\text{Peakwork})_i = \beta_o + \beta_1(\text{Acres})_i + \beta_2[(\text{Acres})^2_i] + \beta_3[\text{Sqrt}(\text{Acres})_i] + \beta_4[\text{Ln}(\text{Acres})_i] + \beta_5(\text{No. of Head})_i + \beta_6[(\text{No. of Head})^2_i] + \beta_7[\text{Sqrt}(\text{No. of Head})_i] + \beta_8[\text{Ln}(\text{No. of Head})_i] + \beta_9(\text{Farm Value of Sales})_i + \beta_{10}[(\text{Farm Value of Sales})^2_i] + \beta_{11}[\text{Sqrt}(\text{Farm Value of Sales})_i] + \beta_{12}[\text{Ln}(\text{Farm Value of Sales})_i] + \epsilon_i.$$

The farm type variable was omitted in the full model since it was not significant in the first-order model. Farm type was utilized only in stratification. Table 1 and Table 2 of Appendix C show the R-squares for the full regression model without number of head variables in the analysis and the R-square values using the full model, respectively.

The R-square values were generated using PROC REG /SELECTION = R-SQUARE; there was a stop placed after the fourth variable entered the model. Although the R-squares improved (see Table 2), none of the variables showed R-squares large enough to pursue fitting peak number of workers into the models. Many of the records did not contain all the necessary control data to compute the regression. The full model analyzed 402 records with acres variables and number of head variables. Of the 5,749 records used in the regression with the acreage variables (see Table 1), 5,347 records contained missing data for the number of head variables. Because of the missing values, the full model analyzed the remaining 402 records. See Table 2.

Regression analyses were then attempted using stratified data. Stratifications were first performed on farm value of sales alone without farm type. Table 3a and Table 3b show the resulting R-squares. Stratifications were then performed on farm value of sales based on farm type, and regression analyses using the full model were performed on these two-way strata. See Tables 4a, 4b, 4c, and 4d. Reductions in the number of records were

again noted in each of these tables. Refer to Appendix C for these tables.

Because of the low correlations and inconclusive scatterplots, the data, both unstratified and stratified, generally did not produce R-squares of at least 0.80. Fitting the data into regression models to predict peak number of workers proved inconclusive.

FLORIDA ANALYSIS

List frame records were extracted from Florida's list master file based on zero or positive data for peak number of workers. Of the 12,236 extracted records, 4,706 records had a value of zero assigned to peak number of workers, and 521 records contained inconsistent control data in Florida. Any type of farm variable assigned to zero or missing in the control data, or any farm value of sales variable assigned to missing in the control data did not show any promise for analysis. With these additional deletions, the resulting data set for Florida consisted of 11,715 records.

Like Indiana, specific correlation tests examined selected control data from Florida's master that may show relationships with peak number of workers. The following variables were selected: vegetable land, cropland, all land, tobacco, corn, soybeans, wheat, hogs, all cattle, dairy cows, chickens, aquaculture, and floriculture.

In Appendix D, results are shown from the correlation tests that were performed with peak number of workers using unstratified data and with the data stratified by farm type and/or farm value of sales. Table 1 shows the correlation coefficient and the number of records for each of the selected control variables, excluding the aquaculture and floriculture control variables. Table 2 stratifies the selected control variables by farm value of sales, and Table 3a stratifies the variables by farm type and farm value of sales. Table 3b stratifies the all land control variable based on each of the farm types by farm value of sales. Because aquaculture and floriculture are indicator variables, correlation analyses could only be conducted based on stratifications with farm value of sales and farm type. For this reason Table 1 does not contain the aquaculture and floriculture variables.

In comparing the four tables, the following results can be mentioned. Table 1 showed that dairy cow control data had the highest correlation with peak number of workers. However, no improvements were made when dairy cow data were stratified by farm type and farm value of sales. Chicken control data had the second highest correlation, and improvements were noted in the first two strata of Table 2. No other apparent gains were made when stratifying the original control data.

Since the acreage and most of the livestock selected control data did not show promise for additional analyses, several acreage variables and livestock number of head variables were summed together. In this way, it was hoped that the number of records would increase and also improve any correlation. The summing of records was contrived so that these summed records could still be stratified by the farm type variable.

Correlation tests were run between peak

number of workers and the following: the summed acreage variables, the number of head variables, the acreage variables and the farm type variable, and the number of head variables and the farm type variable. The analyses were performed with unstratified data first and then with stratified data by farm type and/or farm value of sales.

The acreage variables and number of head variables were generated in the following manner. According to the "Agricultural Labor Survey: Supervising and Editing Manual" for 1994, a cash grain farm acre variable can be generated by summing control data for corn, sorghum, oats, rice, rye, and soybeans. An "other field crop" farm acre variable summed hay, peanuts, potatoes, and sugarcane. Vegetable farm acres consisted of snapbeans, beets, broccoli, cabbage, cantaloupe, carrots, collards, kale, cauliflower, celery, eggplant, endive, cucumber, lettuce, okra, onions, peppers, spinach, squash, peas, sweet corn, tomatoes, watermelon, radishes, and mustard greens. Fruit farm acres consisted of avocados, blackberries, strawberries, nectarines, grapes, plums, peaches, mangoes, grapefruit, oranges, lemons, limes, tangerines, and tangelos. The nuts variable consisted of pecan control data. The livestock variable summed the number of head of all cattle, hogs, sheep, and goats. Finally, a poultry number of head variable summed chicken control data. Because Florida and Indiana produce different types of commodities, the consequential total farm acres variables and number of head variables summed different control data.

The following tables were produced. Table 4a shows the results of the correlation tests with peak number of workers and the acreage variables. Table 4b shows the results of a similar test with peak number of workers and the acreage variables, where total farm acres is associated with their farm type variable. Similarly, Table 5a shows the results with peak number of workers and number of head. Table 5b shows the results with peak number of workers and number of head, where number of head is associated with its farm type variable.

Correlation analyses produced low to average correlation coefficients. Notice that when the total farm acres was stratified by the farm type variable, the number of records on a whole dropped by approximately forty percent. When the total farm acres and number of head variables were stratified by farm type, the number of records consequently decreased and the correlation coefficients did not improve across all variables. However, the correlation coefficients for bees, horses, and livestock number of head variables improved when each was stratified by its farm type variable. The coefficients for bees, horses, and livestock jumped from 0.330 to 0.972, 0.093 to 0.718, and 0.055 to 0.245, respectively.

With the hopes of improving correlation, further stratifications were performed based on farm value of sales. Table 6 shows the correlation coefficients with peak number of workers and farm value of sales. Table 7a looks at the stratification of the acreage variables by farm value of sales. Table 7b presents the analysis based on number of head by farm value of sales. Table 8a shows the correlations with peak number of workers and the acreage variables stratified by farm type and farm

value of sales. Similarly, Table 8b shows the correlations with peak number of workers and number of head stratified by farm type and farm value of sales. Table 8a and Table 8b both possessing the farm type variable are subsets of Table 7a and Table 7b, respectively.

The stratification by farm value of sales did not improve the correlations with peak number of workers. Minimal gains were made with correlation generally as the data was further stratified. Again, records were lost in stratification because individual records did not contain all the control data necessary for the analyses. Due to missing control data across many of the records, these records were excluded from the correlation analyses.

Like the Indiana data, various scatterplots of the Florida data based on these variables were constructed to find possible non-linear trends. None of the plots showed resemblance to a quadratic, logarithmic, or other higher-order fit. Duplication and/or random scattering of the data made detection of trends difficult.

Tables 1, 2, and 3 in Appendix E show scatterplots representative of the correlations. These scatterplots correspond to the correlations in Tables 4a, 5a, and 6 of Appendix D. Cash grain, poultry, and farm value of sales control data were selected because they provided some of the scattering detailed in the previous discussion.

Since the correlation analyses showed that the relationships with peak number of workers produced low correlation coefficients, first-order linear regression models did not seem possible. In a last attempt, transformations on the independent variables, i.e., total acres, number of head, and farm value of sales, were performed. Instead of fitting a first-order model, a polynomial model was tried. Transformations included the square, the log, and the square root transformations. The following was the full model used to predict peak number of workers:

$$(\text{Peakwork})_i = \beta_o + \beta_1(\text{Acres})_i + \beta_2[(\text{Acres})^2_i] + \beta_3[\text{Sqrt}(\text{Acres})_i] + \beta_4[\text{Ln}(\text{Acres})_i] + \beta_5(\text{No. of Head})_i + \beta_6[(\text{No. of Head})^2_i] + \beta_7[\text{Sqrt}(\text{No. of Head})_i] + \beta_8[\text{Ln}(\text{No. of Head})_i] + \beta_9(\text{Farm Value of Sales})_i + \beta_{10}[(\text{Farm Value of Sales})^2_i] + \beta_{11}[\text{Sqrt}(\text{Farm Value of Sales})_i] + \beta_{12}[\text{Ln}(\text{Farm Value of Sales})_i] + \epsilon_i.$$

The farm type variable was omitted in the full model since it was not significant in the first-order model. Farm type was utilized only in stratification. Table 1 and Table 2 of Appendix F show the R-squares for the regression model without number of head in the analysis and the R-square values using the full model, respectively. The R-square values were generated using PROC REG /SELECTION = R-SQUARE; there was a stop placed after the fourth variable entered the model. Like Indiana, none of the variables showed R-squares large enough to pursue fitting peak number of workers into the models. Refer to Appendix F for Florida's regression results.

Regression analyses were then attempted using stratified data. Stratifications were first performed on farm value of sales alone without farm type. Table 3a and Table 3b show the resulting R-squares.

Stratifications were then performed on farm value of sales based on farm type, and regression analyses using the full model were performed on these two-way strata. See Tables 4(a1) and 4(a2), Tables 4(b1) and 4(b2), Tables 4(c1) and 4(c2), and Tables 4(d1) and 4(d2).

Although the R-squares improved in Table 2, none of the variables showed R-squares large enough to pursue fitting peak number of workers into the models. Most of the records did not contain all the necessary control data to perform regression. The full model analyzed 232 records with acreage variables and number of head variables. Of the 6,069 records used in the regression with the acreage variables (see Table 1), 5,837 records possessed missing data for the number of head variables. Because of the missing values, the full model analyzed the remaining 232 records. See Table 2.

Because of the low correlations and inconclusive scatterplots, the data, both unstratified and stratified, generally did not produce R-squares of at least 0.80. Fitting the data into regression models to predict peak number of workers proved inconclusive.

CONCLUSIONS AND RECOMMENDATIONS

When extracting records from a large list of operations, it is hoped that they are representative of the total population. In the given research, the data were extracted from Indiana's and Florida's list master file based on zero or positive control data for peak number of workers. The only sample taken was the 100 records selected to evaluate data quality for inconsistencies. The data sets were the largest possible data sets for peak number of workers. The data sets were approximately seven and seventeen percent of Indiana's and Florida's master file, respectively, and they likely were not representative of the general population. Since many of the records had missing control data for farm type and farm value of sales, stratifications of peak number of workers by farm type and/or farm value of sales eliminated records from analysis. Too many inconsistencies and potential outliers for the peak number of worker control data tended to flaw and skew any possible significant results that could have resulted. This was apparent in the low correlation coefficients, the scatterplots, and the R-squares. Fitting the data to regression models was not investigated further because of these factors. Because the large amount of missing control data affected this study and respective analysis, it also affected the operational stratification and sampling effectiveness.

Maybe there were relationships left unexamined, but these relationships were difficult to expose with the given control data. The control data used in the study consisted of farm acres, number of head, farm type, and farm value of sales variables. Consequently, a broader study could not be investigated. Moreover, if seasonality-type control data were used in this study, the number of records would still be reduced as stratifications are performed. The use of seasonality-type control data would not be practical as the foregoing analyses portrayed. In Florida it appeared that the most promising variables involved specialty commodities or those not generally produced throughout the state, i.e., bees, horses, tobacco, and

cotton. Although there were some regression R-squares of at least 0.80 among the specialty, livestock, and poultry variables, these limited relationships would not be cost effective to make a system of generating peak number of worker indices operational just on those records.

At this point the research concluded that a peak worker index variable could not be generated using Indiana's and Florida's List Sampling Frame. It is recommended, based upon this analysis, that a peak number index not be created using current control information on the List Sampling Frame.

REFERENCES

"Agricultural Labor Survey: Supervising and Editing Manual" (1994), Unites States Department of Agriculture, National Agricultural Statistics Service.

Moussa, Effat A. (1991). Computer Integrated Probability and Statistical Methods for Science and Management: Part III, DePaul University, Chicago, IL.

SAS Institute Inc. (1988), SAS\STAT User's Guide, Release 6.03 Edition.

United States Department of Agriculture (1983), Scope and Methods of the Statistical Reporting Service, Publication No 1308. Washington, D.C.

APPENDIX A: INDIANA CORRELATION ANALYSES TABLES

Table 1: Correlation Analysis with Peak Number of Workers and Selected Control Variables

Selected Control Variables	Number of Records	Correlation Coefficient
Vegetable Land	174	0.187
Cropland	5,935	0.174
Tobacco	249	0.432
Corn	5,294	0.144
Soybeans	4,868	0.100
Wheat	3,954	0.133
Hogs	5,907	0.116
Dairy Cows	5,462	0.037
All Cattle	5,913	0.069
Poultry	374	0.770
All Land	5,952	0.175

Table 2: Stratified Correlation Analysis with Peak Number of Workers and Selected Control Variables (SCV) by Farm Value of Sales (FS)

Number of Records Correlation Coefficient	FS < $20,000	$20,000 <= FS < $100,000	$100,000 <= FS < $250,000	FS >= $250,000
SCV = Vegetable Land	17 -0.152	38 0.350	50 0.068	69 0.078
SCV = Cropland	826 0.035	1,770 -0.017	1,580 -0.085	1,759 0.168
SCV = Tobacco	64 0.370	133 0.281	36 0.596	16 0.344
SCV = Corn	452 0.204	1,600 -0.037	1,532 -0.045	1,710 0.114
SCV = Soybeans	336 -0.142	1,393 -0.057	1,461 -0.049	1,678 0.060
SCV = Wheat	250 -0.111	1,096 0.024	1,170 -0.060	1,438 0.114
SCV = Hogs	819 -0.025	1,765 -0.006	1,571 -0.018	1,752 0.098
SCV = Dairy Cows	667 -0.063	1,646 -0.013	1,489 0.024	1,660 0.021
SCV = All Cattle	822 0.116	1,767 0.029	1,569 0.030	1,755 0.046
SCV = Chickens	75 0.589	157 0.387	79 0.673	63 0.798
SCV = All Land	829 0.026	1,774 0.002	1,582 0.060	1,767 0.167

Table 3a: Stratified Correlation Analysis With Peak Number of Workers and Selected Control Variables (SCV) and Farm Type (FT) by Farm Value of Sales (FS)

Number of Records Correlation Coefficient	FS < $20,000	$20,000 < = FS < $100,000	$100,000 < = FS < $250,000	FS > = $250,000
FT= Cash Grain SCV= Cropland	321 -0.025	1,072 0.026	1,022 0.030	1,086 0.171
FT= Cash Grain SCV= Corn	293 -0.018	1,046 0.027	1,021 0.0004	1,082 0.167
FT= Cash Grain SCV= Soybeans	258 -0.041	1,034 0.002	1,016 0.010	1,085 0.084
FT= Cash Grain SCV= Wheat	165 0.027	754 0.069	782 -0.045	912 0.008
FT= Tobacco SCV= Tobacco	25 0.171	41 -0.061	1 *****	1 *****
FT= Vegetable SCV= Vegetable Land	7 -0.374	14 0.274	11 -0.253	30 0.085
FT= Livestock SCV= Hogs	363 -0.032	572 -0.090	502 -0.140	568 0.314
FT= Livestock SCV= Dairy Cows	293 -0.114	542 -0.034	487 0.118	548 0.077
FT= Livestock SCV= All Cattle	366 0.209	572 0.027	503 0.133	570 0.157
FT= Poultry SCV= Chickens	3 0.913	4 -0.580	19 0.586	32 0.783

Table 3b: Stratified Correlation Analysis with Peak Number of Workers and All Land Control Variable and Farm Type (FT) by Farm Value of Sales (FS)

Number of Records Correlation Coefficient	FS < $20,000	$20,000 < = FS < $100,000	$100,000 < = FS < $250,000	FS > = $250,000
FT= Cash Grain All Land	321 -0.030	1,073 0.019	1,022 0.043	1,088 0.167
FT= Tobacco All Land	25 -0.206	42 -0.142	1 *****	1 *****
FT= Oth. Fld. Crop All Land	78 0.124	36 0.430	4 0.815	5 0.291
FT= Vegetable All Land	7 -0.612	14 0.678	11 -0.006	31 -0.080
FT= Fruit All Land	10 -0.108	16 0.043	13 0.356	7 0.052
FT= Livestock All Land	366 0.074	572 0 003	503 0.054	570 0.520
FT= Poultry All Land	0 *****	4 -0.509	11 -0.152	38 0.262

Table 4a: Correlation Analysis with Peak Number of Workers and Acreage Variables

Acreage Variables	Number of Records	Correlation Coefficient
Cash Grain	5,443	0.137
Tobacco	249	0.432
Other Field Crops	3,702	0.337
Vegetables	110	0.268
Fruits	67	0.015

Table 4b: Correlation Analysis with Peak Number of Workers and Acreage Variables by Farm Type

Farm Type	Number of Records	Correlation Coefficient
Cash Grain	3,499	0.170
Tobacco	68	0.362
Other Field Crops	122	0.598
Vegetables	43	0.392
Fruits	38	-0.012

Table 5a: Correlation Analysis with Peak Number of Workers and Number of Head Variables

Number of Head	Number of Records	Correlation Coefficient
Livestock	5,917	0.124
Poultry	417	0.501

Table 5b: Correlation Analysis with Peak Number of Workers and Number of Head Variables by Farm Type

Farm Type	Number of Records	Correlation Coefficient
Livestock	2,026	0.349
Poultry	52	0.476

Table 6: Correlation Analysis with Peak Number of Workers and Farm Value of Sales (FS)

Farm Value of Sales (FS)	Number of Records	Correlation Coefficient
FS < $20,000	889	0.060
$20,000 <= FS < $100,000	1,814	0.214
$100,000 <= FS < $250,000	1,827	-0.011
FS >= $250,000	1,620	0.049

Table 7a: Stratified Correlation Analysis with Peak Number of Workers and Acreage Variables by Farm Value of Sales (FS)

Number of Records Correlation Coefficient	FS < $20,000	$20,000 <= FS < $100,000	$100,000 <= FS < $250,000	FS >= $250,000
Cash Grain	525 -0.101	1,657 0.077	1,539 0.013	1,570 0.000
Tobacco	64 0.114	133 0.281	36 0.596	16 0.344
Other Field Crops	544 0.147	1,184 0.055	943 0.081	1,031 0.368
Vegetable	6 -0.175	23 0.757	34 -0.081	47 0.169
Fruit	9 0.457	24 0.443	20 0.472	14 -0.183
Greenhouse	9 0.457	24 0.443	20 0.472	14 -0.183

Table 7b: Stratified Correlation Analysis with Peak Number of Workers and Number of Head Variables by Farm Value of Sales (FS)

Number of Records Correlation Coefficient	FS < $20,000	$20,000 <= FS < $100,000	$100,000 <= FS < $250,000	FS >= $250,000
Livestock	823 0.098	1,768 0.007	1,573 -0.008	1,555 0.104
Poultry	76 0.573	162 0.384	92 0.633	87 0.497

Table 8a: Stratified Correlation Analysis with Peak Number of Workers and Acreage Variables and Their Associated Farm Type by Farm Value of Sales (FS)

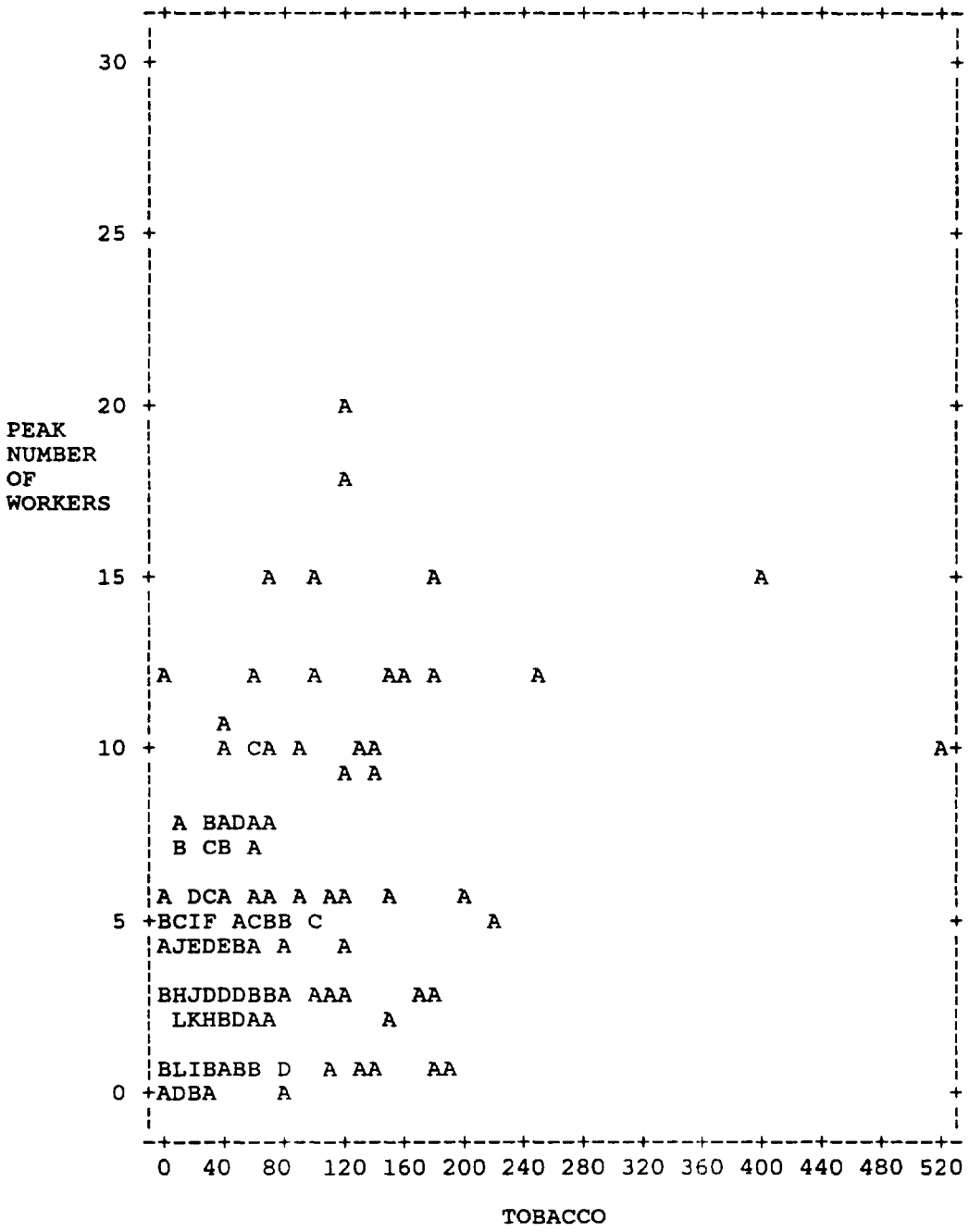
Number of Records Correlation Coefficient	FS < $20,000	$20,000 <= FS < $100,000	$100,000 <= FS < $250,000	FS >= $250,000
Cash Grain	320 0.019	1,072 0.032	1,022 -0.003	1,085 0.140
Tobacco	25 0.171	41 0.705	1 ****	1 ****
Other Field Crops	78 0.286	36 0.324	4 0.647	4 0.673
Vegetable	5 -0.310	10 0.956	9 -0.598	19 0.060
Fruit	7 0.459	14 0.340	12 0.541	5 -0.229
Greenhouse	9 0.457	22 0.449	20 0.472	14 -0.183

Table 8b: Stratified Correlation Analysis with Peak Number of Workers and Number of Head Variables and Their Associated Farm Type by Farm Value of Sales (FS)

Number of Records Correlation Coefficient	FS < $20,000	$20,000 <= FS < $100,000	$100,000 <= FS < $250,000	FS >= $250,000
Livestock	379 0.049	574 -0.079	503 -0.103	570 0.343
Poultry	4 -0.425	11 0.128	37 0.468	38 0.778

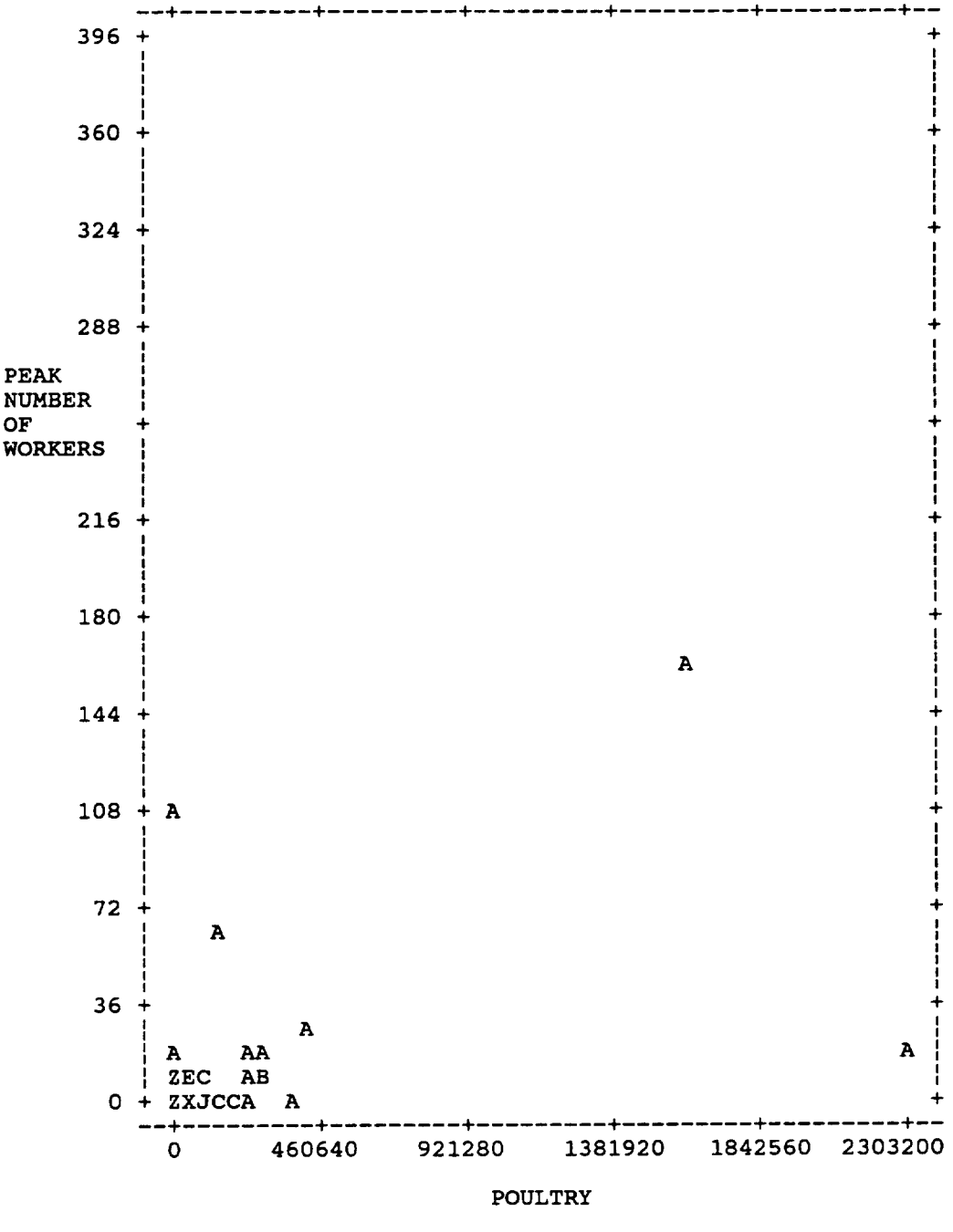
APPENDIX B: INDIANA SCATTERPLOT ANALYSES

Table 1: Scatterplot of Peak Number of Workers and Tobacco
Legend: A = 1 obs, B = 2 obs, etc.



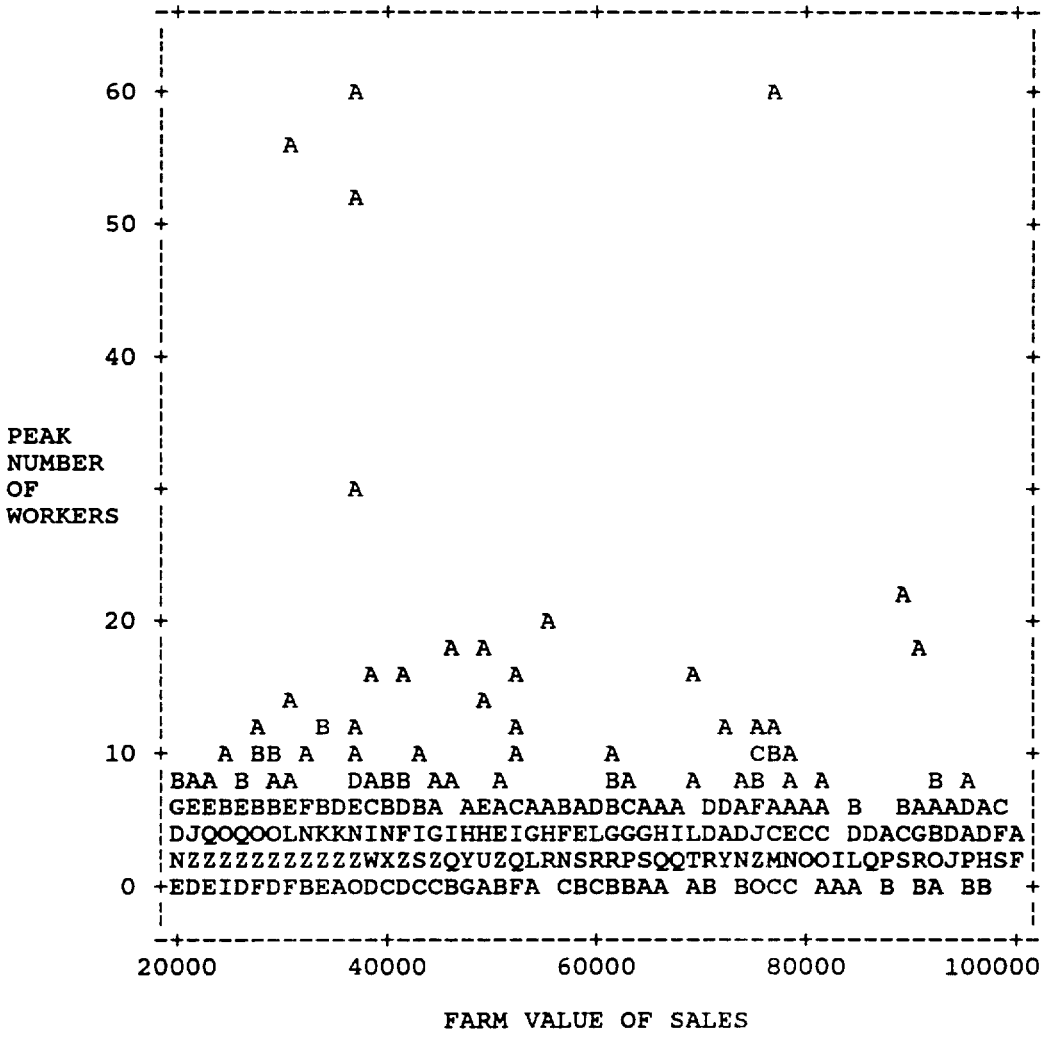
NOTE: 33 obs out of range

Table 2: Scatterplot of Peak Number of Workers and Poultry
Legend: A = 1 obs, B = 2 obs, etc.



NOTE: 304 obs hidden

Table 3: Scatterplot of Peak Number of Workers and Farm Value of Sales (FS):
$20,000 < = FS < $100,000
Legend: A = 1 obs, B = 2 obs, etc.



NOTE: 68 obs hidden.

APPENDIX C: INDIANA REGRESSION ANALYSES TABLES

Table 1: Regression Analysis for Dependent Variable Peak Number of Workers (Number of Head Variables not in Full Model)

R-Square (n = 5,749)	Variables in Model
0.043	Farmsale
0.050	Acres Sq_Sale
0.052	Acres Sq_Sale Rt_Acres
0.054	Acres Sq_Sale Lg_Sale Rt_Acres

Table 2: Regression Analysis for Dependent Variable Peak Number of Workers (Full Model)

R-Square (n = 402)	Variables in Model
0.535	Poultry
0.601	Livestock Poultry
0.618	Sq_Acres Livestock Poultry
0.647	Acres Rt_Sale Livestock Poultry

Table 3a: Regression Analysis for Dependent Variable Peak Number of Workers Stratified by Farm Value of Sales (Number of Head Variables not in the Full Model)

Farm Value of Sales (FS)	R-Square	Variables in Model
FS < $20,000 (n=737)	0.016	Lg_Sale
	0.017	Rt_Sale Lg_Sale
	0.017	Farmsale Sq_Sale Rt_Sale
	0.018	Acres Lg_Sale Sq_Acres Rt_Acres
$20,000 <= FS < $100,000 (n=1,724)	0.002	Sq_Acres
	0.002	Sq_Sale Sq_Acres
	0.003	Acres Rt_Acres Lg_Acres
	0.004	Acres Sq_Acres Rt_Acres Lg_Acres
$100,000 <= FS < $250,000 (n=1,556)	0.005	Lg_Acres
	0.010	Lg_Sale Rt_Acres
	0.011	Acres Lg_Sale Sq_Acres
	0.012	Acres Lg_Sale Rt_Acres Lg_Acres
FS >= $250,000 (n=1,732)	0.042	Farmsale
	0.048	Acres Sq_Sale
	0.050	Acres Sq_Sale Rt_Acres
	0.051	Acres Sq_Sale Sq_Acres Rt_Acres

Table 3b: Regression Analysis for Dependent Variable Peak Number of Workers Stratified by Farm Value of Sales (Full Model)

Farm Value of Sales (FS)	R-Square	Variables in Model
FS < $20,000 (n=71)	0.330	Poultry
	0.416	Rt_Sale Poultry
	0.428	Sq_Sale Lg_Acres Poultry
	0.450	Farmsale Sq_Sale Rt_Sale Poultry
$20,000 <= FS < $100,000 (n=158)	0.148	Poultry
	0.157	Lg_Acres Poultry
	0.202	Acres Sq_Acres Poultry
	0.218	Acres Sq_Acres Rt_Acres Poultry
$100,000 <= FS < $250,000 (n=89)	0.442	Poultry
	0.474	Lg_Sale Poultry
	0.479	Acres Lg_Sale Poultry
	0.491	Farmsale Rt_Sale Lg_Sale Poultry
FS >= $250,000 (n=84)	0.606	Livestock
	0.668	Poultry Livestock
	0.696	Acres Poultry Livestock
	0.706	Acres Lg_Sale Poultry Livestock

Table 4a: Regression Analysis for Dependent Variable Peak Number of Workers by Farm Type and Farm Value of Sales (FS): FS < $20,000

Farm Type	R-Square	Variables in Model
Cash Grains (n=320)	0.018	Acres
	0.019	Acres Sq_Sale
	0.020	Farmsale Lg_Acres Sq_Sale
	0.025	Acres Sq_Acres Rt_Acres Lg_Acres
Tobacco (n=25)	0.135	Lg_Sale
	0.140	Sq_Acres Lg_Sale
	0.181	Acres Sq_Acres Lg_Sale
	0.212	Farmsale Acres Sq_Acres Sq_Sale
Other Field Crops (n=78)	0.144	Farmsale
	0.153	Lg_Acres Rt_Sale
	0.169	Rt_Acres Lg_Acres Lg_Sale
	0.177	Rt_Acres Lg_Acres Rt_Sale Lg_Sale
Fruit (n=8)	0.480	Sq_Acres
	0.555	Lg_Acres Sq_Sale
	0.613	Sq_Acres Rt_Acres Sq_Sale
	0.628	Farmsale Rt_Acres Rt_Sale Lg_Sale
Livestock (n=297)	0.010	Lg_Acres
	0.013	Rt_Acres Sq_Sale
	0.018	Acres Rt_Sale Lg_Sale
	0.018	Farmsale Acres Sq_Sale Rt_Sale

Table 4b: Regression Analysis for Dependent Variable Peak Number of Workers by Farm Type and Farm Value of Sales (FS): $20,000 <= FS < $100,000

Farm Type	R-Square	Variables in Model
Cash Grains (n=1,072)	0.003	Sq_Acres
	0.003	Acres Sq_Acres
	0.004	Acres Rt_Acres Lg_Acres
	0.004	Farmsale Sq_Acres Rt_Sale Lg_Sale
Tobacco (n=42)	0.017	Sq_Sale
	0.127	Rt_Acres Sq_Sale
	0.180	Acres Sq_Acres Lg_Sale
	0.299	Acres Sq_Acres Lg_Acres Sq_Sale
Other Field Crops (n=36)	0.193	Sq_Acres
	0.200	Acres Lg_Sale
	0.215	Farmsale Acres Sq_Sale
	0.254	Acres Sq_Acres Rt_Acres Lg_Acres
Vegetables (n=12)	0.919	Sq_Acres
	0.936	Sq_Acres Lg_Acres
	0.941	Sq_Acres Rt_Acres Lg_Acres
	0.943	Farmsale Sq_Acres Rt_Acres Lg_Acres
Fruit (n=16)	0.484	Sq_Sale
	0.612	Sq_Acres Sq_Sale
	0.620	Farmsale Acres Sq_Acres
	0.625	Sq_Acres Lg_Acres Rt_Sale Lg_Sale
Livestock (n=538)	0.013	Lg_Acres
	0.037	Lg_Acres Sq_Sale
	0.043	Farmsale Lg_Acres Sq_Sale
	0.048	Farmsale Lg_Acres Rt_Sale Lg_Sale

Table 4c: Regression Analysis for Dependent Variable Peak Number of Workers by Farm Type and Farm Value of Sales (FS): $100,000 <= FS < $250,000

Farm Type	R-Square	Variables in Model
Cash Grains (n=1,022)	0.001	Acres
	0.002	Rt_acres Sq_Sale
	0.003	Acres Rt_Acres Lg_Acres
	0.007	Acres Sq_Acres Rt_Acres Lg_Acres
Vegetables (n=11)	0.047	Sq_Sale
	0.093	Acres Sq_Acres
	0.177	Acres Rt_Acres Lg_Acres
	0.188	Acres Rt_Acres Lg_Acres Sq_Sale
Fruit (n=12)	0.149	Rt_Sale
	0.261	Sq_Acres Sq_Sale
	0.369	Farmsale Sq_Sale Rt_Sale
	0.516	Farmsale Sq_Acres Sq_Sale Rt_Sale
Livestock (n=495)	0.041	Sq_Acres
	0.048	Acres Sq_Acres
	0.054	Sq_Acres Lg_Acres Lg_Sale
	0.061	Acres Sq_Acres Rt_Acres Lg_Acres
Poultry (n=9)	0.179	Sq_Acres
	0.333	Rt_Acres Lg_Acres
	0.466	Acres Sq_Acres Rt_Acres
	0.755	Farmsale Lg_Acres Sq_Sale Rt_Sale

Table 4d: Regression Analysis for Dependent Variable Peakwork by Farm Type and Farm Value of Sales (FS): FS >= $250,000

Farm Type	R-Square	Variables in Model
	0.029	Rt_Acres
Cash Grains	0.038	Acres Sq_Sale
(n=1,085)	0.041	Rt_Acres Lg_Acres Sq_Sale
	0.043	Acres Rt_Acres Sq_Sale Rt_Sale
	0.017	Lg_Acres
Vegetables	0.074	Farmsale Rt_Sale
(n=30)	0.097	Acres Sq_Acres Lg_Sale
	0.100	Sq_Acres Rt_Acres Lg_Acres Lg_Sale
	0.026	Lg_Sale
Greenhouse	0.093	Rt_Sale Lg_Sale
(n=12)	0.194	Acres Rt_Acres Lg_Acres
	0.775	Acres Sq_Acres Rt_Acres Lg_Acres
	0.440	Sq_Acres
Livestock	0.504	Acres Sq_Acres
(n=554)	0.528	Acres Sq_Acres Rt_Acres
	0.549	Farmsale Sq_Acres Sq_Sale Rt_Sale
	0.128	Sq_Sale
Poultry	0.209	Lg_Acres Sq_Sale
(n=37)	0.261	Lg_Acres Rt_Sale Lg_Sale
	0.303	Farmsale Acres Sq_Sale Rt_Sale

APPENDIX D: FLORIDA CORRELATION ANALYSES TABLES

Table 1: Correlation Analysis with Peak Number of Workers and Selected Control Variables

Selected Control Variables	Sample Size	Correlation Coefficient
Vegetable Land	1,131	0.335
Cropland	11,715	0.409
All Land	9,383	0.185
Tobacco	199	0.350
Corn	966	0.197
Soybeans	250	0.189
Wheat	158	0.103
Hogs	7,811	0.001
All Cattle	8,142	0.062
Dairy Cows	420	0.876
Chickens	258	0.604

Table 2: Stratified Correlation Analysis with Peak Number of Workers and Selected Control Variables (SCV) by Farm Value of Sales (FS)

Number of Records Correlation Coefficients	FS < $20,000	$20,000 <= FS < $100,000	$100,000 <= FS < $250,000	FS >= $250,000
SCV= Vegetable Land	106 -0.104	289 0.231	275 0.269	461 0.298
SCV= Cropland	3,925 -0.065	2,702 -0.014	2,942 -0.027	2,146 0.414
SCV= All Land	3,644 -0.037	2,428 0.032	1,503 -0.010	1,808 0.169
SCV= Tobacco	19 -0.188	55 0.180	60 0.171	65 0.067
SCV= Corn	249 -0.002	327 0.148	194 0.032	196 0.105
SCV= Soybeans	23 0.281	70 0.003	92 -0.112	65 0.129
SCV= Wheat	11 -0.243	45 -0.006	62 -0.196	40 -0.009
SCV= Hogs	3,452 0.003	2,039 -0.017	1,076 -0.025	1,244 -0.007
SCV= All Cattle	3,548 -0.037	2,137 -0.067	1,149 -0.044	1,308 0.022
SCV= Dairy Cows	62 0.106	63 0.131	57 -0.219	238 0.867
SCV= Chickens	51 0.967	43 0.735	96 0.470	68 0.312
SCV= Aquaculture	26 -0.257	13 0.011	36 0.115	8 0.361
SCV= Floriculture	72 -0.137	439 0.270	577 -0.261	418 0.449

Table 3a: Stratified Correlation Analysis with Peak Number of Workers and Selected Control Variables (SCV) and Farm Type (FT) by Farm Value of Sales (FS)

Number of Records Correlation Coefficient	FS < $20,000	$20,000 <= FS < $100,000	$100,000 <= FS < $250,000	FS >= $250,000
FT= Cash Grain SCV= Cropland	53 0.201	54 0.187	25 0.251	6 -0.022
FT= Cash Grain SCV= Corn	40 0.403	43 0.218	22 0.177	6 0.759
FT= Cash Grain SCV= Soybeans	13 0.366	26 -0.029	19 -0.253	4 0.311
FT= Cash Grain SCV= Wheat	2 ****	10 -0.026	10 -0.134	3 0.945
FT= Tobacco SCV= Tobacco	14 -0.102	34 0.320	38 0.058	27 0.050
FT= Vegetable SCV= Vegetable Land	62 0.124	179 0.217	171 0.155	341 0.293
FT= Livestock SCV= Hogs	2,568 0.034	1,247 -0.002	358 -0.004	253 0.015
FT= Livestock SCV= All Cattle	2,607 0.138	1,278 0.059	370 0.091	263 0.245
FT= Livestock SCV=Dairy	13 -0.022	8 0.290	34 -0.215	221 0.867
FT= Poultry SCV= Chickens	1 ****	4 0.544	82 0.100	41 0.776
FT= Other Livestock SCV= Aquaculture	1 ****	2 ****	31 -0.101	3 0.882
FT= Greenhouse SCV= Floriculture	72 -0.137	433 0.272	575 -0.277	412 0.475

Table 3b: Stratified Correlation Analysis with Peak Number of Workers and All Land Control Variable and Farm Type (FT) by Farm Value of Sales (FS)

Number of Records Correlation Coefficient	FS < $20,000	$20,000 <= FS < $100,000	$100,000 <= FS < $250,000	FS >= $250,000
FT= Cash Grain All Land	53 0.425	54 0.385	25 -0.069	6 0.208
FT= Tobacco All Land	14 0.568	34 -0.183	38 0.087	27 -0.276
FT= Cotton All Land	1 ****	18 -0.300	33 0.197	24 0.293
FT= Other Field Crops All Land	196 -0.026	171 -0.004	87 -0.054	102 0.499
FT= Vegetables All Land	60 -0.057	175 0.065	164 -0.095	322 0.278
FT= Fruit All Land	201 0.226	461 0.213	251 0.068	516 0.115
FT= Nuts All Land	360 0.086	116 0.076	17 -0.086	9 0.355
FT= Greenhouse All Land	11 0.920	81 -0.010	392 0.423	215 0.025
FT= Livestock All Land	2,605 0.094	1,278 0.029	370 -0.023	263 0.176
FT= Poultry All Land	1 ****	4 -0.019	88 0.033	103 0.186
FT= Dairy All Land	13 -0.055	8 0.234	34 0.526	220 0.695
FT= Other Livestock All Land	128 0.451	28 0.112	4 0.742	1 ****

Table 4a: Correlation Analysis with Peak Number of Workers and Acreage Variables

Acreage Variables	Number of Records	Correlation Coefficient
Cash Grain	1,293	0.617
Tobacco	199	0.350
Other Field Crops	3,262	0.461
Vegetables	942	0.474
Fruit	2,158	0.449
Cotton	173	0.067
Sod Farm	81	0.499
Nuts	951	0.092

Table 4b: Correlation Analysis with Peak Number of Workers and Acreage Variables by Farm Type

Farm Acres	Number of Records	Correlation Coefficient
Cash Grain	134	0.247
Tobacco	113	0.357
Other Field Crops	555	0.476
Vegetables	621	0.475
Fruit	1,437	0.543
Cotton	74	0.327
Sod Farm	66	0.568
Nuts	504	0.317

Table 5a: Correlation Analysis with Peak Number of Workers and Number of Head Variables

Number of Head	Number of Records	Correlation Coefficient
Livestock	8,269	0.055
Poultry	334	0.569
Horses	2,251	0.093
Bees	108	0.330

Table 5b: Correlation Analysis with Peak Number of Workers and Number of Head Variables by Farm Type

Number of Head	Number of Records	Correlation Coefficient
Livestock	4,518	0.245
Poultry	199	0.359
Horses	139	0.718
Bees	18	0.972

Table 6: Correlation Analysis with Peak Number of Workers and Farm Value of Sales (FS)

Farm Value of Sales	Number of Records	Correlation Coefficient
FS < $20,000	3,718	0.031
$20,000 <= FS < $100,000	2,805	0.095
$100,000 <= FS < $250,000	3,044	0.047
FS >= $250,000	2,148	0.333

Table 7a: Stratified Correlation Analysis With Peak Number of Workers and Acreage Variables by Farm Value of Sales (FS)

Number of Records Correlation Coefficient	FS < $20,000	$20,000 <= FS < $100,000	$100,000 <= FS < $250,000	FS >= $250,000
Cash Grain	329 0.060	438 0.093	267 -0.067	259 0.639
Tobacco	19 -0.188	55 0.180	60 0.171	65 0.229
Cotton	3 0.000	42 0.093	68 -0.101	60 -0.104
Other Field Crops	909 0.051	1,088 0.009	553 0.111	712 0.456
Vegetables	71 0.669	232 0.156	252 0.121	387 0.425
Fruit	314 0.068	636 0.089	435 0.127	773 0.447
Nuts	491 0.002	273 -0.016	103 0.031	84 0.119
Sod Farm	0 ****	0 ****	41 0.911	40 0.262

Table 7b: Stratified Correlation Analysis With Peak Number of Workers and Number of Head Variables by Farm Value of Sales (FS)

Number of Records Correlation Coefficient	FS < $20,000	$20,000 <= FS < $100,000	$100,000 <= FS < $250,000	FS >= $250,000
Livestock	3,550 -0.032	2,156 -0.078	1,174 -0.0687	1,389 -0.012
Poultry	52 0.736	43 0.735	113 0.432	126 0.253
Horses	871 0.275	672 0.581	326 0.198	382 0.034
Bees	47 0.106	28 -0.018	23 -0.255	10 0.262

Table 8a: Stratified Correlation Analysis with Peak Number of Workers and Acreage Variables and Their Associated Farm Type (FT) by Farm Value of Sales (FS)

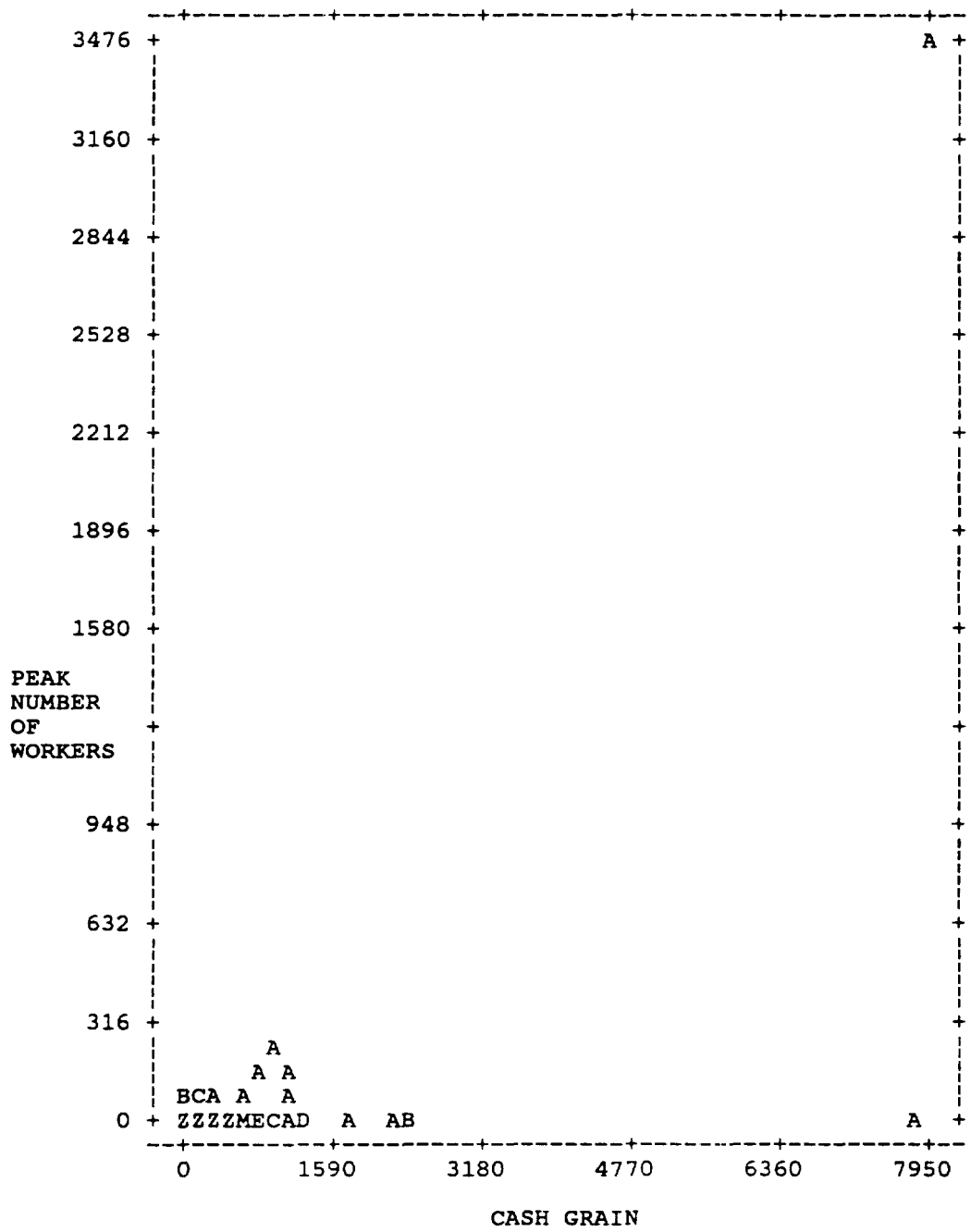
Number of Records Correlation Coefficient	FS < $20,000	$20,000 <= FS < $100,000	$100,000 <= FS < $250,000	FS >= $250,000
Cash Grain FT = Cash Grain	53 0.253	51 0.137	24 0.336	6 0.660
Tobacco FT = Tobacco	14 -0.102	34 0.320	38 0.058	27 0.219
Cotton FT = Cotton	1 ****	17 -0.214	32 0.247	24 0.166
Other Field Crops FT = Other Field Crops	196 0.006	170 -0.017	87 0.074	102 0.462
Vegetables FT = Vegetables	42 0.745	152 0.108	152 -0.019	275 0.411
Fruit FT = Fruit	195 0.089	462 0.042	257 0.151	523 0.573
Sod Farm FT = Greenhouse	0 ****	0 ****	41 0.911	25 0.488
Nuts FT = Nuts	361 0.032	116 0.193	17 0.534	10 0.507

Table 8b: Stratified Correlation Analysis with Peak Number of Workers and Number of Head Variables and Their Associated Farm Type (FT) by Farm Value of Sales (FS)

Number of Records Correlation Coefficient	FS < $20,000	$20,000 <= FS < $100,000	$100,000 <= FS < $250,000	FS >= $250,000
Livestock FT= Livestock	2,607 0.132	1,278 0.053	370 0.080	263 0.239
Poultry FT= Poultry	1 ****	4 0.544	98 0.054	96 0.611
Horses FT= Other Livestock	114 0.467	23 0.712	2 1.00	0 ****
Bees FT= Other Livestock	13 0.577	3 0.189	1 ****	1 ****

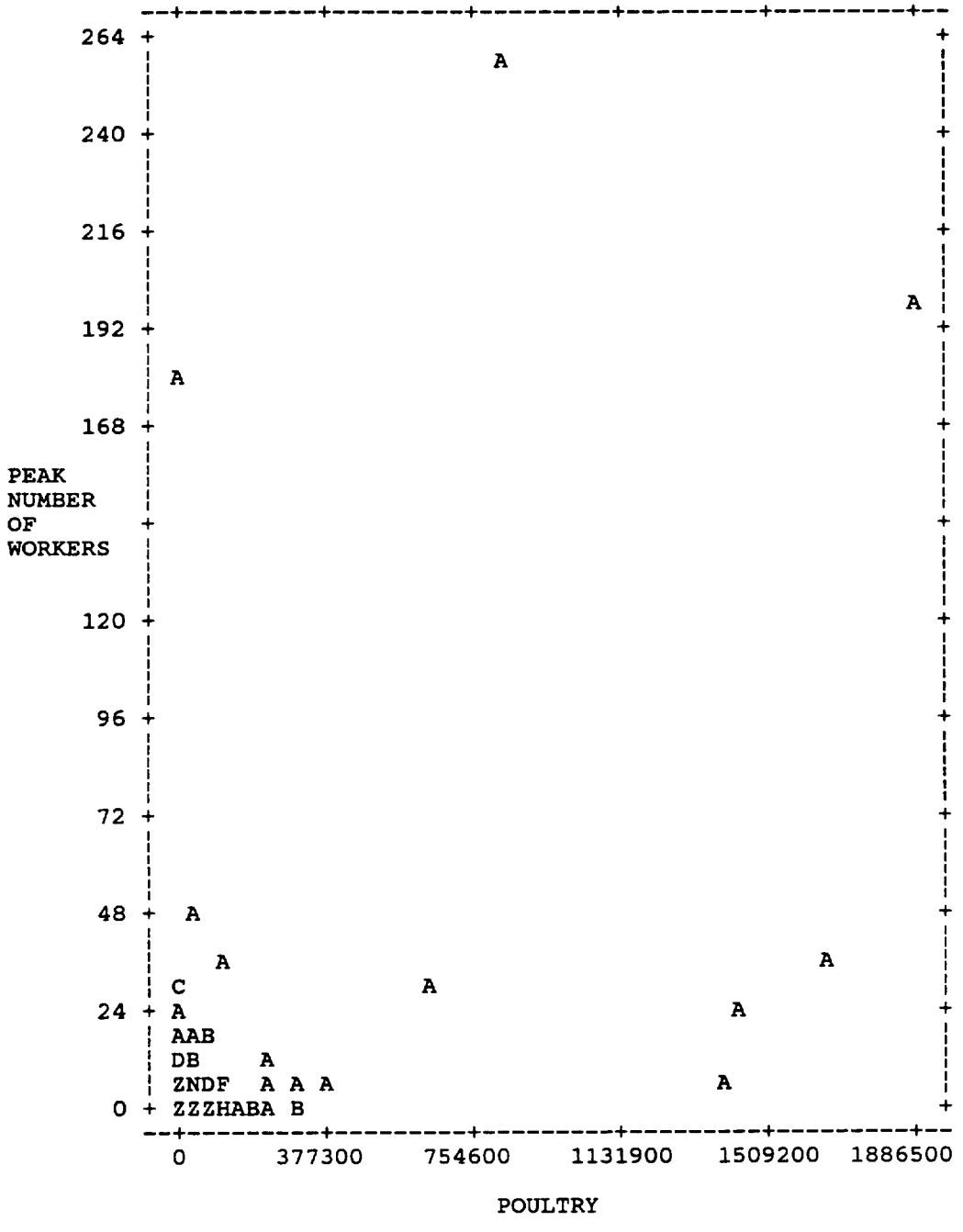
APPENDIX E: FLORIDA SCATTERPLOT ANALYSES

Table 1: Scatterplot of Peak Number of Workers and Cash Grain
Legend: A = 1 obs, B = 2 obs, etc.



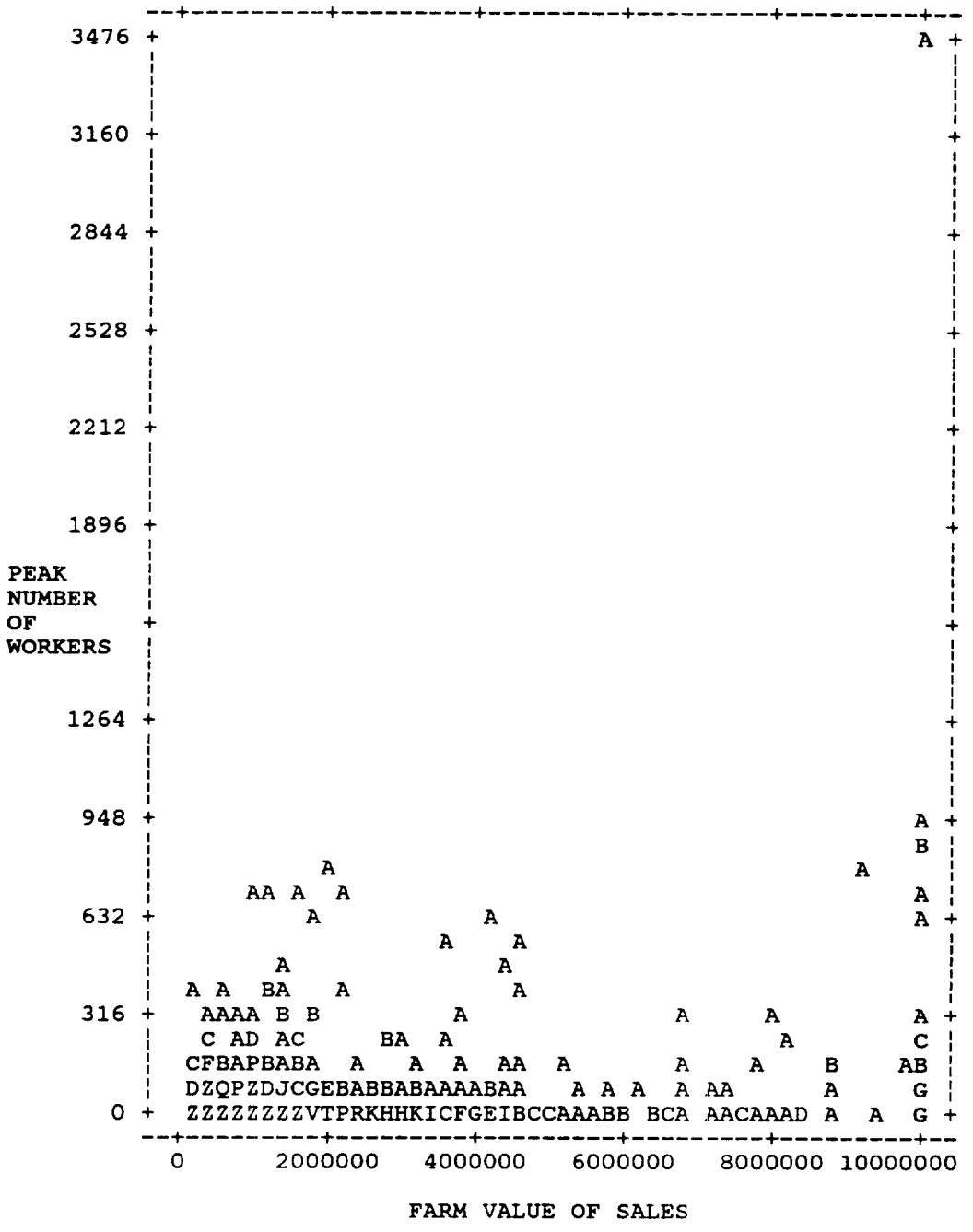
NOTE: 1146 obs hidden

Table 2: Scatterplot of Peak Number of Workers and Poultry
Legend: A = 1 obs, B = 2 obs, etc.



NOTE: 169 obs. hidden and 43 obs. out of range

Table 3: Scatterplot of Peak Number of Workers and Farm Value of Sales (FS):
FS >= $250,000
Legend: A = 1 obs, B = 2 obs, etc.



NOTE: 1491 obs hidden

APPENDIX F: FLORIDA REGRESSION ANALYSES TABLES

Table 1: Regression Analysis for Dependent Variable Peak Number of Workers (Number of Head Variables not in Full Model)

R-Square (n = 6069)	Variables in Model
0.205	Acres
0.291	Sq_Acres Acres
0.299	Rt_Acres Sq_Acres Acres
0.314	Ln_Acres Rt_Acres Sq_Acres Acres

Table 2: Regression Analysis for Dependent Variable Peak Number of Workers (Full Model)

R-Square (n = 232)	Variables in Model
0.418	Poultry
0.455	Farmsale Poultry
0.490	Acres Farmsale Poultry
0.495	Sq_Acres Ln_Acres Farmsale Poultry

Table 3a: Regression Analysis for Dependent Variable Peak Number of Workers Stratified by Farm Value of Sales (Number of Head Variables not in the Model)

Farm Value of Sales (FS)	R-Square	Variables in Model
FS < $20,000 (n=1,647)	0.004	Sq_Sale
	0.006	Lg_Acres Sq_Sale
	0.008	Sq_Acres Sq_Sale Acres
	0.008	Farmsale Sq_Acres Sq_Sale Acres
$20,000 <= FS < $100,000 (n=1,831)	0.010	Sq_Sale
	0.017	Farmsale Rt_Acres
	0.018	Farmsale Sq_Acres Rt_Acres
	0.018	Farmsale Rt_Acres Lg_Acres Acres
$100,000 <= FS < $250,000 (n=1,057)	0.004	Acres
	0.008	Rt_Sale Acres
	0.012	Farmsale Rt_Acres Lg_Acres
	0.014	Farmsale Sq_Sale Rt_Acres Acres
FS >= $250,000 (n=1,534)	0.198	Acres
	0.284	Sq_Acres Acres
	0.303	Sq_Acres Rt_Acres Acres
	0.323	Sq_Acres Rt_Acres Lg_Acres Acres

Table 3b: Regression Analysis for Dependent Variable Peak Number of Workers Stratified by Farm Value of Sales (Full Model)

Farm Value of Sales (FS)	R-Square	Variables in Model
FS < $20,000 (n=32)	0.023	Sq_Acres
	0.060	Lg_Sale Acres
	0.134	Sq_Acres Rt_Sale Livestock
	0.152	Farmsale Sq_Acres Poultry Livestock
$20,000 <= FS < $100,000 (n=35)	0.538	Poultry
	0.563	Sq_Sale Poultry
	0.587	Sq_Sale Acres Poultry
	0.622	Sq_Sale Acres Poultry Livestock
$100,000 <= FS < $250,000 (n=61)	0.395	Poultry
	0.517	Poultry Livestock
	0.551	Rt_Acres Poultry Livestock
	0.559	Rt_Acres Sq_Sale Poultry Livestock
FS >= $250,000 (n=104)	0.157	Lg_Acres
	0.208	Sq_Acres Rt_Acres
	0.219	Rt_Acres Acres Poultry
	0.249	Farmsale Sq_Acres Rt_Acres Poultry

Table 4(a1): Regression Analysis for Dependent Variable Peak Number of Workers by Farm Type (Acres Only) and Farm Value of Sales (FS): FS < $20,000

Farm Type	R-Square	Variables in Model
Cash Grains (n=53)	0.052	Acres
	0.067	Sq_Sale Acres
	0.070	Rt_Acres Lg_Acres Sq_Sale
	0.075	Farmsale Rt_Acres Sq_Sale Rt_Sale
Tobacco (n=14)	0.014	Acres
	0.038	Rt_Acres Lg_Acres
	0.277	Acres Rt_Acres Lg_Acres
	0.319	Sq_Sale Acres Rt_Acres Lg_Acres
Other Field Crops (n=196)	0.003	Lg_Sale
	0.007	Sq_Acres Lg_Sale
	0.009	Farmsale Acres Rt_Sale
	0.010	Farmsale Sq_Sale Rt_Sale Lg_Sale
Vegetable (n=45)	0.554	Sq_Acres
	0.649	Sq_Acres Lg_Acres
	0.654	Sq_Acres Lg_Acres Lg_Sale
	0.678	Farmsale Sq_Acres Rt_Sale Lg_Sale
Fruit (n=168)	0.003	Rt_Acres
	0.012	Farmsale Sq_Sale
	0.017	Farmsale Sq_Sale Rt_Sale
	0.040	Farmsale Sq_Sale Rt_Sale Lg_Sale
Nuts (n=361)	0.016	Lg_Sale
	0.023	Farmsale Sq_Sale
	0.027	Farmsale Acres Sq_Sale
	0.030	Acres Sq_Acres Rt_Acres Rt_Sale

Table 4(a2): Regression Analysis for Dependent Variable Peak Number of Workers by Farm Type (Number of Head Only) and Farm Value of Sales (FS): FS < $20,000

Farm Type	R-Square	Variables in Model
Livestock (n=764)	0.004	Lg_Sale
	0.005	Sq_Acres Lg_Sale
	0.005	Sq_Acres Lg_Acres Lg_Sale
	0.006	Acres Sq_Acres Rt_Acres Lg_Sale
Other Livestock (n=13)	0.259	Lg_Acres
	0.747	Lg_Acres Lg_Sale
	0.791	Lg_Acres Rt_Sale Lg_Sale
	0.908	Acres Lg_Acres Rt_Sale Lg_Sale

Table 4(b1): Regression Analysis for Dependent Variable Peak Number of Workers by Farm Type (Acres Only) and Farm Value of Sales (FS): $20,000 < = FS < $100,000

Farm Type	R-Square	Variables in Model
	0.058	Sq_Sale
Cash Grains	0.061	Sq_Acres Sq_Sale
(n=54)	0.070	Acres Rt_Acres Rt_Sale
	0.101	Acres Sq_Acres Rt_Acres Lg_Acres
	0.189	Lg_Sale
Tobacco	0.227	Sq_Acres Rt_Sale
(n=34)	0.229	Farmsale Sq_Acres Sq_Sale
	0.240	Acres Rt_Acres Lg_Acres Lg_Sale
	0.042	Sq_Acres
Cotton	0.086	Rt_Acres Lg_Acres
(n=18)	0.131	Acres Rt_Acres Lg_Acres
	0.185	Acres Rt_Acres Lg_Acres Lg_Sale
	0.018	Lg_Acres
Other Field Crops	0.031	Rt_Acres Lg_Acres
(n=171)	0.038	Acres Rt_Acres Lg_Acres
	0.049	Acres Sq_Acres Rt_Acres Lg_Acres
	0.054	Lg_Sale
Vegetable	0.109	Rt_Acres Lg_Sale
(n=173)	0.118	Farmsale Rt_Acres Sq_Sale
	0.133	Farmsale Rt_Acres Sq_Sale Rt_Sale
	0.003	Sq_Sale
Fruit	0.007	Lg_Acres Sq_Sale
(n=461)	0.008	Sq_Acres Rt_Acres Sq_Sale
	0.008	Farmsale Acres Sq_Acres Sq_Sale
	0.014	Lg_Sale
Nuts	0.017	Farmsale Sq_Sale
(n=116)	0.023	Farmsale Rt_Sale Lg_Sale
	0.027	Farmsale Acres Rt_Sale Lg_Sale
	0.009	Lg_Sale
Greenhouse	0.091	Acres Sq_Acres
(n=11)	0.579	Acres Sq_Acres Lg_Acres
	0.624	Acres Sq_Acres Rt_Acres Lg_Sale

Table 4(b2): Regression Analysis for Dependent Variable Peak Number of Workers by Farm Type (Number of Head) and Farm Value of Sales (FS): $20,000 <= FS < $100,000

Farm Type	R-Square	Variables in Model
Livestock (n=783)	0.005	Lg_Sale
	0.005	Sq_Acres Lg_Sale
	0.007	Farmsale Sq_Sale Rt_Sale
	0.008	Farmsale Sq_Acres Sq_Sale Rt_Sale

Table 4(c1): Regression Analysis for Dependent Variable Peak Number of Workers by Farm Type (Acres Only) and Farm Value of Sales (FS): $100,000 < = FS < $250,000

Farm Type	R-Square	Variables in Model
	0.155	Sq_Sale
Cash Grains	0.253	Farmsale Sq_Sale
(n=25)	0.263	Farmsale Sq_Sale Rt_Sale
	0.411	Farmsale Acres Sq_Acres Sq_Sale
	0.016	Sq_Acres
Tobacco	0.045	Sq_Acres Sq_Sale
(n=38)	0.135	Acres Sq_Acres Rt_Acres
	0.205	Acres Sq_Acres Rt_Acres Lg_Acres
	0.030	Sq_Acres
Cotton	0.037	Sq_Acres Lg_Sale
(n=33)	0.045	Acres Sq_Acres Rt_Acres
	0.058	Acres Sq_Acres Rt_Acres Lg_Acres
	0.104	Lg_Acres
Other Field Crops	0.171	Rt_Acres Lg_Acres
(n=87)	0.235	Acres Rt_Acres Lg_Acres
	0.265	Acres Sq_Acres Rt_Acres Lg_Acres
	0.041	Rt_Acres
Vegetable	0.050	Rt_Acres Lg_Sale
(n=164)	0.054	Sq_Acres Rt_Acres Lg_Sale
	0.077	Farmsale Rt_Acres Sq_Sale Rt_Sale
	0.014	Farmsale
Fruit	0.015	Acres Sq_Acres
(n=257)	0.019	Farmsale Sq_Sale Rt_Sale
	0.019	Farmsale Sq_Acres Sq_Sale Rt_Sale
	0.173	Rt_Acres
Nuts	0.210	Acres Sq_Acres
(n=17)	0.393	Acres Sq_Acres Rt_Acres
	0.709	Acres Sq_Acres Rt_Acres Lg_Acres
	0.005	Lg_Acres
Greenhouse	0.008	Rt_Acres Lg_Acres
(n=116)	0.018	Acres Rt_Acres Lg_Acres
	0.019	Acres Sq_Acres Rt_Acres Lg_Acres

Table 4(c2): Regression Analysis for Dependent Variable Peak Number of Workers by Farm Type (Number of Head Only) and Farm Value of Sales (FS): $100,000 < = FS < $250,000

Farm Type	R-Square	Variables in Model
Livestock (n=244)	0.040	Sq_Sale
	0.054	Rt_Acres Sq_Sale
	0.076	Farmsale Sq_Sale Rt_Sale
	0.090	Farmsale Rt_Acres Sq_Sale Rt_Sale
Poultry (n=51)	0.031	Lg_Acres
	0.129	Farmsale Rt_Sale
	0.138	Sq_Acres Sq_Sale Rt_Sale
	0.139	Farmsale Sq_Acres Rt_Sale Lg_Sale
Dairy (n=23)	0.229	Rt_Acres
	0.246	Acres Sq_Acres
	0.285	Farmsale Rt_Acres Sq_Sale
	0.297	Farmsale Acres Sq_Acres Sq_Sale

Table 4(d1): Regression Analysis for Dependent Variable Peak Number of Workers by Farm Type (Acres Only) and Farm Value of Sales (FS): FS > = $250,000

Farm Type	R-Square	Variables in Model
	0.010	Lg_Acres
Tobacco	0.093	Rt_Acres Lg_Acres
(n=27)	0.172	Acres Sq_Acres Rt_Acres
	0.248	Acres Sq_Acres Rt_Acres Lg_Acres
	0.013	Sq_Sale
Cotton	0.058	Acres Sq_Acres
(n=24)	0.067	Farmsale Sq_Sale Rt_Sale
	0.109	Farmsale Sq_Sale Rt_Sale Lg_Sale
	0.323	Rt_Acres
Other Field Crops	0.606	Acres Sq_Acres
(n=102)	0.769	Acres Sq_Acres Sq_Sale
	0.816	Acres Sq_Acres Rt_Acres Sq_Sale
	0.138	Farmsale
Vegetable	0.143	Sq_Sale Lg_Sale
(n=335)	0.146	Rt_Acres Lg_Acres Sq_Sale
	0.151	Acres Sq_Acres Rt_Acres Sq_Sale
	0.353	Sq_Acres
Fruit	0.365	Sq_Acres Lg_Sale
(n=523)	0.370	Acres Sq_Acres Lg_Sale
	0.371	Sq_Acres Rt_Acres Lg_Acres Rt_Sale
	0.292	Farmsale
Nuts	0.385	Sq_Acres Sq_Sale
(n=10)	0.588	Farmsale Rt_Sale Lg_Sale
	0.642	Farmsale Lg_Acres Rt_Sale Lg_Sale
	0.347	Sq_Sale
Greenhouse	0.355	Sq_Acres Sq_Sale
(n=101)	0.372	Acres Sq_Acres Sq_Sale
	0.379	Acres Rt_Acres Lg_Acres Sq_Sale

Table 4(d2): Regression Analysis for Dependent Variable Peak Number of Workers by Farm Type (Number of Head Only) and Farm Value of Sales (FS): FS > = $250,000

Farm Type	R-Square	Variables in Model
Livestock (n=175)	0.061	Lg_Sale
	0.065	Sq_Acres Lg_Sale
	0.070	Farmsale Sq_Sale Rt_Sale
	0.072	Acres Sq_Acres Lg_Acres Lg_Sale
Poultry (n=83)	0.172	Rt_Acres
	0.248	Acres Sq_Acres
	0.258	Sq_Acres Rt_Acres Lg_Acres
	0.259	Acres Sq_Acres Rt_Acres Lg_Acres
Dairy (n=148)	0.756	Sq_Sale
	0.772	Sq_Acres Sq_Sale
	0.804	Farmsale Sq_Sale Rt_Sale
	0.836	Farmsale Sq_Sale Rt_Sale Lg_Sale