University of Arkansas, Fayetteville

ScholarWorks@UARK

Graduate Theses and Dissertations | Graduate School and International Education

12-2015

Evaluating the Impacts of Respondent Errors in ARMS: A Case of Farm Service Agency Loans

Louise Nicole McMinn

University of Arkansas, Fayetteville

Follow this and additional works at: https://scholarworks.uark.edu/etd

Part of the Agricultural Economics Commons, Finance Commons, and the Finance and Financial Management Commons

Click here to let us know how this document benefits you.

Citation

McMinn, L. N. (2015). Evaluating the Impacts of Respondent Errors in ARMS: A Case of Farm Service Agency Loans. *Graduate Theses and Dissertations* Retrieved from https://scholarworks.uark.edu/etd/1339

This Thesis is brought to you for free and open access by the Graduate School and International Education at ScholarWorks@UARK. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of ScholarWorks@UARK. For more information, please contact uarepos@uark.edu.

Evaluating the Impacts of Respondent Errors in ARMS: A Case of Farm Service Agency Loans

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Agricultural Economics

By

Louise Nicole McMinn
Arkansas Tech University
Bachelor of Science in Agricultural Business, 2013

December 2015
University of Arkansas

This thesis is approved for recommendation to the Graduate Council.

Dr. Bruce L. Ahrendsen
Thesis Director

Dr. Bruce L. Dixon
Committee Member

Dr. Charles B. Dodson
Committee Member

Abstract

Many studies have used the U.S. Department of Agriculture's (USDA) Agricultural Resource Management Survey (ARMS) to research various aspects involving the agricultural sector in the United States. Since nonresponse and inaccurate reporting may cause significant bias in statistical analysis, research was conducted to determine the magnitude of response error on the farm debt section of the ARMS Phase III. A multinomial logit model identified demographic, structural, and financial characteristics of FSA Farm Loan Program (FLP) borrowers who refused to indicate if they had end of year farm debt, or who accurately or inaccurately classified their farm operations as having end of year farm debt on the ARMS for 2001, 2004, 2006, and 2007. Additionally, estimates of the magnitude of response errors in ARMS for both FSA direct and guaranteed FLPs were estimated. The current study found that 12.9 percent of the direct FLP respondents and 9.9% of the guaranteed FLP respondents indicated "no" on the "Owe Money" question when they should have indicated "yes". Also, those responding "no" were found to have their ARMS total debt outstanding less than their FSA total debt outstanding. Direct FLP operators were more likely to report "no" and, therefore, under-report end of year debt in the ARMS if they had a lower total FSA debt outstanding balance, had a greater value of crop production relative to total production, or had a lower gross cash farm income. Guaranteed FLP operators were more likely to under-report their debt in the ARMS if they had an operating line of credit loan, had a greater share of production from crops, had a lower gross cash farm income, were in survey year 2004, or were in survey year 2007. They were less likely to under-report their debt if they either had some college education, were socially disadvantaged eligible, or were beginning farmer eligible. These results allow future researchers using ARMS data to appraise operator debt status to be better informed about potential data limitations.

Acknowledgements

I would like to take this opportunity to express my gratitude to everyone that supported me while I was working toward completing my thesis. I am grateful for the superb guidance, constructive criticism, and valuable advice that I was given throughout the process.

First, I would like to thank my advisor Dr. Bruce Ahrendsen for his time, guidance, and advice throughout the thesis process. Next, I would like to thank Diana Danforth for the computer programming support and advice for my thesis. I would also like to thank Dr. Bruce Dixon for his valuable insight and advice. Lastly, I would like to thank Dr. Charles Dodson for giving me access to the FSA loan program data, and for always being available to answer any questions that arose.

I am fortunate to be a part of the Department of Agricultural Economics and Agribusiness at the University of Arkansas, Fayetteville. My two year Master's study was an enjoyable learning experience made possible by all of the faculty, staff, and students.

This work was supported, in part, by the USDA National Institute of Food and Agriculture, Hatch/Multi State project 220091. However, any opinions, findings, conclusions, or recommendations expressed in this thesis are those of the author and do not necessarily reflect the view of the U.S. Department of Agriculture.

Dedication

I dedicate my thesis to my husband, David Stout, who always and continues to be my greatest supporter and biggest fan. He motivates me to work my hardest, and strive to achieve greatness.

I also dedicate this thesis to my parents who have always supported me in all of my endeavors.

Table of Contents

List of Tables

Chapter 1: Introduction

1.1 Introduction

Many studies use the United States Department of Agriculture (USDA)'s Agricultural Resource Management Survey (ARMS) to research various aspects, such as policy impact, financial performance, and economic impact, involving the production agricultural sector in the United States. The ARMS is a comprehensive survey which gathers data on crops, input usage, farmer demographics, and finances, and is a very useful tool for researchers. The ARMS can also be used by those in government to determine the effectiveness of policy and regulation, and to assess effectiveness and funding levels of farm programs. One of the farm programs that uses information provided by ARMS is the USDA's Farm Service Agency's (FSA) Farm Loan Program. ARMS data may be used to determine the impact of FSA in meeting future farmer loan demand and funding levels needed to meet program demand.

1.2 Purpose of Study

Because the ARMS plays a critical role in research and policy, the accuracy of the ARMS data is also of fundamental importance. Many studies have stated the importance of the ARMS to researchers in academia and government who analyze U.S. farm and conservation policy as well as the effects of macroeconomic and other factors on the U.S. farm sector (Blank and Klinefelter, 2012; Featherstone, Park, and Weber, 2012; Weber and Clay, 2013). One of the ways that the ARMS data accuracy may be compromised is due to the respondents themselves. Farm operators may not want to truthfully answer debt questions or operators may not take the time to accurately respond to the debt questions. Nonresponse and inaccurate reporting in the ARMS can cause biased estimates that do not accurately reflect farm financial health or the effectiveness and demand of USDA credit programs. Since nonresponse and inaccurate reporting can cause significant bias in statistical analysis, research needs to be conducted to determine the magnitude of response error on the ARMS.

The analysis in this study is undertaken using matched FSA and ARMS data. ARMS and FSA loan data were matched using the primary operator identifier (POID). The POID is the identifier used by USDA's National Agricultural Statistics Service (NASS) for the ARMS and is unique for farms within each State. USDA constructed a dataset of outstanding direct and guaranteed loans which included the POID. ARMS data were then matched to the FSA loan data by using the POID.

Focus is on the identification of respondent errors when answering farm debt questions in the ARMS and on the magnitude of the errors. A multinomial logit model is estimated to identify demographic, structural, and financial characteristics of FSA borrowers who accurately and inaccurately classify their farm operations as having end of year debt on the ARMS. Additionally, the magnitude of response errors in ARMS for both FSA direct and guaranteed loan programs are estimated. Future researchers using ARMS data to appraise borrower debt status will be better informed about potential data limitations.

1.3 ARMS Overview

ARMS is an annual survey administered annually by the NASS, and is broken into three phases (USDA, ERS, 2015a). Phase I determines which operations are still in business. Phase II gathers information on production practices and input usage. Phase III assesses the finances of farm businesses and farm households. Of the three phases of the ARMS, the current study is interested in the Phase III survey which collects data on income, expenses, assets, liabilities, and operator demographics. Phase III tends to be long and complex in order to capture all the information needed to fully evaluate farm financial health and policy effects (Millar and O'Connor, 2012).

1.4 FSA Overview

The USDA's FSA provides loans to farm operators who are unable to obtain credit from conventional sources at equitable rates and terms. The FSA loan program benefits beginning farmers, socially disadvantaged (SDA) farmers, and established farmers facing temporary financial setbacks. The FSA sets aside some funds specifically for beginning and SDA farmers. A beginning farmer is an operator who has operated a farm ten or less years, does not have a farm thirty percent or larger than the average farm in their county, meets all FSA loan eligibility criteria, and contributes significantly in the operation of the farm. SDA borrowers have to meet all the FSA loan program eligibility criteria, and belong to a group that has historically been underserved because of ethnicity, race, and/or gender. According to the FSA, these SDA groups are: "American Indians or Alaskan Natives, Asians, Blacks or African-Americans, Native Hawaiians or other Pacific Islanders, Hispanics and women" (USDA, FSA, 2012).

FSA provides two main loan programs to eligible borrowers: direct and guaranteed (USDA, FSA, 2015a). Under the direct loan program, FSA provides loans directly to the borrower. The direct loan program has a number of loan types, of which four are considered: farm ownership (FO), operating loan (OL), emergency loan (EM), and economic emergency (EE). FO loans may be used to make land purchases and farm improvements. OL loans may be used to purchase livestock and equipment, pay for operating and family living expenses, and refinance debt under certain circumstances. EM loans are to help producers who have had production and physical losses as the result of drought, flooding, other natural disasters or quarantine. Although EE loans have not been originated since the early 1980s, they were for 40 year terms with some still having unpaid balances at the time of analysis. Farmers who were unable to receive credit from their usual lender due to national or area-wide economic stresses, such as general tightening of agricultural credit or an unfavorable relationship between production costs and prices received for agricultural commodities were eligible to receive EE loans (USDA, FSA, 2011).

The FSA guaranteed loan program guarantees loans made and serviced by commercial banks, the cooperative Farm Credit System, and other credit providers to eligible borrowers. This guarantee protects lenders against losses if the borrower does not meet their loan obligations by providing a guarantee of up to 95 percent of the loss of loan principal and interest. The guaranteed loan program consists of FO and OL loan types. In addition to direct FO loan purposes, guaranteed FO loans may be used to refinance debt.

1.5 Organization

Chapter 2 of this thesis contains a literature review of studies related to this study. Chapter 3 covers the data and methods. Summary statistics are discussed in Chapter 4. Chapter 5 discusses the model estimation results. Chapter 6 summarizes and interprets the econometric analysis, presents the conclusions, and includes future research recommendations.

Chapter 2: Literature Review

2.1 Literature Review

Many studies have been conducted in the recent past that use ARMS data and/or FSA data. While none of these studies have used FSA data to identify the frequency of respondent error in ARMS, they have looked into issues that surround the ARMS and the effectiveness of FSA loan programs. The following review of these studies covers four main topics: credit usage and debt, ARMS usability, nonresponse in ARMS, and FSA's farm loan programs. Lastly, a study that influences the current study's model specification is reviewed.

2.2 Credit Usage and Debt

Katchova (2005) examined the borrower's decision to use credit. Katchova's study identified the characteristics of individuals and farm operations that determine what influenced farm credit usage, the amount of credit, and number of loans. Katchova's research on credit usage provides insights into factors which are more likely to impact non-response or inaccurate reporting on questions regarding debt. Katchova's study sought to examine from the demand-side of agricultural credit by using 2001 ARMS data from borrowers and non-borrowers by farm typology: rural residence farms, intermediate farms, and commercial farms (Hoppe and MacDonald, 2013). One stipulation of the demand-side of agricultural credit is that the credit decision is jointly made between the lender and borrower. Therefore, the lender can affect the availability of funds to the farmer by restricting, rejecting, or modifying the funds. Katchova used a Probit model to identify the characteristics of farmers who are more likely to have debt. For those farmers that do use credit, Katchova used a truncated regression model to estimate the level of indebtedness. Lastly, she used a truncated Poisson model to determine the number of loans. Katchova's results show farms most likely to have debt are rural residence, intermediate farms, and commercial farms that have higher gross farm income. Older farmers of rural residences and intermediate size farms were less likely to have debt. Operator's age and income were the biggest influences on the degree in indebtedness. Lastly, farms with higher gross income and crop insurance tend to use more loans to finance their farming operation.

Briggeman, Koenig, and Moss (2012) discussed the importance of accurate and reliable farm debt data so that such data may be used to identify and lessen the effect of economic downswings in the agricultural sector. As an example, they discussed the buildup to and the occurrence of the farm debt crisis in the early 1980s. Up until the late 1990s and early 2000s, the USDA had more information to estimate U.S. farm debt.

USDA suspended state level accounts of farm debt because of the complication of using commercial bank data to estimate farm debt at the state level. With the reduction of these data sources, the ARMS became more important in estimating farm debt levels but the farm debt estimate also became less consistent (Briggeman, Koenig, and Moss, 2012).

Briggeman, Koenig, and Moss compare USDA's estimate of farm sector real estate debt to the sum of Farm Credit System and commercial bank real estate debt for 1985-2010. They argue there was a break in the estimate of farm sector real estate debt in 2000. While before 2000, the positive difference between the two series did not drastically change from year to year. However, the difference went to about zero in 2003 and remained near zero. In fact, the farm sector real estate debt reported by USDA was less than the debt reported by the Farm Credit System and commercial banks for several years. If real estate debt from other lenders was added, the difference would be even greater. This suggested there may be a sizable under-estimation of real estate debt by the ARMS.

Briggeman, Koenig, and Moss suggested the ARMS's respondents and the questionnaire structure may limit the ability to compute accurate, total farm debt amount. One category of respondents the ARMS does not cover is landowners who do not farm, which is also known as non-operator landlords. Farmland has influenced the amount of assets and collateral for debt reported on the agricultural sector balance sheet. While ARMS used debt numbers reported by the Farm Credit System and commercial banks, ARMS also subtracted out any non-farm debt. Subsequently, ARMS may underestimate farmland value and the debt on the farmland. ARMS does ask about loans from non-traditional sources; however, some questions and the structure of the farm debt section of the survey may be difficult for respondents to comprehend. The study suggested altering some questions and reordering other questions.

2.3 ARMS Usability

Ellinger, Ahrendsen, and Moss (2012) analyzed the possible implications of the economic measures listed on the farm firm's financial statements. By examining accounting principles and the ARMS questionnaire, certain items were found that limit ARMS data from fully gauging economic and financial conditions. These limitations impacted asset valuation, income and expense recognition, and extraordinary income reporting. In particular, data limitations could cause understatement of leverage, overstatement of liquidity, and under reporting of year-to-year farm income variability measures. Improved farm financial condition measures could be made by lessening the data limitations according to Ellinger, Ahrendsen, and Moss.

Featherstone, Park, and Weber (2012) examined how to obtain more information from the ARMS. They considered issues revolving around survey nonresponse, pseudo panels, and frequent updating of cost-of-production data on an enterprise basis. Featherstone, Park, and

Weber found three issues were in need of further research: nonresponse, refining methods to develop pseudo panel data, and developing methods to compile commodity specific financial data between the different commodity surveys.

Ahrendsen and Katchova (2012) discussed the financial performance calculated and reported from ARMS by the USDA Economic Research Service (ERS). They compared the financial measures to the recommended financial measures of the Farm Financial Standards Council (FFSC, 2015). The financial measures reported by ERS were duplicated and compared and contrasted to four other methods to calculate financial performance measures. Ahrendsen and Katchova recommended that ERS: 1) use the FFSC financial measure recommendations, 2) assess the policy for flagging estimates as statistically unreliable, 3) report medians, and 4) add the percentage of farm businesses that have financial values within critical zones.

Blank and Klinefelter (2012) posited that the usefulness and relevance of the ARMS data can be improved by refining the sample frame. Originally, the ARMS was created to meet Congress's requirement that data be collected on the production costs of wheat, feed grains, cotton, and dairy; therefore, not all states were sampled. This meant the farmers sampled provided a sufficient cross section of production cost data for those commodities listed, but did not provide sufficient information on other commodities or regions. Additionally, the people included in the survey may not be sufficiently representative because economic performance varies across both farm size and commodity specialization. The handful of targeted commodities did not provide a representative sample across all commodities since states not sampled may be more livestock extensive (Mountain region) while small grain operations in the Midwest region are over-represented. Therefore, ARMS needs to sample people by farm size and commodity strata. Furthermore, the survey's respondent burden tended to be heavier for large farms since they are fewer in number, and are surveyed more often than smaller operations. Blank and Klinefelter suggested respondent burden could be reduced for large farms by designing the survey for farm size, base questions on relevance, and make questions easier to understand by rewording them.

2.4 Nonresponse in ARMS

Miller, Robbins, and Habiger (2010) examined the challenges of missing data in the ARMS Phase III. According to their study, unit non-response (whole survey refused) and item non-response (certain items refused) lead to missing data that requires special handling during statistical analysis. To help alleviate this issue, NASS imputed information into missing items for any variables used in the published summary statistics. The NASS procedure eliminated outliers in the data then they use conditional averages equivalent to a regression on categorical variables.

Millar, Robbins, and Habiger (2010) selected around 100 variables from the ARMS Phase III to analyze. They conducted a detailed study of data plots and various aspects of the variables (continuous, categorical, or censored). One effect of mean imputation is that it reduces the variation in the data set. While imputation was usually done at a small rate, higher rates of imputation can cause a large downward bias in the variance in the data set. Millar, Robbins, and Habiger examined the distributions before and after machine imputation, and confirmed imputation altered the distribution. Lastly, the mean becomes biased if non-respondents are not like the respondents in regards to the item's value.

Earp et al. (2008) examined the effect calibration has on non-response bias in the ARMS Phase III. Non-response bias was potentially higher for the ARMS Phase III due to lower response rates, and NASS weighted the respondent sample so the estimated/calibrated variable totals for a large subset of items match target values from other sources (Earp et al., 2008). They compared the 2002 Census of Agriculture values of ARMS respondents to the values in the full sample of the 2005 ARMS respondents. Additionally, they used respondent data from the 2005 ARMS and non-respondent data from the 2002 Census of Agriculture.

Earp et al. found calibration weighting reduced bias in 90 percent of the study variables so that they were no longer significantly different from zero via a t-test. Of these variables, 50 percent had a significant reduction in bias levels via a paired t-test. One variable, fertilizer expenses, still had a significant bias via t-test after calibration. According to Earp et al., calibration appears to be an effective tool for reducing non-response bias.

The study by Gerling, Tran, and Earp (2008) examined the most common reasons for nonresponse in the 2006 ARMS III for Washington State. The ARMS Phase III generally has response rates lower than 80%, and has a potential to have a higher nonresponse bias than the other ARMS phases (Gerling, Tran, and Earp, 2008). While administering the ARMS Phase III, field enumerators asked operators who had declined to cooperate on the ARMS to explain why they refused to complete the survey. Gerling, Tran, and Earp found the top three reasons for refusal were: would not take time, will not do financial surveys, and information too personal. However, further research needs to be conducted to determine if results are survey, regional, or national specific. In summary, they recommended adding a cell to the survey for recording the reason for nonresponse.

Next, Weber and Clay (2013) analyzed non-response in the ARMS. According to Weber and Clay, approximately a third of sampled farming operators ignore the entire ARMS. Weber and Clay use the Census of Agriculture data in their study because the data provides information on ARMS respondents and non-respondents, and the data comes from the same questionnaire collected for both respondents and non-respondents. Initially, Weber and Clay began exploring the motivations and characteristics associated with non-response, and the differences in point estimates of two econometric models when estimated on two subsamples. One subsample consists of respondents, and the other subsample consists of a random drawing from a group of respondents and non-respondents. Weber and Clay found response rates decrease monotonically with increasing farm size. Also, non-responding farm operators have greater sales than respondent farm operators. Weber and Clay found minimal nonresponse bias in the two econometric models they estimated.

Reasons for non-response are many, but the most common reason for nonresponse is that the respondents will not take the time or are too busy to take the survey (Weber and Clay 2013). Some other reasons are the respondents will not fill out financial surveys, and the survey is too personal. The reasons imply that farm operators may be unwilling to provide personal information or farm operators may believe the information will be used against them. According to Weber and Clay, larger farms take more time to fill out the survey, and incur greater disutility from the task. Additionally, production has moved to larger operations that may have a greater legal and contractual complexity.

Every five years, NASS administers both the ARMS and the Census of Agriculture survey. Weber and Clay obtained the principal operator identifier (POID) of all the surveyed farming operators of the ARMS Phase III for the years 2003-2006 and 2008-2010. The POIDs were used to match the surveyed farming operators in 2003-2006 with their 2002 Census of Agriculture, and the surveyed farming operators in 2008-2010 with their 2007 Census of Agriculture. Across all of the years, Weber and Clay had 189,474 matched observations of which 67% were ARMS respondents, 28% were refusals, and 5% were inaccessible. Weber and Clay grouped variables into three categories: household characteristics, farm characteristics, and farm specialization.

Weber and Clay estimated a Probit model to explore nonresponse patterns. Their Probit model outcome was whether or not the farm operator responded to the survey. They found that the multivariate analysis supported most of the findings from the descriptive comparisons. As farm size increases, the probability of response decreases. Grain specialty farms had the lowest probability of responding while dairy farms were most likely to respond. Weber and Clay found that with every hour decrease in time necessary to fill out the ARMS, the probability of response increases by seven percent.

Weber and Clay used two econometric models to examine nonresponse bias. One was a model of labor market participation of the principal farm operator. The other model examined farm diversification discount. The model on labor market participation of the principal farm operator was a Probit model, and used variables such as operator age, operator experience, household size, commodity specialization, and region. The second model was constructed to examine whether diversified farms are discounted by the market similar to what has been found for corporate firms. The main explanatory variable was an indicator of farm diversification. The farm was either diversified across livestock and crops or they were specialized in either crops or livestock. The labor participation model coefficient estimates found age increased the likelihood of working off-farm, but at a decreasing rate. The probability of working off-farm decreased with farm size as well. Next, Weber and Clay found nonresponse bias in coefficient estimates was small in degree and low in frequency across both models. Lastly, they found nonresponse bias was unlikely to weaken conclusions based on econometric models using ARMS data. While Weber and Clay did not include loan debt in their analysis, their findings may give context to why respondents may refuse or inaccurately report their debt. For example, Weber and Clay found grain farms had the lowest response probability, and the current study utilizes a crop variable as an explanatory variable.

2.5 FSA's Farm Loan Programs

Ahrendsen et al. (2011) estimated a triple hurdle model of U.S. commercial bank usage of the FSA's guaranteed OL and interest assistance programs in order to identify the farm and banking variables that affect the bank's decision to use loan guarantee and interest assistance. They examined annual data on U.S. commercial banks from 1995-2003 in their model. Ahrendsen et al. found the farm debt servicing ratio, individual bank loan-to-asset ratio, bank size, and general guaranteed loan and interest assistance environment variables to all be statistically significant in all three hurdles. Another study by Dixon, Ahrendsen, and McCollum (1999) examined characteristics of banks and/or economic forces that influence the level of FSA loan guarantee programs commercial banks had within Arkansas, and they examined factors affecting the volume of loss claims. They used a six-equation model that is comprised of three double hurdle sub-models, and estimated as Probit equations. Dixon, Ahrendsen, and McCollum found more OL loans than FO loans in Arkansas, and commercial banks used guaranteed loans to add to the security of riskier loans. Commercial banks with certified or approved lender status were more likely to use guaranteed loans. Commercial banks with more guaranteed loans tend to have loss claims, and commercial banks who filed loss claims in the past were associated with filing loss claims in the present. Despite the consolidation of commercial banks, research did not find a reduction in the amount of FSA guaranteed loans made by commercial banks.

Lastly, Nwoha et al. (2007) focused on whether FSA direct loan targeting for beginning farmers and SDA farmers was financially necessary and utilized ARMS data from 2000-2003.

Nwoha et al. utilized the delete-a-group jackknife procedure to determine whether differences in means for certain variables were statistically significant between: 1) FSA direct loan eligible recipient and non-FSA direct loan eligible recipient, 2) SDA race and SDA gender, 3) SDA race and non-SDA, 4) SDA gender and non-SDA, and 5) beginning farmer and non-beginning farmer. Nwoha et al. used the financial characteristics solvency, liquidity, profitability, repayment capacity, and financial efficiency to compare the two groups. Nwoha et al. found weaker financial characteristics for FSA direct loan eligible farm operations relative non-FSA direct loan eligible operations. The financial ratios of SDA race and SDA gender were found to not be statistically different from one another. However, Nwoha et al. found SDA gender farms have significantly less farm assets, liabilities, equity, and gross and net cash farm incomes and smaller debt-to-asset ratio than non-SDA farms. Nwoha et al. found beginning farmers had a much smaller financial size (income statement and balance sheet measures) than non-beginning farmers. This indicated that the beginning farmer program was targeted to a set of farmers that were vastly different than regular FSA borrowers.

2.6 Model Specification

Dixon et al. (2007) researched FSA direct farm loan program (FLP) graduation rates, and the reasons behind borrowers exiting the program. Direct loans can be considered a transitory step for borrowers so that they graduate from FSA direct FLP assistance and obtain guaranteed FLP assistance as soon as they become financially able. Dixon et al. used a survey of borrower applications originating in fiscal years 1994-1996. The 2004 survey asked farm loan managers at the FSA field office level to specify why borrowers with no active direct loans exited the direct FLP. Additionally, financial information and demographic information on borrowers were provided by the farm loan managers.

Dixon et al. estimated a multinomial logit model in order to identify the relevant variables in predicting outcome type. The model identified indicators used to predict if a borrower: 1) remained a longer term client, 2) exited and continued farming, 3) voluntarily left for another occupation or retirement, or 4) involuntarily left farming. The multinomial logit model broke the outcomes (STATUS) into four categories. STATUS=1 were borrowers who had active loans on November 30, 2004. STATUS=2 were borrowers who exited the FLP, and were still farming with conventional credit, guaranteed credit, or no need for credit. STATUS=3 were borrowers who left farming voluntarily or retired. STATUS=4 were borrowers who left farming involuntarily excluding those who had died. These four outcomes stratified the independence of irrelevant alternatives (IIA).[1] Dixon et al. grouped the independent variables into four categories: borrower demographics, characteristics of the current loan, prior financial distress and involvement with FSA direct loans, and borrower financial characteristics. Demographic variables included borrower age, race, and gender. Current loan characteristics included: FO, OL, beginning farmer, and/or SDA, and were all binary variables equaling 1 if the loan had that characteristic. For prior distress, Dixon et al. used a variable (FINDIS) that indicated prior financial distress prior to loan application. Other variables included a count of the number of each loan type to indicate reliance and experience with the FSA. They hypothesized a higher FSA reliance would be inversely related to exiting direct FLPs. The financial characteristics included debt-to-asset ratio, net worth, ratio of non-farm income to total cash farm income sources, ratio of balance available for debt service to total debt service due that year, and total annual household net cash income. Dixon et al. posited that borrowers with higher net worth,

[1] Greene, W.H. *Econometric Analysis*, *Seventh edition.* Upper Saddle River: Prentice Hall, Inc. 2011: pages 767-768.

income diversification, repayment capacity, and income should have graduated from the direct FLP sooner.

Dixon et al. found borrowers were less likely to exit when they had higher numbers of active FSA direct loans at loan origination. Borrowers were more likely to exit and continue farming or voluntarily leave farming when they had less FSA direct loan involvement. Borrowers were more likely to leave farming voluntarily, and less likely to graduate from the direct FLP when they had higher debt-to-asset ratios; however, the opposite was found for those borrowers with higher net worth at origination. Borrowers were more likely to exit involuntarily when they had prior financial difficulties. Non-white borrowers were less likely to voluntarily leave farming. Beginning farmer loan borrowers were less likely to continue in the FLP and more likely to voluntarily leave farming. The results in general indicated FSA borrowers were not becoming permanent FSA clients, and FSA's goals were being met. Dixon et al. suggested strengthening financial requirements to loan origination in order to minimize farmers who come across financial hardship and left farming. However, strengthened financial requirements could exclude some of the farmers who were the intended recipients of the FSA direct FLP.

Chapter 3: Data and Methods

3.1 Data Sources

FSA provided data on active direct and guaranteed loans as of December 31 for calendar years 2001, 2004, 2006, and 2007 for borrowers who were ARMS respondents in the corresponding year. The FSA direct FLP data included the following information: state ID, POID, principal outstanding, interest outstanding, delinquent loan amount, adverse action loan amount (foreclosure, bankruptcy, or acceleration loan amounts), number of loans 90 days past due, average days past due, FSA score (like a credit rating), average interest rate, total direct FLP principal and interest outstanding, total FO principal and interest outstanding, total OL principal and interest outstanding, total emergency loan (EM) principal and interest outstanding, total economic emergency loan (EE) principal and interest outstanding, Black borrower identifier, Hispanic borrower identifier, Asian/Pacific Islander identifier, American Indian identifier, woman identifier, and beginning farmer (BF) identifier. The FSA guaranteed FLP data included the following borrower loan level information: state ID, POID, outstanding balance, initial loan amount, closing date, fiscal year, loan obligation date, maturity date, loan obligation number, assistance type, program type, fund code, interest assistance percentage, line of credit (LOC) indicator, LOC amount, originating lender, lender branch number, lender type code, fixed or variable rate indicator, borrower interest rate, lender interest rate, loan purpose code, and delinquency code. The direct FLP data were aggregated to the POID for the 2006 and 2007. In other words, all loan data per borrower was summed or averaged to a single line of information for that borrower; whereas, direct 2001 and 2004 and guaranteed data had each borrower’s loan(s) listed. Since the guaranteed FLP data and direct FLP data for 2001 and 2004 were at the loan level, aggregation to the borrower level was done so the data were similar to the 2006 and

2007 direct FLP data. This aggregation made it impossible to use variables on specific loan dates or originating lender.

The project also used data from the USDA's Phase III of ARMS. Using the POID, the FSA and ARMS data were merged for each year. In particular, the ARMS includes production, financial, and demographic information for the farm operator and household. ARMS data also includes summary variables calculated by ERS and NASS as well as data added to correct for inconsistencies and missing data. Specific variables include information such as: acres operated, rent received/paid, livestock production, income data, operating and capital expenditures, use of time, farm assets, farm debt, operator and spouse information, and operator and household information. The enumerator includes information on operator records use and operator records type.

The project is concerned with farm debt. The ARMS farm debt section includes information on: whether the operation has a positive debt balance at the end of the year, and if it does, there is a table with questions on what is the lender type, loan balance outstanding, loan interest rate, and other items for each loan up to a maximum of five loans in years 2001, 2004, and 2006, and four loans in year 2007. The farm debt section also includes how many additional loans and additional aggregated loan amount the operator has that are not in the maximum of four or five loans reported in detail in the debt-by-lender table. ERS computes the following variables: rate of return on assets (ROA), rate of return on equity (ROE), debt-to-asset ratio (DAR), and operating profit margin (OPM). By using the FSA data and the ARMS Phase III data, comparison of debt information provided by the FSA to the debt information reported in ARMS was able to be conducted.

3.2 Dependent Variable

The current study seeks to identify factors related to whether the respondent accurately or inaccurately reports having positive debt levels on the ARMS. In particular, a multinomial logit model[2] is estimated to identify those variables that indicate how a respondent answers the ARMS III farm debt question: "Did this operation owe any money to any banks, co-ops, individuals, merchants, or Federal agencies at the end of" the survey's respective calendar year.[3] This will be referred to as the "Owe Money" question and is the dependent variable of the multinomial logit model. The question has three possible responses: yes, no, and refusal. If the respondent answers "yes", they continue in the Farm Debt section to the designated debt-by-lender table of loan-specific questions. As previously discussed, the debt-by-lender table has questions on a per loan basis about lender type, loan balance outstanding, loan interest rate, etc. For example, the 2007 ARMS also asks questions on loan type, when loan will be repaid, loan origination date, loan term, loan purpose, interest rate type, frequency of repricing, and loan payments.[4] If the respondent answers "no" to the "Owe Money" question, they are directed to the next section of the ARMS, and zeros are recorded in the debt-by-lender table. When the respondent refuses to answer, a negative one is recorded for the response to the "Owe Money" question and for the responses in the debt-by-lender table.

[2] Greene (2011) equation 18.5, page 763.

[3] Phase III of ARMS for 2006 and 2007 added after the question "include money owed against your line of credit. Exclude CCC loans."

[4] The 2006 ARMS also asks questions on frequency of loan principal repayment and periodic repayments. The 2004 ARMS also asks questions on farm purpose percentage and loan guarantee. More information can be found at http://www.ers.usda.gov/data-products/arms-farm-financial-and-crop-production-practices/questionnaires-and-manuals.aspx#27921

When a question's response is recorded by the USDA, a number is associated with the variable corresponding to the question. The "Owe Money" question does not have a variable number associated with it for 2001 and 2004, but it does have a variable number (1058) associated with it for 2006 and 2007. Since recorded information is only available for the "Owe Money" question for 2006 and 2007, a proxy for the "Owe Money" question needs to be used so the 2001 and 2004 observations can be used. The ERS imputed a proxy variable (P999) for the "Owe Money" question, and a detailed description of the variable is given in Banker et al. (2010). When the entire debt-by-lender table was refused by the respondent, P999 equals one (and corresponds to refusing to answer the Owe Money question). When the respondent indicated no debt was outstanding, P999 equals three (and corresponds to answering "no" to the "Owe Money" question). Lastly, P999 equals zero when the respondent provided outstanding loan information in the debt-by-lender table (and corresponds to answering "yes" to the "Owe Money" question).

Since all respondents in the current study have outstanding FSA FLP debt at the end of the calendar year, all non-refusal respondents should answer "yes" to the "Owe Money" question, and all should have a P999 value of zero. By using the constructed P999 variable (DEBT_PROX) for all of the four years in the sample, respondents accurately reporting whether they have debt at the end of the year can be determined.

3.3 Independent Variable Considerations for the Models

When hypothesizing relevant independent variables for the models, the variables available in the combined data from the FSA FLP loan database and ARMS are examined. The independent variables are sorted into five categories: FSA Direct FLP loan characteristics, FSA

Guaranteed FLP loan characteristics, operator demographics, farm operation characteristics, and farm operation financial characteristics. Variable definitions are in Table 3.1.

Table 3.1 Variable Definitions for Direct and Guaranteed FLP and ARMS Data

Dependent Variable	**Definition**
DEBT_PROX	Equals 0 if operator responded "yes", equals 3 if operator responded "no", and equals 1 if operator refused to respond to the “Owe Money” question
Independent Variables	
FSA FLP variables	
FSADEBTTOTK	Total FSA direct and guaranteed FLP debt per borrower in thousands of dollars
INTRATE	Average interest rate of borrower’s direct loans
BORR_GUAR_INT_RATE	Average interest rate of borrower’s guaranteed loans
PASTDUE_IND	Equals 1 if any direct loans of borrower has days past due > 0, 0 otherwise
FO_DIR	Equals 1 if borrower has direct FO loan(s) only, 0 otherwise
FO_GTE	Equals 1 if borrower has guaranteed FO loan(s) only, 0 otherwise
OL_DIR	Equals 1 if borrower has direct OL loan(s) only, 0 otherwise
OL_GTE	Equals 1 if borrower has guaranteed OL loan(s) only, 0 otherwise
EMEE_DIR	Equals 1 if borrower has direct emergency loan(s) only, 0 otherwise
OL_LOC_GTE	Equals 1 if borrower has guaranteed OL line of credit loan(s) only, 0 otherwise
MULT_LN_DIR	Equals 1 if borrower has multiple direct loan types, 0 otherwise
MULT_LN_GTE	Equals 1 if borrower has multiple guaranteed loan types, 0 otherwise
MULT_PROG	Equals 1 if borrower has both direct and guaranteed loan types, 0 otherwise
Operator demographic variables	
OP_AGE	Age of primary operator in years
MARRIED	Equals 1 if operator is married, 0 otherwise
HS_EDUC	Equals 1 if operator has high school or less education, 0 otherwise
SC_EDUC	Equals 1 if operator has some college education, 0 otherwise
CGB_EDUC	Equals 1 if operator has college and/or beyond education, 0 otherwise
BF_ELIG	Equals 1 if primary operator is beginning farmer eligible (10 or fewer years since operating any operation), 0 otherwise
OP_SDA_P	Equals 1 if primary operator is SDA eligible, 0 otherwise

Table 3.1 Variable Definitions for Direct and Guaranteed FLP and ARMS Data Cont.

Independent Variables	
Operation characteristics Cont.	
HH_SIZE	Number of household members
CROP_RATIO	Value of crop production divided by total value of production
RECORD_USE_YES	Equals 1 if operator referred to records, 0 otherwise
RECORD_USE_SOME	Equals 1 if operator referred to loose receipts, 0 otherwise
RECORD_USE_NONE	Equals 1 if operator never referred to any records or receipts, 0 otherwise
RECORD_USE_MOT	Equals 1 if operator referred to records most of the time, 0 otherwise
RECORD_USE_SOT	Equals 1 if operator referred to records some of the time, 0 otherwise
RECORD_USE_NEV	Equals 1 if operator never refers to records, 0 otherwise
Y2001	Survey year 2001
Y2004	Survey year 2004
Y2006	Survey year 2006
Y2007	Survey year 2007
Operation financial characteristics	
IGCFIK	Gross cash farm income in thousands of dollars
ETOTK	Total expenses in thousands of dollars
INCFIK	Net cash farm income in thousands of dollars
EARNEDK	Household earned income in thousands of dollars
CAPEXP_TOTK	Total capital expenses in thousands of dollars
EFINTK	Interest expense in thousands of dollars
INTFEE_NREK	Non-real estate interest expense in thousands of dollars
INTFEE_REK	Real estate interest expense in thousands of dollars
NETWK	Net worth in thousands of dollars
ATOTK	Total assets in thousands of dollars
ACTOTK	Current assets in thousands of dollars
DTOTK	Total liabilities in thousands of dollars
LCTOTK	Current liabilities in thousands of dollars
Financial ratios	
NWC_EXPENSE_RATIO	Net working capital to total expense ratio measured in percent
CR	Current ratio (current assets / current liabilities)
DAR	Debt-to-asset ratio measured in percent
ROA	Rate of return on assets ((net farm income + interest expenses – estimated charges for operator labor and management) / total assets) measured in percent
OPM	Operating profit margin (net farm income / value of farm production) measured in percent

Table 3.1 Variable Definitions for Direct and Guaranteed FLP and ARMS data Cont.

Independent Variable	
Financial ratio variables cont.	
DRCU	Debt repayment capacity utilization (debt / debt repayment capacity) measured in percent
OER	Operating expense ratio (cash operating expenses / gross cash farm income) measured in percent
DEPER	Depreciation expense ratio
Source: Merged ARMS-FSA dataset (2001, 2004, 2006, 2007)	

For the FSA Direct FLP loan characteristics variables, the following variables are used: FSADEBTTOTK, PASTDUE_IND, INTRATE, FO_DIR, OL_DIR, EMEE_DIR, and MULT_LN_DIR. FSADEBTTOTK is the total outstanding FSA loan balance of direct and guaranteed loans measured in thousands of dollars at the end of the calendar year per borrower. The amount of outstanding debt (FSADEBTTOTK) should be a highly relevant factor related to whether the borrower accurately reports their debt or not. If a respondent's FSA debt is small, the respondent may not report the operation owes money, especially if the operation does not have other debt. PASTDUE_IND is a delinquency indicator equaling one when any of the aggregated loans has a days past due of one or more days for that year; otherwise, zero. INTRATE is the average interest rate for the direct loans for a particular borrower. Katchova 2005 found that interest rate impacted borrower demand for credit. Interest rate could impact respondent accuracy since the respondent may not remember their loan(s) interest rate.

FO_DIR, OL_DIR, EMEE_DIR, and MULT_LN_DIR indicate whether the borrower only has FO Direct FLP loans, only has OL Direct FLP loans only has emergency type Direct FLP loans, or has multiple Direct FLP loan types (FO_DIR, OL_DIR, and/or EMEE_DIR). EM and EE loans are combined into a single binary variable (EMEE_DIR) since there are relatively few EE loans with a remaining balance (Table 3.2).

Table 3.2 Total FSA EM and EE Observations

Total EM Loan Observations				
TOTEM	**Sample N**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
No	1,978	68.37	1,978	68.37
Yes	915	31.63	2,893	100
Total EE Loan Observations				
TOTEE	**Sample N**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
No	2,814	97.27	2,814	97.27
Yes	79	2.73	2,893	100

Source: Merged FSA-ARMS data set (2001, 2004, 2006, and 2007)
Note: TOTEM = Yes, if the operation has one or more EM loans; No otherwise.
Note: TOTEE = Yes, if the operation has one or more EE loans; No otherwise.

Selection of loan type usually depends on the intended use of the borrowed funds; therefore, it was important to capture any effects that loan type may have on accuracy of debt reporting. Since operating loans are short to intermediate term loans that are typically paid off early in the year, borrowers may not report them because they are paid off by the time the ARMS Phase III survey is administered in March and April; whereas, FO loans are long term loans and will likely have an outstanding balance for many years. A borrower with more than one type of loan should have an easier time remembering to report their FSA debt.

FSA Guaranteed FLP variables are: FSADEBTTOTK, BORR_GUAR_INT_RATE, FO_GTE, OL_GTE, OL_LOC_GTE, and MULT_LN_GTE. FSADEBTTOTK is the same as defined previously. BORR_GUAR_INT_RATE is the average interest rate of the borrower's guaranteed loans. FO_GTE, OL_GTE, OL_LOC_GTE, and MULT_LN_GTE indicate whether the borrower has only FO loans, only OL loans, only OL LOC loans, or has multiple loans types (FO, OL, and/or OL LOC). OL line of credit loans are unique to the guaranteed FLP dataset because FSA does not make direct FLP OL line of credit loans. The reasoning behind choosing these variables is the same as for direct loans.

Operator demographic variables give insight into the differences in characteristics among those operators who accurately report their debt, those operators who do not, and those who refuse to report. Operator demographic variables include: OP_AGE, HS_EDUC, SC_EDUC, CGB_EDUC, OP_SDA_P, and BF_ELIG. OP_AGE is a continuous age (in years) variable to reflect any effects age has on reporting accuracy. Previous studies indicated operator age has an impact on response outcome on the ARMS (Weber and Clay, 2013). Katchova (2005) also found older farmers of rural residences and intermediate farms were less likely to have debt. HS_EDUC, SC_EDUC, and CGB_EDUC are binary education variables. These are included because previous studies have indicated education impacts credit usage (Katchova, 2005). The level of education an operator has obtained may have an effect on the level of debt reporting accuracy. HS_EDUC is a binary variable that equals one when an operator has a high school or less education; otherwise, zero. SC_EDUC is a binary variable equaling one when an operator has some college education; otherwise, zero. CGB_EDUC is a binary variable equaling one when an operator is a college graduate and beyond; otherwise, zero. OP_SDA_P is a binary variable created using ARMS and FSA data. Reporting race or ethnicity on the FSA farm loan application is voluntary unless borrowers are applying for a SDA loan (USDA, FSA, 2015b). To circumvent this limitation, both FSA data and ARMS data were used to construct a SDA eligible variable equaling one when the borrower is SDA eligible; otherwise, zero. This is similar to the process used by Nwoha et al. (2007). A borrower is considered as an SDA when identified as such by either the FSA or ARMS data. This includes: 1) identified as a racial-ethnic minority or female in the FSA guaranteed FLP or direct FLP data or 2) when they indicated to be a racial-ethnic minority or female in the ARMS survey. BF_ELIG is a binary variable indicating whether an operator is beginning farmer eligible. The BF_ELIG variable is constructed by subtracting the year the operator began operating any operation from the ARMS survey year. Any operator with ten years or less of farming experience is considered to be a beginning farmer. Previous studies have indicated that socially disadvantaged farmers and beginning farmers have different characteristics and/or exhibit different behaviors compared to other FSA borrowers (Dixon et al., 2007; Nwoha et al., 2007).

The variables included in operation characteristics are: crop ratio (CROP_RATIO) and gross cash farm income measured in thousands of dollars (IGCFIK). A crop ratio (intensity) variable was computed by dividing the value of crop production by total production. Previous studies have determined agricultural type by proportion revenues from crops because crop operations have greater variation in revenues due to weather events (Settlage et al., 2001; Dixon et al. 2004). Weber and Clay (2013) found farm production specialization to impact the probability of responding to ARMS. Additionally, crop operations tend to have more borrowed capital for operating expenses (Settlage et al., 2001). IGCFIK is a good indicator of operation size. As stated in the literature review, larger operations have a higher ARMS non-response rate (Weber and Clay 2013). Katchova (2005) found farms with higher gross farm income are more likely to report debt and tend to report a greater number of loans. IGCFIK was included to capture any effects operation size has on reporting accuracy.

In order to capture any effects from a particular survey year, a binary variable was computed for each year. Y2001 indicates ARMS survey year 2001. Y2004 indicates 2004, Y2006 indicates 2006, and Y2007 indicates 2007.

3.4 Other Independent Variables Considered

Other variables were also considered in preliminary models, but were not included in the final model. In both the Direct FLP (DIR) model and the Guaranteed FLP (GTE) model, a multiple loan program variable (MULT_PROG) was substituted for the loan type variables. Although most borrowers had only direct loans or only guaranteed loans, there are some borrowers with both direct and guaranteed loans.[5] MULT_PROG was tried to determine if multiple loan program borrowers reported their debt differently from those with just one loan program type. When MULT_PROG was included, the significance of the gross cash farm income coefficient in the refusal vector of the direct model became insignificant and the multi-program coefficients were insignificant, and there were no other changes. Guaranteed borrowers operate larger farms than direct borrowers; therefore, in the direct model MULT_PROG is a proxy for farm size. Having more than one FLP type was concluded to not have a major impact on the outcome; therefore, the multiple loan program variable was omitted from the models. For the guaranteed FLP model, lender guaranteed interest rate (LEND_GUAR_INT_RATE) and interest assistance indicator (IA_IND) were originally included. Having these two variables instead of the borrower guaranteed interest rate (BORR_GUAR_INT_RATE) did not change the parameter estimates notably and significance levels. BORR_GUAR_INT_RATE was used due to its similarity with the Direct FLP interest rate (INTRATE) variable. Initially, MARRIED was included in the direct FLP and guaranteed FLP models; however, MARRIED showed no significance and made no substantive changes to either model when it was excluded. Also, two percent of MARRIED had missing values (55 of 2,696 observations for the direct FLP model and 51 of 2,714 for the guaranteed FLP model), which contributed to the decision to exclude it from the final models.

[5] Eighty-five percent of Direct FLP multinomial model observations are direct loans only, and 63 percent of Guaranteed FLP multinomial model observations are guaranteed loans only.

Number of household members (HH_SIZE) was also included in the direct FLP and guaranteed FLP models but it was found to be statistically insignificant. Some studies have used number of household members when determining who does not respond to the ARMS (Weber and Clay, 2013). However, the number of household members did not make any substantial changes to the estimated model coefficients when it was included; therefore, the variable was excluded from the final models. Farm organizational type (LEGAL_STAT) was also initially included since more complex organizations may be more likely to refuse to answer questions since their finances may be more difficult to incorporate in the debt-by-lender table. However, the "other" organizational type had very few observations. This is consistent with FSA's charge to lend to family-sized operations, which are primarily sole proprietorships. Weber and Clay (2013) found farm organizational type had an effect on response to the ARMS. However, when they re-estimated their model without farm size, the sign on the organizational type coefficient changed suggesting a relationship between farm size and organizational complexity.

As previously mentioned, some responses were provided by the NASS enumerator after the completion of the interview. One of these questions asked how often the respondent refers to records: all of the time, most of the time, some of the time, almost never, or never (USDA, NASS, p. 32, 2001). In order to reduce the five responses to three responses, all of the time and most of the time were combined into one response: most of the time. Some of the time and almost never was combined into one response: some of the time. The three responses to the question were: most of the time, some of the time, and never. Hypothetically, respondents who refer to records most of the time should be more accurate when reporting their debt on the survey. The record usage question (RECORD_USE) showed up as highly significant, supporting the hypothesis. However, many previously significant variables became insignificant in the direct and guaranteed models. Since RECORD_USE had 22.5 percent missing observations for the direct FLP model and 27.9 percent missing observations for the guaranteed FLP when estimated without RECORD_USE, the impact of RECORD_USE was pursued in more detail. The sample was restricted to observations with non-missing RECORD_USE values. Then RECORD_USE was omitted from the models to see if the change in coefficient significance resulted from RECORD_USE being included in the model or as a result of the large decrease in non-missing observations from the original sample. The significant variables in the original model lost significance because of the large change in sample size in the direct FLP and guaranteed FLP models; therefore, RECORD_USE was omitted from the models to allow use of the larger sample.

Another variable initially included in the models was household earned income (EARNED). Inclusion of this variable did not change other model parameter estimates nor was it significant. Total capital expenses (CAPEXP_TOTK) was considered, but CAPEXP_TOTK has a large number of missing observations so it was not included in the models. Some other variables considered were: depreciation expense ratio (DEPER), ROA, and total assets (ATOT). Adding DEPER to the models caused the significant coefficient on IGCFIK to lose significance in both the direct and guaranteed FLP multinomial models. Also, the significance on the "no" intercept in the Direct FLP multinomial model and the significance on the "refusal" intercept in the guaranteed FLP multinomial model lost significance. Adding ROA to the models caused the significant IGCFIK coefficient for the "refusal" outcome to lose significance. Otherwise, ROA did not change other model parameter estimates nor was it significant. Lastly, the addition of ATOT to the models caused the significant IGCFIK coefficient to lose significance in both direct and guaranteed models. The significant coefficient on OL LOC for the "no" outcome also lost significance. Otherwise, ATOT did not change other model parameter estimates nor was it significant.

Other variables considered were found to be either computed or imputed from the ARMS farm debt portion of the survey, i.e., these variables are not independent of the “Owe Money” responses so they are at least partially endogenous. For example in the debt-by-lender table in the ARMS, respondents give their total principal and interest outstanding per loan listed, and this information is used to compute total debt (DTOT). However, principal and interest outstanding is only in the debt-by-lender table if the respondent answered yes to the “Owe Money” question, which is the dependent variable for the models. Therefore, DTOT is dependent on “Owe Money”. In another example, interest expense (EFINT) is imputed in some situations using debt information from the debt-by-lender table. If a respondent indicated having zero interest expense in the expense section of ARMS, but indicated having loan debt in the debt section, then NASS estimates interest expense from the reported loan debt (USDA, NASS, 2009a). Likewise, if loan debt is indicated to be zero, but the respondent reported interest expense in the expense section, then the loan debt amount is estimated from the reported interest expense. Hence, EFINT is computed from the debt-by-lender table in some cases. Other variables initially considered, but were found to be dependent on the farm debt portion of the survey are: DAR, operating profit margin (OPM), and net working capital to expense ratio (NWC_EXPENSE_RATIO).

3.5 Model Type and Specification

As was briefly discussed earlier, a multinomial logit model (Greene, 2011) is estimated to identify those variables that indicate how a respondent answers the “Owe Money” question. The dependent variable is the ARMS question asking respondents if the operation owes money to any banks, co-ops, individuals, merchants, or Federal agencies at the end of that survey’s respective calendar year. This question may be used to determine the proportion of respondents responding erroneously since all respondents should answer “yes” if they have not refused to answer the question. The dependent variable, DEBT_PROX, has three nominal outcomes (yes, no, refusal), making the multinomial logit regression model an appropriate empirical model.

Since the direct FLP and guaranteed FLP are different loan programs and have borrowers with different circumstances influencing the selection of one program over the other, a separate model is estimated for each program. The Direct FLP multinomial model has DEBT_PROX as the dependent variable and the independent variable groups: Direct FLP, operator demographics, operation, and operation financial characteristics. The Guaranteed FLP multinomial model has DEBT_PROX as the dependent variable and the independent variable groups: guaranteed FLP, operator demographics, operation, and operation financial characteristics. The multinomial logit models expected coefficient signs of no and refusal vectors when yes is base, and their data sources are displayed in Table 3.3.

Table 3.3 Multinomial Logit Model Expected Coefficient Signs of No and Refusal Vectors when Yes is Base and Data Sources

Variable Name	**Outcome: No** Expected Coefficient Sign	**Outcome: Refusal** Expected Coefficient Sign	**Data Source**
Dependent Variable			
DEBT_PROX	Na	Na	ARMS
FSA Direct FLP Variables			
FSADEBTTOTK	-	+	FSA
PASTDUE_IND	+	+	FSA
INTRATE	-/+	-/+	FSA
OL_DIR	+	-/+	FSA
EMEE_DIR	-/+	-/+	FSA
MULT_LN_DIR	-	-/+	FSA
FSA Guaranteed FLP Variables			
FSADEBTTOTK	-	+	FSA
BORR_GUAR_INT_RATE	-/+	-/+	FSA
OL_GTE	+	-/+	FSA
OL_LOC_GTE	+	-/+	FSA
MULT_LN_GTE	-	-/+	FSA
Operator Demographics			
SC_EDUC	-	-	ARMS
CGB_EDUC	-	-	ARMS
OP_AGE	+	+	ARMS
OP_SDA_P	-/+	-/+	ARMS and FSA
BF_ELIG	-	-/+	ARMS
Operation Characteristics			
CROP_RATIO	-/+	-/+	ARMS
Y2004	-/+	-/+	ARMS
Y2006	-/+	-/+	ARMS
Y2007	-/+	-/+	ARMS
Operation Financial Characteristics			
IGCFIK	-	+	ARMS

One assumption underlying the multinomial logistic regression model is the independence of irrelevant alternatives (IIA). This assumption requires that if one outcome is dropped from the choice set, then the parameter estimates corresponding to the remaining alternatives will not change significantly when the omitted outcome is truly irrelevant (Greene, 2011). Normally the IIA is tested by computing appropriate Wald statistics via a Hausman-McFadden test. In order to do this, bootstrap estimates of the appropriate covariance matrices would have to be used. Given the complex ARMS sampling strategy, utilizing bootstrap covariance matrix estimates in constructing Wald statistics was deemed unadvisable since it is not clear what the appropriate distribution would be of the Wald statistic. Instead, the two binomial logit sub-models that would logically flow from the conventional IIA testing were estimated. In the first sub-model for a given program (direct FLP or guaranteed FLP), a binary logit model was estimated by deleting the refusal observations. In the second sub-model, the no responses were eliminated while the refusal responses were included. The yes responses were not eliminated since they make up 82.8 percent of the Direct FLP multinomial model observations and 85.5 percent of the Guaranteed FLP multinomial model observations.

The practical approach to addressing IIA concerns was to compare the resulting parameter estimates from the two sub-models with the corresponding vector of coefficients from the full model. The percentage changes in the coefficients were computed from the full model and the sub models. The resulting differences in percentage terms generally ranged from -10 percent to 10 percent. However, the guaranteed FLP program had a larger percentage change for most coefficients with a range of -20 percent to 20 percent and one coefficient in the direct FLP, (PASTDUE_IND), had a change of 90 percent. Therefore, an approximate z-test on coefficient equality was computed, and the null of equal coefficient values on three of the significant variables that varied by more than plus or minus 10 percent could not be rejected. This implies rejection of the IIA would not be likely if the test was able to be done formally, and the coefficients resulting from the combined sample are reliable with reference to the IIA. Furthermore, the validity of the IIA tests have been questioned in research. In particular, simulation research has shown that the Hausman-McFadden test and the Small-Hsiao test for IIA have performed poorly; hence, the IIA assumption tests are unsatisfactory and not recommended (Allison, 2012; Cheng and Long, 2007).

In the event that the IIA assumption could be rejected, two binomial models were estimated: direct FLP and guaranteed FLP. The “yes” and “refusal” outcomes were combined together since the “yes” and “refusal” outcomes had similar summary statistics on their independent variables when compared to those of the “no” outcome. The estimates for the Direct FLP binomial model can be found in Table 5.2 and the Guaranteed FLP binomial model in Table 5.4. The summary statistics for the Direct FLP and Guaranteed FLP binomial models can be found in Appendix A and Appendix B. The implications of these binomial models relative to the multinomial models are discussed below.

Chapter 4: Summary and Debt Statistics

4.1 Direct FLP Summary Statistics

Table 4.1 displays the summary statistics for the dependent and independent variables in the Direct FLP multinomial model and other variables of interest. The summary statistics show the mean and bootstrap standard errors for each of the dependent variable (DEBT_PROX) outcomes (yes, no, refusal). Of the 156,693 weighted observations for the dependent variable (DEBT_PROX) outcomes (yes, no, refusal), 129,682 (82.8 percent) are in DEBT_PROX's "yes" outcome and 6,806 (4.3 percent) are in DEBT_PROX's "refusal" outcome. Interestingly, 20,204 (12.9 percent) of the weighted observations are in DEBT_PROX's "no" outcome. This shows inaccuracy of reporting since an estimated 12.9 percent (13.6 percent of the sample) of the respondents answered "no" when they should have answered yes because they have outstanding FSA debt. This result indicates that respondents are not always reporting their information on the ARMS correctly, and other sections of the ARMS may experience inaccurately reporting as well.

Table 4.1 Direct FLP Mean and Bootstrap Standard Error Summary Statistics

Variables	**Outcomes**									
DEBT_PROX	**Yes**		**Refusal**			**No**			**All**	
	Mean	Btsp Std Err	Mean		Btsp Std Err	Mean		Btsp Std Err	Mean	Btsp Std Err
Sample N	2,162 (80.2%)		166 (6.2%)			368 (13.6%)			2,696 (100%)	
Weighted N	129,682 (82.8%)		6,806 (4.3%)			20,204 (12.9%)			156,693 (100%)	
Direct FLP Variables										
FSADEBTTOTK ($1000)	133.002	4.615	136.105		27.493	80.256	di	8.550	126.336	4.149
INTRATE	0.047	0.001	0.049		0.002	0.046		0.002	0.047	0.001
PASTDUE_IND	0.049	0.010	0.041		0.019	0.095		0.051	0.055	0.011
FO_DIR	0.380	0.023	0.544		0.107	0.341	i	0.059	0.382	0.021
OL_DIR	0.218	0.018	0.112	a	0.036	0.276	h	0.057	0.221	0.018
EMEE_DIR	0.266	0.018	0.262		0.079	0.324		0.060	0.273	0.017
MULT_LN_DIR	0.136	0.015	0.082		0.041	0.059	d	0.018	0.123	0.013
MULT_PROG	0.165	0.015	0.148		0.047	0.078	d	0.020	0.153	0.012
Borrower Demographics										
OP_AGE	51.018	0.528	55.070	b	1.944	54.931	d	1.328	51.699	0.462
HS_EDUC	0.472	0.022	0.684	a	0.077	0.578		0.062	0.495	0.021
SC_EDUC	0.331	0.022	0.185	a	0.047	0.283		0.061	0.318	0.020
CGB_EDUC	0.197	0.017	0.132		0.054	0.139	f	0.031	0.187	0.014
OP_SDA_P	0.208	0.018	0.071	a	0.027	0.203	h	0.047	0.202	0.017
BF_ELIG	0.161	0.018	0.022	a	0.012	0.136	h	0.044	0.152	0.016
MARRIED	0.893	0.012	0.902		0.040	0.860		0.047	0.889	0.011

Table 4.1 Direct FLP Mean and Bootstrap Standard Error Summary Statistics Cont.

Variables	**Outcomes**									
DEBT_PROX	**Yes**		**Refusal**			**No**			**All**	
	Mean	Btsp Std Err	Mean		Btsp Std Err	Mean		Btsp Std Err	Mean	Btsp Std Err
Sample N	2,162 (80.2%)		166 (6.2%)			368 (13.6%)			2,696 (100%)	
Weighted N	129,682 (82.8%)		6,806 (4.3%)			20,204 (12.9%)			156,693 (100%)	
Operation Characteristics										
HH_SIZE	3.070	0.073	2.738	b	0.131	2.631	d	0.126	2.998	0.062
RECORD_USE_MOT	0.683	0.025	0.297	a	0.116	0.378	d	0.080	0.633	0.025
RECORD_USE_SOT	0.148	0.019	0.487	b	0.164	0.199	i	0.060	0.166	0.019
RECORD_USE_NEV	0.169	0.020	0.217		0.090	0.423	di	0.084	0.200	0.020
RECORD_TP_YES	0.661	0.025	0.673		0.115	0.348	dh	0.073	0.625	0.025
RECORD_TP_SOME	0.215	0.021	0.126		0.071	0.354	h	0.084	0.228	0.021
RECORD_TP_NONE	0.109	0.016	0.176		0.079	0.275	e	0.079	0.131	0.016
CROP_RATIO	0.464	0.020	0.450		0.090	0.542		0.054	0.473	0.018
Y2001	0.261	0.021	0.338		0.131	0.322		0.073	0.272	0.021
Y2004	0.320	0.022	0.391		0.087	0.291		0.055	0.320	0.020
Y2006	0.204	0.017	0.178		0.060	0.164		0.036	0.198	0.015
Y2007	0.215	0.017	0.093	a	0.036	0.223	h	0.044	0.210	0.016
Operation Financial Characteristics										
IGCFIK ($1000)	208.862	8.666	268.239		46.095	127.068	dg	17.033	200.894	7.690
ETOTK ($1000)	162.548	6.490	178.921		36.831	98.822	dh	12.707	155.042	5.812
INCFIK ($1000)	46.314	3.191	89.319	a	14.655	28.246	dg	6.084	45.852	2.772
EARNEDK ($1000)	34.718	2.074	37.286		5.067	31.424		6.775	34.402	2.004
EFINTK ($1000)	16.248	0.673	11.981	c	2.450	5.215	dh	1.155	14.640	0.592
INTFEE_REK ($1000)	10.597	0.525	8.945		1.855	4.105	dh	1.105	9.688	0.470
INTFEE_NREK ($1000)	5.651	0.427	3.036	a	0.784	1.110	dh	0.207	4.952	0.364
CAPEXP_TOTK ($1000)	5.093	0.681	5.766		3.853	2.802	e	0.779	4.851	0.604

Table 4.1 Direct FLP Mean and Bootstrap Standard Error Summary Statistics Cont.

Variables	**Outcomes**							
DEBT_PROX	**Yes**		**Refusal**		**No**		**All**	
	Mean	Btsp Std Err	Mean	Btsp Std Err	Mean	Btsp Std Err	Mean	Btsp Std Err
Sample N	2,162 (80.2%)		166 (6.2%)		368 (13.6%)		2,696 (100%)	
Weighted N	129,682 (82.8%)		6,806 (4.3%)		20,204 (12.9%)		156,693 (100%)	
Operation Financial Characteristics Cont.								
NETWK ($1000)	694.148	27.741	969.147	181.818	702.269	63.133	707.140	25.816
ATOTK ($1000)	916.511	31.137	1,117.085	208.297	705.585 di	63.488	898.026	28.771
ACTOTK ($1000)	111.100	6.566	126.103	37.312	52.720 di	9.778	104.224	5.819
DTOTK ($1000)	222.363	8.315	147.938 b	36.502	3.316 dg	0.524	190.886	7.284
LCTOTK ($1000)	67.064	3.295	57.773	14.576	3.316 dg	0.524	58.441	2.853
Financial Ratio Variables								
Liquidity								
NWC_EXPENSE_RATIO (%)	35.227	0.098	41.469 a	0.136	63.624 dg	0.112	39.160	0.083
CR	5.012	1.189	16.785 b	5.154	38.032 dh	8.854	9.745	1.546
Solvency								
DAR (%)	29.940	1.173	12.302 a	3.053	0.709 dg	0.150	25.405	1.078
Profitability								
ROA (%)	-0.891	0.586	1.784	4.804	-2.853	1.787	-1.028	0.590
OPM (%)	-34.333	5.806	-44.714	30.128	-67.763	19.751	-39.094	6.001

Table 4.1 Direct FLP Mean and Bootstrap Standard Error Summary Statistics Cont.

Variables	**Outcomes**									
DEBT_PROX	**Yes**		**Refusal**			**No**			**All**	
	Mean	Btsp Std Err	Mean		Btsp Std Err	Mean		Btsp Std Err	Mean	Btsp Std Err
Sample N	2,162 (80.2%)		166 (6.2%)			368 (13.6%)			2,696 (100%)	
Weighted N	129,682 (82.8%)		6,806 (4.3%)			20,204 (12.9%)			156,693 (100%)	
Financial Ratio Variables Cont.										
Debt Repayment										
DRCU (%)	4.588	1.586	1.477	c	0.580	0.050	dh	0.285	3.868	1.307
Efficiency										
DEPER	0.201	0.028	0.115	b	0.019	0.131		0.034	0.188	0.023
OER (%)	125.942	11.218	77.415	b	15.558	111.663	h	7.676	121.993	9.454
Source: Merged FSA-ARMS data set (2001, 2004, 2006, and 2007)										
For most variables sample n=2,696 and weighted n=156,693. Record type variables, record use variables, and CAPEXP_TOTK have sample n=1,775 and weighted n=121,302. MARRIED has sample n=2,641 and weighted n=154,414. HH_SIZE and EARNEDK have sample n=2,612 and weighted n=153,689. CR has sample n=2,689 and weighted n=156,317. OPM, DEPER, and OER have sample n=2,695 and weighted n=156,692.										
Footnotes signifying significance levels for the difference in means. Yes-Refusal: a ($p < 0.01$); b ($p < 0.05$); c ($p < 0.10$). Yes-No: d ($p < 0.01$); e ($p < 0.05$); f ($p < 0.10$). Refusal-No: g ($p < 0.01$); h ($p < 0.05$); i ($p < 0.10$).										

When examining the summary statistics for the independent variables, a discrepancy between the mean amount of FSADEBTTOTK from the FSA data compared to the mean amount of DTOTK (total debt in thousands) from the ARMS data was discovered. DTOTK accounts for current and long-term liabilities and is the summation of farm liabilities from all lender types. For the "yes" and "refusal" outcomes, DTOTK of $222 thousand and $148 thousand are more than FSADEBTTOTK of $133 thousand and $136 thousand which is expected. Since the data are from only one lender (FSA), DTOTK should be far greater than FSADEBTTOTK. And the difference is significant ($p < 0.01$) for the "yes" outcome, although the difference is insignificant for the refusal respondents. The "no" outcome has a mean DTOTK of $3 thousand and FSADEBTTOTK has a mean of $80 thousand and the difference is statistically significant ($p < 0.01$). Since "no" respondents indicate they have zero debt in the Farm Debt section of the ARMS, the amount of debt reported on the ARMS is greatly under-estimated. For the estimated 20,204 operators responding "no" over the four years 2001, 2004, 2006, and 2007, an estimate of the amount of debt under reported is $1.554 billion or a simple average of $389 million per year. This is 6.5 percent of the $5.980 billion average reported by the Economic Research Service for FSA direct loans for those same four years (USDA, ERS, 2015b). However, an estimated 7.8 percent of the 20,204 operators had guaranteed loans in addition to direct loans so that the $389 million includes mostly direct FLP loans but also some guaranteed loans originated by other lenders. As stated in the literature review, the study by Briggeman, Koenig, and Moss (2012) found the amount of debt reported by lenders when added together was more than the amount of debt reported on the ARMS. The current study shows that one problem area for under-reporting of debt lies with those respondents indicating they have no outstanding debt at the end of the year when they should be indicating yes. Additionally, the summary statistics show DTOTK and

LCTOTK are the same amount which indicates DTOTK is only reflecting current liabilities for the respondents answering “no” on the ARMS. Operations responding “no” to the “Owe Money” question have significantly less FSA debt than those responding “yes” ($p < 0.01$) or refusing to answer ($p < 0.10$).

For the other independent variables in the Direct FLP multinomial model, the average age for respondents range between 51-55 years of age with 51.7 years of age being the overall mean. The “yes” respondents are about four years younger than the “refusal” ($p < 0.10$) and “no” ($p < 0.01$) respondents on average. Overall, 49.5 percent of the respondents have a high school or less education. A lower share of respondents refusing to answer the “Owe Money” question (0.185) have some college education than respondents answering “yes” (0.331) or “no” (0.283), although the difference is only significant ($p < 0.01$) for the “yes’ respondents. A smaller share of respondents refusing to answer the “Owe Money” question (0.07) are SDA eligible operators compared to those answering “yes” (0.21, $p < 0.01$) or “no” (0.20, $p < 0.05$). Additionally, respondents refusing to answer the “Owe Money” question have a lower mean ratio (0.02) of beginning farmer eligible operators compared to those answering “yes” (0.16, $p < 0.01$) or “no” (0.14, $p < 0.05$).

The summary statistics for RECORD_USE support our hypothesis since respondents answering “yes” to the “Owe Money” question have a mean ratio of 0.68 for RECORD_USE_MOT compared to a mean ratio of 0.38 for RECORD_USE_MOT for those answering “no.” Also, respondents answering “no” have a higher mean ratio (0.42) of RECORD_USE_NEV compared to the mean ratio (0.38) for using RECORDS_USE_MOT. We also observe operators answering “no” have a slightly higher mean ratio for CROP_RATIO (0.54) than those answering “yes” (0.46) or refusing to answer (0.45), although the difference is statistically insignificant. The summary statistics show respondents refusing to answer and those responding “yes” have a statistically higher mean IGCFI ($268 thousand and $209 thousand) than those responding “no” ($127 thousand). The summary statistics partially reflect what was stated in the literature review. Weber and Clay (2013) found that the probability of response decreases as farm size increases. However, the summary statistics presented here show that the “refusal” and “yes” respondent operations are larger than the “no” operations when size is measured by IGCFI, total expenses (ETOT), net cash farm income (INCFI), and total assets (ATOT).

4.2 Guaranteed FLP Summary Statistics

Table 4.2 displays the summary statistics for the dependent and independent variables in the guaranteed FLP multinomial model and other variables of interest. The summary statistics show the mean and bootstrap standard errors for each of the dependent variable (DEBT_PROX) outcomes (yes, no, refusal). Of the 91,771 weighted observations for the dependent variable (DEBT_PROX) outcomes (yes, no, refusal), 78,486 (85.5 percent) are in DEBT_PROX’s “yes” outcome and 4,197 (4.6 percent) are in DEBT_PROX’s “refusal” outcome and 9,087 (9.9 percent) are in DEBT_PROX’s “no” outcome. This shows inaccuracy of reporting since an estimated 9.9 percent (9.5 percent of the sample) of the respondents answered “no” when they should have answered yes because they have outstanding debt guaranteed by FSA.

Table 4.2 Guaranteed FLP Mean and Bootstrap Standard Error Summary Statistics

Variables	**Outcomes**							
DEBT_PROX	**Yes**		**Refusal**		**No**		**All**	
	Mean	Btsp Std Err	Mean	Btsp Std Err	Mean	Btsp Std Err	Mean	Btsp Std Err
Sample N	2,299 (84.7%)		158 (5.8%)		257 (9.5%)		2,714 (100%)	
Weighted N	78,486 (85.5%)		4,197 (4.6%)		9,087 (9.9%)		91,771 (100%)	
Guaranteed FLP Variables								
FSADEBTTOTK ($1000)	269.717	8.492	311.603	30.789	272.926	36.102	271.950	8.035
BORR_GUAR_INT_RATE (%)	7.288	0.116	7.156	0.225	7.680	0.277	7.321	0.101
FO_GTE	0.425	0.023	0.344	0.063	0.395	0.075	0.419	0.021
OL_GTE	0.123	0.015	0.120	0.041	0.07 f	0.024	0.117	0.014
OL_LOC_GTE	0.088	0.009	0.075	0.025	0.188 eh	0.042	0.097	0.009
MULT_LN_GTE	0.364	0.021	0.461	0.065	0.348	0.081	0.367	0.019
MULT_PROG	0.280	0.021	0.240	0.053	0.170 e	0.047	0.267	0.019
Borrower Demographics								
OP_AGE	49.498	0.513	50.632	1.262	48.739	1.006	49.475	0.451
HS_EDUC	0.453	0.021	0.473	0.071	0.607 f	0.083	0.470	0.020
SC_EDUC	0.316	0.021	0.390	0.068	0.257	0.089	0.314	0.020
CGB_EDUC	0.230	0.018	0.137 c	0.045	0.136 e	0.038	0.217	0.017
OP_SDA_P	0.107	0.014	0.080	0.027	0.050 d	0.014	0.10\\0	0.012
BF_ELIG	0.150	0.015	0.030 a	0.011	0.141 h	0.048	0.144	0.014
MARRIED	0.896	0.014	0.901	0.054	0.923	0.028	0.899	0.012

Table 4.2 Guaranteed FLP Mean and Bootstrap Standard Error Summary Statistics Cont.

Variables	**Outcomes**							
DEBT_PROX	**Yes**		**Refusal**		**No**		**All**	
	Mean	Btsp Std Err	Mean	Btsp Std Err	Mean	Btsp Std Err	Mean	Btsp Std Err
Sample N	2,299 (84.7%)		158 (5.8%)		257 (9.5%)		2,714 (100%)	
Weighted N	78,486 (85.5%)		4,197 (4.6%)		9,087 (9.9%)		91,770 (100%)	
Operation Characteristics								
HH_SIZE	3.165	0.066	3.348	0.256	3.089	0.149	3.166	0.059
RECORD_USE_MOT	0.717	0.025	0.313 a	0.099	0.403 d	0.106	0.676	0.025
RECORD_USE_SOT	0.139	0.020	0.144	0.059	0.280	0.126	0.152	0.022
RECORD_USE_NEV	0.143	0.018	0.543 a	0.110	0.317	0.105	0.172	0.018
RECORD_TP_YES	0.722	0.023	0.442 b	0.109	0.687 i	0.098	0.709	0.022
RECORD_TP_SOME	0.170	0.018	0.044 a	0.020	0.101	0.047	0.160	0.017
RECORD_TP_NONE	0.093	0.015	0.439 a	0.109	0.173 h	0.075	0.112	0.015
CROP_RATIO	0.518	0.017	0.542	0.056	0.716 dh	0.054	0.539	0.017
Y2001	0.216	0.021	0.042 a	0.037	0.090 e	0.045	0.195	0.019
Y2004	0.290	0.019	0.397	0.067	0.349	0.083	0.301	0.020
Y2006	0.257	0.018	0.252	0.068	0.281	0.064	0.259	0.017
Y2007	0.237	0.016	0.309	0.062	0.281	0.056	0.246	0.015
Operation Financial Characteristics								
IGCFIK ($1000)	363.806	15.247	400.367	52.220	242.727 dg	28.996	353.488	14.038
ETOTK ($1000)	282.774	10.805	298.569	32.999	176.871 dg	20.701	273.009	9.922
INCFIK ($1000)	81.031	6.520	101.800	31.180	65.856	13.194	80.479	5.979
EARNEDK ($1000)	33.121	2.279	28.658	4.097	38.306	6.724	33.438	2.067
EFINTK ($1000)	28.646	1.205	18.885 a	3.080	8.018 dg	1.355	26.156	1.098
INTFEE_REK ($1000)	19.082	0.787	13.276 b	2.662	5.030 dg	1.009	17.425	0.738
INTFEE_NREK ($1000)	9.563	0.774	5.609 b	1.659	2.988 dg	0.618	8.731	0.670

Table 4.2 Guaranteed FLP Mean and Bootstrap Standard Error Summary Statistics Cont.

Variables	**Outcomes**									
DEBT_PROX	**Yes**		**Refusal**			**No**			**All**	
	Mean	Btsp Std Err	Mean		Btsp Std Err	Mean		Btsp Std Err	Mean	Btsp Std Err
Sample N	2,299 (84.7%)		158 (5.8%)			257 (9.5%)			2,714 (100%)	
Weighted N	78,486 (85.5%)		4,197 (4.6%)			9,087 (9.9%)			91,770 (100%)	
Operation Financial Characteristics Cont.										
CAPEXP_TOTK ($1000)	10.148	1.360	1.793	a	0.754	2.341	d	0.869	9.178	1.200
NETWK ($1000)	854.954	37.561	1,209.685	a	97.265	825.610	g	104.783	868.272	33.786
ATOTK ($1000)	1,234.556	43.796	1,484.507	a	102.618	831.993	dg	105.134	1,206.123	39.231
ACTOTK ($1000)	182.214	9.325	143.989		25.219	99.836	d	19.439	172.309	8.380
DTOTK ($1000)	379.601	13.822	274.822	b	43.771	6.382	dg	0.992	337.851	12.858
LCTOTK ($1000)	121.866	6.295	106.356		17.423	6.382	dg	0.992	109.721	5.494
Financial Ratio Variables										
Liquidity										
NWC_EXPENSE_RATIO (%)	15.941	0.058	15.354	a	0.074	95.014	dg	0.165	23.745	0.052
CR	2.907	0.281	14.291	c	6.437	83.850	dg	24.988	11.300	2.305
Solvency										
DAR (%)	36.206	1.015	20.672	a	2.891	1.925	dg	0.549	32.101	1.023
Profitability										
ROA (%)	2.062	0.739	0.878		1.494	-6.549		7.452	1.155	0.951
OPM (%)	-14.006	3.865	-22.137		14.550	-4.070		6.986	-13.394	3.384

Table 4.2 Guaranteed FLP Mean and Bootstrap Standard Error Summary Statistics Cont.

Variables	**Outcomes**							
DEBT_PROX	**Yes**		**Refusal**		**No**		**All**	
	Mean	Btsp Std Err	Mean	Btsp Std Err	Mean	Btsp Std Err	Mean	Btsp Std Err
Sample N	2,299 (84.7%)		158 (5.8%)		257 (9.5%)		2,714 (100%)	
Weighted N	78,486 (85.5%)		4,197 (4.6%)		9,087 (9.9%)		91,770 (100%)	
Financial Ratio Variables Cont.								
Debt Repayment								
DRCU (%)	8.821	2.650	1.642 a	0.835	-0.543 dh	0.607	7.566	2.282
Efficiency								
DEPER	0.136	0.012	0.159	0.036	0.070 dh	0.016	0.131	0.011
OER (%)	92.235	3.455	87.405	9.427	108.121	19.028	93.587	3.538
Source: Merged FSA-ARMS data set (2001, 2004, 2006, and 2007)								
For most variables sample n=2714 and weighted n=91,770. Record type variables, record use variables, and CAPEXP_TOTK have sample n=1,758 and weighted n=66,204. MARRIED has sample n=2,663 and weighted n=89,695. HH_SIZE has sample n=2,646 and weighted n=89,384. CR has sample n=2,709 and weighted n=91,589.								
Footnote signifying significance levels for the difference in means. Yes-Refusal: a ($p < 0.01$); b ($p < 0.05$); c ($p < 0.10$). Yes-No: d ($p < 0.01$); e ($p < 0.05$); f ($p < 0.10$). Refusal-No: g ($p < 0.01$); h ($p < 0.05$); i ($p < 0.10$).								

As observed from the direct FLP summary statistics, a discrepancy between the mean amounts of FSADEBTTOTK from the FSA data compared to the mean amount of DTOTK (total debt in thousands) from the ARMS survey was discovered. The "no" outcome in the guaranteed FLP has a mean DTOTK of $6.4 thousand and FSADEBTTOTK has a mean of $272.9 thousand. The difference in the means was found to be significantly different from zero ($p < 0.010$. The weighted total of under-reported debt for the estimated 9,087 operators answering "no" over the four years of 2001, 2004, 2006 and 2007 is $2.422 billion or a simple average of $606 million per year. Almost $868 million dollars or $217 million per year more than what is under-reported in the direct FLP. However, the under-reported estimates of $1.554 billion from the direct FLP summary statistics and $2.442 billion from the guaranteed FLP summary statistics are not additive since 7.8 percent of the observations in the direct FLP summary statistics are also in the guaranteed FLP summary statistics. This is because the FSA debt variable includes both direct and guarantee indebtedness for those borrowers with loans from both programs. Again, the summary statistics show DTOTK and LCTOTK are the same amount which indicates DTOTK is only reflecting current liabilities for the respondents answering "no" on the ARMS survey. For those refusing to answer, it is surprising their FSADEBTTOTK mean is greater than the DTOTK mean by $37 thousand indicating an under-reporting of debt for an estimated total of $154 million, or a simple average of about $39 million per year. Although the difference is statistically insignificant, the under-reporting would only become greater as debts from other lenders are added. There is no statistical difference in the means of FSADEBTTOTK for those responding "yes", "no", and refusing to answer the owe money question of the ARMS survey.

For the other independent variables in the Guaranteed FLP multinomial model, more than twice the share of borrowers responding "no" (0.19) have only OL LOC loans than those responding “yes” (0.09) or those refusing to respond (0.08). The average age for respondents ranges between 49-51 years of age with 49.5 years of age being the overall mean. The mean age for guaranteed FLP operators is 2.2 years lower than the mean age reported in the direct FLP summary statistics ($p < 0.01$). Overall, nearly half (47.5 percent) of the respondents have a high school or less education, which is just 2.5 percentage points lower than the percent of respondents in the direct FLP summary statistics. A greater share of respondents answering “yes” (0.23) has a college education than those answering “no” (0.14, $p < 0.05$) or refusing to answer the “Owe Money” question (0.0.14, $p < 0.10$). Overall, 10 percent of the respondents are SDA eligible which is half the percentage of respondents in the direct FLP model. Respondents answering “no” to the “Owe Money” question (0.05) have a smaller mean ratio of SDA eligible operators compared to those answering “yes” (0.11, $p < 0.01$) or refusing to answer (0.08), although the latter is insignificant. Additionally, respondents refusing to answer the “Owe Money” question have a lower mean ratio (0.03) of beginning farmer eligible operators compared to those answering “yes” (0.15, $p < 0.01$) or “no” (0.14, $p < 0.05$) and is similar to what was found in the direct FLP summary statistics.

The summary statistics for RECORD_USE support the hypothesis since respondents answering “yes” to the “Owe Money” question have a mean ratio of 0.72 for RECORD_USE_MOT compared to a mean ratio of 0.40 and 0.31 for RECORD_USE_MOT for those answering “no” and for those refusing to answer. Operators answering “no” have a higher mean ratio for CROP_RATIO (0.72) than those answering “yes” (0.52) or refusing to answer (0.54). The summary statistics show respondents refusing to answer and those answering “yes” have a higher mean IGCFIK ($400 thousand and $364 thousand) than those responding “no”

($243 thousand). Additionally, respondents refusing to answer and those answering “yes” have a higher total expenses (ETOTK) and total assets (ATOTK) compared to those responding “no.”

4.3 DTOTK Under-reporting: Direct and Guaranteed FLPs

Since under-reporting of mean debt on the ARMS compared to the amount of total FSA mean debt was observed, a variable was constructed that equals one when DTOTK is less than FSADEBTTOTK, zero otherwise for the total number of observations available. For the Direct FLP, 66,998 out of 172,789 (38.8 percent) weighted operators have a DTOTK less than FSADEBTTOTK (Table 4.3). Also, those responding “no” have 21,970 out of 22,194 (99.0 percent) of weighted operators with DTOTK less than FSADEBTTOTK. Those responding “yes” have 39,523 out of 142,596 (27.7 percent) of weighted operators with DTOTK less than FSADEBTTOTK. Refusal respondents have 5,504 out of 7,993 (68.8 percent) of weighted operators with a DTOTK less than FSADEBTTOTK.

Table 4.3 Number of Observations with ARMS Total Debt less than FSA Total Direct and Guaranteed Debt for Direct FLP Operations

		Debt Proxy Outcomes			
	DEBT_UNDR	**Yes**	**Refusal**	**No**	**Total**
		Number			
Do not under report	Weighted N	103,073	2,494	223	105,791
	Sample N	1,685	89	12	1,786
	Respondent Percentage	59.65	1.44	0.13	61.23
	Yes, Refusal or No Percentage	72.28	31.18	1.01	
Do under report	Weighted N	39,523	5,504	21,970	66,998
	Sample N	626	97	385	1,108
	Respondent Percentage	22.87	3.19	12.72	38.77
	Yes, Refusal or No Percentage	27.72	68.82	98.99	
Total	Weighted N	142,596	7,993	22,194	172,789
	Sample N	2,311	186	397	2,894
	Percent (%)	82.53	4.63	12.84	100
Source: Merged FSA-ARMS data set (2001, 2004, 2006, and 2007)					

For the guaranteed FLP, 40,746 out of 99,176 (41.1 percent) weighted operators have a DTOTK lower than FSADEBTTOTK (Table 4.4). Also, those responding “no” have 8,958 out of 9,298 (96.4 percent) of weighted operators with DTOTK lower than FSADEBTTOTK. Those responding “yes” have 28,625 out of 84,579 (33.9 percent) of weighted operators with DTOTK lower than FSADEBTTOTK. Refusal respondents have 3,134 out of 5,298 (59.2 percent) of weighted operators with a DTOTK lower than FSADEBTTOTK. Both the Direct FLP and Guaranteed FLP show that operators responding “no” have a higher percentage that have DTOTK lower than FSADEBTTOTK followed by respondents refusing to answer. Operators responding “yes” have the smallest percentage with DTOTK less than FSADEBTTOTK. Those borrowers responding “no” are a definite problem area for ARMS estimation and accuracy. This result indicates that imputation for those respondents in the “no” outcome is difficult and needs improvement, although the difficulty of estimating total debt for many “yes” and “refusal” respondents is also apparent.

Table 4.4 Number of Observations with ARMS Total Debt less than FSA Total Direct and Guaranteed Debt for Guaranteed FLP Operations

		Debt Proxy Outcomes			
	DEBT_UNDR	**Yes**	**Refusal**	**No**	**Total**
		Number			
Do not under report	Weighted N	55,927	2,163	339	58,430
	Sample N	1,647	70	11	1,728
	Respondent Percentage	56.39	2.18	0.34	58.92
	Yes, Refusal or No Percentage	66.12	40.83	3.65	
Do under report	Weighted N	28,652	3,135	8,958.	40,746
	Sample N	749	110	256	1,115
	Respondent Percentage	28.89	3.16	9.03	41.08
	Yes, Refusal or No Percentage	33.88	59.17	96.35	
Total	Weighted N	84,579	5,298	9,298	99,176
	Sample N	2,396	180	267	2,843
	Percent (%)	85.28	5.34	9.38	100
Source: Merged FSA-ARMS data set (2001, 2004, 2006, and 2007)					

Chapter 5: Model Estimation Results

This chapter identifies those respondent characteristics that inaccurately report "no" to having current end of year debt in the Farm Debt section of the ARMS. The first section presents the estimated Direct FLP multinomial logistic model and discusses the variables with significant coefficients. Within that section, implications of the Direct FLP binomial logistic model in relation to the Direct FLP multinomial logistic model are discussed. The second section of this chapter presents the Guaranteed FLP multinomial logistic model and discusses the variables with significant coefficients. The implications of the Guaranteed FLP binomial logistic model in relation to the Guaranteed FLP multinomial logistic model are also discussed.

5.1 Direct FLP Multinomial Logistic Model Estimation Results

The estimated coefficients for the 16 independent variables (32 coefficients) in the Direct FLP multinomial model are presented in Table 5.1. The "no" intercept coefficient is highly significant ($p < 0.01$) and negative reflecting that "no" responses are generally less likely than "yes" responses. FSADEBTTOTK is highly significant ($p < 0.01$) and negative for the "no" outcome indicating that as total FSA debt decreases, the more likely a respondent will indicate "no" on the Farm Debt section of the ARMS. This outcome is plausible because a respondent with a small amount of FSA debt may not remember or bother to report their debt (Table 3.3). This result is consistent with the summary statistics, where the "no" mean is significantly less than the "yes" mean ($p < 0.01$).

Table 5.1 Direct FLP Multinomial Logistic Model Results and Odds Ratios

Analysis of Maximum Likelihood Estimates					
Parameter	**Outcome**	**Estimate**	**Wald ChiSq**	**Pr>ChiSq**	**Odds Ratio Est**
INTERCEPT	No	-2.388	7.637	p<0.01	na
INTERCEPT	Refusal	-2.529	1.458	ns	na
FSADEBTTOTK	No	-0.004	7.044	p<0.01	0.996
FSADEBTTOTK	Refusal	0.001	0.644	ns	1.001
INTRATE	No	0.719	0.649	ns	2.052
INTRATE	Refusal	-0.184	0.076	ns	0.832
PASTDUE_IND	No	-3.691	0.275	ns	0.025
PASTDUE_IND	Refusal	1.813	0.051	ns	6.131
OL_DIR	No	0.366	0.998	ns	1.443
OL_DIR	Refusal	-0.796	2.212	ns	0.451
EMEE_DIR	No	0.310	0.679	ns	1.363
EMEE_DIR	Refusal	-0.622	1.588	ns	0.537
MULT_LN_DIR	No	-0.473	1.148	ns	0.623
MULT_LN_DIR	Refusal	-0.682	1.036	ns	0.506
OP_AGE	No	0.018	2.570	ns	1.018
OP_AGE	Refusal	0.011	0.325	ns	1.011
SC_EDUC	No	-0.264	0.577	ns	0.768
SC_EDUC	Refusal	-0.893	5.710	p<0.05	0.409
CGB_EDUC	No	-0.498	2.327	ns	0.608
CGB_EDUC	Refusal	-0.634	1.649	ns	0.530
OP_SDA_P	No	0.191	0.272	ns	1.210
OP_SDA_P	Refusal	-0.850	3.971	p<0.05	0.428
BF_ELIG	No	0.235	0.286	ns	1.265
BF_ELIG	Refusal	-1.616	5.102	p<0.05	0.199
CROP_RATIO	No	0.700	4.697	p<0.05	2.013
CROP_RATIO	Refusal	-0.264	0.284	ns	0.768
IGCFIK	No	-0.001	3.327	p<0.10	0.999
IGCFIK	Refusal	0.0004	4.378	p<0.05	1.000
Y2004	No	-0.208	0.243	ns	0.812
Y2004	Refusal	0.011	0.000	ns	1.011
Y2006	No	-0.152	0.142	ns	0.859
Y2006	Refusal	-0.352	0.197	ns	0.704
Y2007	No	0.094	0.053	ns	1.098
Y2007	Refusal	-1.028	1.912	ns	0.358
Source: Merged FSA-ARMS data set (2001, 2004, 2006, and 2007)					
Notes: Sample N= 2,696; Weighted N= 156,693					

SC_EDUC is significant ($p < 0.05$) and negative for the "refusal" outcome, and indicates operators with some college education are less likely to refuse to answer compared to operators with high school or less education. In the summary statistics, the "refusal" mean is statistically different from the "yes" mean ($p < 0.01$). Since SC_EDUC is a binary variable, this indicates that proportions of respondents with some college education are different for those refusing to answer the "Owe Money" question and for those responding "yes." One plausible reason the some college coefficient is significant, but the CGB_EDUC is not, may be due to the fact that more education is synonymous with a more complex farming operation structure. More complexity in operation structure and finances may make reporting more difficult for respondents. The negative signs ($p < 0.05$) on both the OP_SDA_P and BF_ELIG "refusal" coefficients indicate SDA and beginning farmer eligible operators are less likely to refuse to answer the "Owe Money" question on the Farm Debt portion of the ARMS than operators not in these classes. It is likely women SDA comprise more of the OP_SDA_P observations than the race/ethnic SDA. According to the 2002 Census of Agriculture, women principal operators comprised 11.2 percent of the total farm operations while race/ethnic principal operators comprised 5.2 percent of the total farm operations (USDA, NASS, 2009b). In the 2007 Census of Agriculture, women principal operators comprised 13.9 percent of the total farm operations while race/ethnic principal operators comprised 6.6 percent of the total farm operations (USDA, NASS, 2009b). While sample size restrictions will not allow a breakdown by gender and race/ethnicity, it is likely that there are disparities in reporting debt by group. The 2002 Census of Agriculture indicated women principal operators were more likely to use computers for business and have internet access than male principal operators (USDA, NASS, 2005). Women operators may be better record keepers and may report more accurately than their male counterparts, and the current study's results support this assertion. The result for beginning farmer is expected because beginning farmers are required to participate in borrower training which may make them better at handling and understanding their finances and have financial records (USDA, FSA, 2015c). In the summary statistics, OP_SDA_P and BF_ELIG have "refusal" means significantly different from the "yes" means ($p < 0.01$). Since OP_SDA_P and BF_ELIG are binary variables, this indicates that the proportions of respondents that are either SDA or beginning farmers are different for those refusing to answer the "Owe Money" question and for those responding "yes."

CROP_RATIO is significant and positive for the "no" outcome ($p < 0.05$) and indicates respondents with more crop intense farms are more likely to respond "no" on the ARMS. Crop operations may have short term operating loans and may pay them off at the beginning of the year before the ARMS survey is administered in March and April. Both the "no" and "refusal" outcome coefficients are significant for IGCFIK. The negative sign on the "no" coefficient ($p < 0.10$) on IGCFIK means respondents with a lower gross cash farm income are more likely to say "no." This is expected because the summary statistics showed respondent's answering "no" had smaller IGCFIK, ATOT, and ETOT. The positive sign on the "refusal" coefficient ($p < 0.05$) on IGCFIK implies respondents with a higher gross cash farm income are more likely to refuse to answer the "Owe Money" question on the Farm Debt section of the ARMS. As noted in the literature review, Weber and Clay (2013) found that the probability of entire survey nonresponse increases as farm size increases. Since IGCFIK is an indicator of farm size, these results are consistent with Weber and Clay's (2013) results.

In regards to the Direct FLP binomial logistic model, the significant coefficients on the variables for "refusal" in the Direct FLP multinomial logistic model are not in the Direct FLP binomial model because the refusal and yes categories are combined into a yes/refusal category. All variables with significant coefficients for the “no” outcome in the Direct FLP multinomial model have significant coefficients at similar significance levels, with the same signs, and similar magnitudes in the Direct FLP binomial model (Table 5.2). Moreover, all of the coefficients for the “no” outcomes have the same sign and similar magnitude for the two Direct FLP models. Considering the similarities between the Direct FLP multinomial model and Direct FLP binomial model, confidence in the “no” coefficient estimates in the Direct FLP multinomial model is boosted.

Table 5.2 Direct FLP Binomial Logistic Model Results and Odds Ratios

Analysis of Maximum Likelihood Estimates				
Parameter	**Estimate**	**Wald ChiSq**	**Pr>ChiSq**	**Odds Ratio Estimate**
INTERCEPT	-2.464	8.238	p<0.01	na
FSADEBTTOTK	-0.004	7.131	p<0.01	0.996
PASTDUE_IND	0.729	0.667	ns	2.074
INTRATE	-3.749	0.287	ns	0.024
OL_DIR	0.406	1.243	ns	1.501
EMEE_DIR	0.347	0.863	ns	1.415
MULT_LN_DIR	-0.442	1.006	ns	0.643
OP_AGE	0.017	2.423	ns	1.017
SC_EDUC	-0.225	0.425	ns	0.799
CGB_EDUC	-0.470	2.099	ns	0.625
OP_SDA_P	0.218	0.356	ns	1.244
BF_ELIG	0.263	0.355	ns	1.301
CROP_RATIO	0.713	4.890	p<0.05	2.04
IGCFIK	-0.001	3.444	p<0.10	0.999
Y2004	-0.205	0.237	ns	0.815
Y2006	-0.130	0.105	ns	0.879
Y2007	0.136	0.111	ns	1.146
Source: Merged FSA-ARMS data set (2001, 2004, 2006, and 2007)				
Notes: Sample N= 2,696; Weighted N: 156,693; The “yes” and “refusal” outcomes are combined into one category and the “no” outcome is the other category.				

5.2 Guaranteed FLP Multinomial Logistic Model Estimation Results

The estimated coefficients for the 30 independent variables in the Guaranteed FLP multinomial model are presented in Table 5.3. The “no” intercept coefficient is highly significant and negative ($p < 0.01$). The same result was found in the Direct FLP multinomial model. In the Direct FLP multinomial model, FSADEBTTOTK is highly significant ($p < 0.01$) and negative for the “no” outcome; however, FSADEBTTOTK is not significant for the “no” or “refusal” Guaranteed FLP multinomial model outcomes. Those answering “no” in the Direct FLP multinomial model have a smaller mean FSADEBTTOTK than those answering “yes” or refusing to answer. Whereas the Guaranteed FLP multinomial model FSADEBTTOTK means between “yes”, “refusal”, or “no” do not vary much. OL_LOC_GTE is marginally significant ($p < 0.10$), positive on the “no” outcome, and indicates respondents with only OL LOC loans are more likely to respond “no” on the “Owe Money” question on the Farm Debt section of the ARMS compared to respondents with only FO loans. This is expected because OL LOC loans are short term loans, and the respondent may not report the loan because they paid it off at the beginning of the year before the ARMS is administered in March and April. Moreover, OL LOC are short term loans and they may have a relatively small balance at the end of the year. The respondent may not remember or bother to report an OL LOC balance at the end the year because the operator does not consider the loan important enough to report. In the summary statistics, the “no” mean is significantly different from the “yes” mean ($p < 0.05$). Since OL_LOC_GTE is a binary variable, this indicates that proportions of respondents with OL line of credit loans are different for those refusing to answer the “Owe Money” question and for those responding “yes.”

Table 5.3 Guaranteed FLP Multinomial Logistic Model Results and Odds Ratios

Analysis of Maximum Likelihood Estimates					
Parameter	**Outcome**	**Estimate**	**Wald ChiSq**	**Pr>ChiSq**	**Odds Ratio Estimate**
Intercept	No	-3.831	7.784	p<0.01	na
Intercept	Refusal	-4.939	2.163	ns	na
FSADEBTTOTK	No	0.001	1.649	ns	1.001
FSADEBTTOTK	Refusal	0.000	0.221	ns	1.000
BORR_GUAR_INT_RATE	No	0.153	2.629	ns	1.166
BORR_GUAR_INT_RATE	Refusal	0.051	0.498	ns	1.052
OL_GTE	No	-0.170	0.157	ns	0.843
OL_GTE	Refusal	0.244	0.253	ns	1.277
OL_LOC_GTE	No	0.711	3.148	p<0.10	2.036
OL_LOC_GTE	Refusal	0.116	0.059	ns	1.123
MULT_LN_GTE	No	-0.084	0.042	ns	0.919
MULT_LN_GTE	Refusal	0.463	1.463	ns	1.589
OP_AGE	No	-0.613	1.865	ns	0.542
OP_AGE	Refusal	0.142	0.165	ns	1.153
SC_EDUC	No	-0.923	6.704	p<0.01	0.397
SC_EDUC	Refusal	-0.542	1.321	ns	0.582
CGB_EDUC	No	-0.016	1.567	ns	0.984
CGB_EDUC	Refusal	-0.004	0.067	ns	0.996
OP_SDA_P	No	-0.810	4.073	p<0.05	0.445
OP_SDA_P	Refusal	0.056	0.015	ns	1.058
BF_ELIG	No	-0.095	0.038	ns	0.909
BF_ELIG	Refusal	-1.853	16.728	p<0.01	0.157
CROP_RATIO	No	1.352	9.144	p<0.01	3.866
CROP_RATIO	Refusal	-0.061	0.028	ns	0.941
IGCFIK	No	-0.002	6.592	p<0.05	0.998
IGCFIK	Refusal	0.000	0.162	ns	1.000
Y2004	No	1.453	3.444	p<0.10	4.275
Y2004	Refusal	2.083	0.440	ns	8.030
Y2006	No	1.184	2.541	ns	3.267
Y2006	Refusal	1.760	0.311	ns	5.813
Y2007	No	1.215	2.835	p<0.10	3.372
Y2007	Refusal	2.008	0.408	ns	7.446
Source: Merged FSA-ARMS data set (2001, 2004, 2006, and 2007)					
Note: Sample N= 2,714; Weighted N= 91,771					

SC_EDUC is highly significant ($p < 0.01$), negative for the "no" outcome, and indicates operators with some college education are less likely to respond "no" compared to operators with high school or less education. The Guaranteed FLP multinomial model result is different from the Direct FLP multinomial model result because the significant Direct FLP multinomial model result is on the refusal outcome. OP_SDA_P is significant and negative for the "no" outcome ($p < 0.05$) which indicates SDA respondents are less likely to respond "no" to the "Owe Money" question. In the summary statistics, the "no" mean is significantly different from the "yes" mean ($p < 0.01$). Since OP_SDA_P is a binary variable, this indicates that proportions of SDA respondents are different for those refusing to answer the "Owe Money" question and for those responding "yes." In the Direct FLP multinomial model, OP_SDA_P is significant for the refusal outcome. BF_ELIG is highly significant, negative for the "refusal" outcome ($p < 0.01$), and indicates beginning farmer respondents are less likely to refuse responding to the "Owe Money" question relative to answering "yes". The summary statistics show BF_ELIG's "refusal" mean is significantly different from the "yes" mean ($p < 0.05$). Since BF_ELIG is a binary variable, this indicates that proportions of beginning farmer respondents are different for those refusing to answer the "Owe Money" question and for those responding "yes." The Direct FLP multinomial model had the same result except at a different significance level ($p < 0.05$).

CROP_RATIO is highly significant and positive for the "no" outcome ($p < 0.01$) and indicates respondents with more crop intense farms are more likely to respond "no" on the ARMS. In the summary statistics, the "no" mean is statistically different from the "yes" mean ($p < 0.01$). Since CROP_RATIO is the share of the total value of production from crops, this indicates that the respondents that answer "no" to the "Owe Money" question have a significantly greater share of total value of production from crops on average than do those respondents that answer “yes.” The Direct FLP multinomial model has the same outcome except the significance level ($p < 0.05$). The negative sign on the “no” coefficient ($p < 0.05$) on IGCFIK means respondents with a lower gross cash farm income are more likely to say “no.” The summary statistics show IGCFIK’s “no” mean is statistically different from the “yes” mean ($p < 0.01$). As stated in the previous section, the summary statistics showed respondents answering “no” had smaller IGCFIK, ATOTK, and ETOTK which is expected. The Direct FLP multinomial model has the same result except the significance level ($p < 0.10$). The “no” Y2004 and Y2007 variables are slightly significant and positive ($p < 0.10$) which indicates that survey years 2004 and 2007 respondents are more likely to respond “no” to the “Owe Money” question compared to survey year 2001 respondents. Between 2003 and 2012, a shorter version of the core survey was mailed out to operators, and the larger sampling size increased usable responses to 20,000 or more compared to 10,000 originally (USDA, ERS, 2015a). The Guaranteed FLP multinomial model results for survey years 2004 and 2007 could be influenced by the increased number of usable responses. Also, survey year 2007 was a census year and the ARMS is longer and appears different compared to non-census years. For instance, ARMS survey year 2007 only has four columns (five normally) for information in the Farm Debt section debt-by-lender table. The Farm Debt section debt-by-lender table is transposed and looks slightly different compared to 2001, 2004, and 2006. However, the Direct FLP multinomial model did not have any year coefficients significant.

In regards to the Guaranteed FLP binomial logistic model, the significant coefficient on BF_ELIG for “refusal” in the Guaranteed FLP multinomial logistic model is not in the Guaranteed FLP binomial model because the refusal and yes categories were combined into a yes/refusal category. Otherwise, all variables (except Y2007) with significant coefficients for the “no” outcome in the Guaranteed FLP multinomial model had significant coefficients, same signs, and similar magnitudes as in the Guaranteed FLP binomial model (Table 5.4).

Table 5.4 Guaranteed FLP Binomial Logistic Model Results and Odds Ratios

Analysis of Maximum Likelihood Estimates				
Parameter	**Estimate**	**Wald ChiSq**	**Pr>ChiSq**	**Odds Ratio Est**
INTERCEPT	-3.824	7.845	p<0.01	na
FSADEBTTOTK	0.001	1.592	ns	1.001
BORR_GUAR_INT_RATE	0.151	2.577	p<0.01	1.163
OL_GTE	-0.183	0.184	ns	0.833
OL_LOC_GTE	0.704	3.157	p<0.01	2.022
MULT_LN_GTE	-0.112	0.076	ns	0.894
OP_AGE	-0.622	1.914	p<0.10	0.537
SC_EDUC	-0.897	6.274	p<0.01	0.408
CGB_EDUC	-0.016	1.547	ns	0.984
OP_SDA_P	-0.814	4.198	p<0.01	0.443
BF_ELIG	-0.034	0.005	ns	0.966
CROP_RATIO	1.353	9.182	p<0.01	3.869
IGCFIK	-0.002	6.545	p<0.01	0.998
Y2004	1.383	3.140	p<0.01	3.985
Y2006	1.134	2.350	p<0.05	3.108
Y2007	1.150	2.553	p<0.05	3.158
Source: Merged FSA-ARMS data set (2001, 2004, 2006, and 2007)				
Note: Sample N= 2,714; Weighted N= 91,771				

The direct FLP offers emergency and economic emergency loans which allows farmers to obtain loans to help with drought, natural disaster, and economic stress. The guaranteed FLP offers operating line of credit loans, and the guaranteed FLP has much higher loan limits than the direct FLP. The guaranteed FLP may also have farm borrowers with slightly better financial characteristics than direct FLP borrowers since guaranteed loans originate with a commercial lender instead of with FSA, although the lender has required a guarantee. The Direct FLP and Guaranteed FLP multinomial models have only a few differences between them. Although both have about the same number of significant coefficients (eight for the Direct FLP and nine for the Guaranteed FLP), the Direct FLP multinomial model had more significant coefficients on the "refusal" outcome and the Guaranteed FLP multinomial model had more on the "no" outcome.

Chapter 6: Conclusion

This chapter summarizes the conclusions to the current study, and offers suggestions for future research.

6.1 Conclusion

The study used FSA and ARMS data to determine the magnitude of respondent errors when answering farm debt questions in the ARMS. Multinomial logistic models were used to identify demographic, structural, and financial characteristics of FSA borrowers who accurately or inaccurately classify their farm operations as having end of year debt as well as those who refuse to indicate if they have end of year debt.

Estimates of 12.9 percent of direct FLP operators and 9.9 percent of guaranteed FLP operators responded “no” to the “Owe Money” question on the Farm Debt section of the ARMS. Inaccurate reporting in the Farm Debt section of the ARMS could also mean that other sections are subject to inaccurate reporting as well. DTOT was also observed as being under-reported for an estimated 38.8 percent of direct FLP operators and 41.1 percent of guaranteed FLP operators. These percentages only consider FSA direct and guaranteed loan indebtedness and are likely much higher if non-FSA related loans are added. Furthermore, DTOT was under-reported the most when operators respond “no” to the “Owe Money” question in both the direct FLP (98.9 percent) and the guaranteed FLP (96.4 percent). Both the direct and guaranteed FLPs show DTOTK is under-reported for those answering “no” to the “Owe Money” question. From the direct FLP statistics, the weighted under-reporting estimate is nearly $1.554 billion for the “no” outcome. From the guaranteed FLP statistics, the weighted under-reporting estimate is $2.442 billion for the “no” outcome. The under-reporting estimates are for the entire four years of the study: 2001, 2004, 2006, and 2007. Those respondents answering “no” are a problematic source of under-reporting of debt in the debt section of the ARMS Phase III. The current study only looks at FSA debt, and the observed under-reporting could be far greater when other lenders, such as commercial banks and Farm Credit System, are added.

The Direct and Guaranteed multinomial logistic models had the intercept, SC_EDUC, OL_SDA, BF_ELIG, CROP_RATIO, and IGCFI as significant coefficients in common. The Direct and Guaranteed multinomial logistic model results showed education, SDA and beginning farmer, and operation type and size as significant characteristics for determining when an operator responds “no” or refuses to respond. Operators with some college education were less likely to refuse in the Direct FLP multinomial logistic model relative to respondents with a high school or less education, and were less likely to respond “no” in the Guaranteed FLP multinomial logistic model. SDA operators were less likely to refuse in the Direct FLP and less likely to respond “no” in Guaranteed FLP multinomial logistic models. This means SDA operators and operators with some college education are more likely to have correctly reported debt. Beginning farmers were less likely to refuse in both the Direct and the Guaranteed FLP multinomial logistic models. As CROP_RATIO increased in both the Direct and Guaranteed FLP multinomial logistic models, the likelihood of responding “no” increased. As the intensity of value of crop production increases, the more likely farm debt is under-reported. As IGCFIK decreased, the likelihood of responding “no” increased in both the Direct and Guaranteed FLP multinomial logistic models, and this indicates the likelihood of under-reporting debt increases as gross cash farm income decreases. OL LOC loan operators were more likely to respond “no” in the Guaranteed FLP multinomial logistic model. Operators with an OL LOC loan are more likely to under-report their debt relative to operators with only guaranteed FO loans. Lastly, size as measured by gross cash farm income is important. Operators were less likely to respond “no” as IGCFIK increased in both the Direct and Guaranteed FLP models and more likely to refuse as IGCFIK increased in the Direct FLP model.

Overall for the “no” outcome, Direct FLP multinomial model operators are more likely to under-report their debt in the ARMS Phase III if they either have a lower total FSA debt outstanding balance, have a greater value of crop production relative to total production, or have a lower gross cash farm income. Guaranteed FLP multinomial model operators are more likely to under-report their debt in the ARMS Phase III if they have only an OL LOC loan, have a greater share of production from crops, have a lower gross cash farm income, are in survey year 2004, or are in survey year 2007. They are less likely to under-report their debt if they either have some college education, are SDA eligible, or are beginning farmer eligible.

6.2 Future Research

Future research could build upon this study by constructing a triple hurdle model to determine if those who respond “yes” accurately indicate their lender and loan amount. The first hurdle would be constructed the same as the current study. The second hurdle would look at those who responded “yes” to the “Owe Money” question to see if they accurately listed their lender. FSA is the correct lender to list for the direct FLP loans. However the lender with an FSA guarantee is the correct lender to list for the guaranteed FLP loans since the loans are originated and serviced by the lender with the guarantee, such as a commercial bank. The third hurdle would look at those respondents accurately reporting their FSA loan to see if they accurately reported their outstanding loan balance. The first hurdle and the current study partially addressed measurement errors for the “no” respondents and non-response errors for the refusals. The last hurdle would consider measurement errors for the “yes” respondents in greater detail. Those respondents who inaccurately list FSA as a lender is another type of measurement error that could be studied.

The results presented here are only for operations with FSA direct and/or guaranteed loans. However, the analysis could be expanded to credit providers such as the Farm Credit System or commercial banks to get a better understanding of the full magnitude of debt under-reporting. Additionally, research could be conducted in other sections of the ARMS survey to determine whether they are prone to inaccurate reporting as well. Also, future research could determine if current NASS imputation techniques have improved the estimation of DTOT from the ARMS Phase III, especially for those respondents indicating “no” on the “Owe Money” question in the Farm Debt section of the ARMS. Lastly, research could further look into reducing non-response by conducting experimental trials using different types of survey instruments to examine whether ARMS response can be improved.

References

Ahrendsen, Bruce L., and Ani L. Katchova. "Financial Ratio Analysis using ARMS Data." *Agricultural Finance Review* 72, no. 2 (2012): 262-272.

Ahrendsen, Bruce L., Bruce L. Dixon, Latisha A. Settlage, Steven R. Koenig, and Charles B. Dodson. "A Triple Hurdle Model of US Commercial Bank use of Guaranteed Operating Loans and Interest Assistance." *Agricultural Finance Review* 71, no. 3 (2011): 310-328.

Allison, Paul. "How Relevant is the Independence of Irrelevant Alternatives?" *Statistical Horizons*, October 8, 2012. Available at: http://statisticalhorizons.com/iia. Accessed July 20, 2015.

Banker, D., M. Morehart, E. O'Donoghue, and D. Milkove. *2009 Farm Business Summary Program*. (2010) USDA-ERS, Washington, DC, July 20, 2010: 17-18.

Blank, Steven C., and Danny Klinefelter. "Keeping ARMS Relevant: Increasing its Usability." *Agricultural Finance Review* 72, no. 2 (2012): 222-232.

Briggeman, Brian C., Steven R. Koenig, and Charles B. Moss. "US Farm Debt: the Role of ARMS." *Agricultural Finance Review* 72, no. 2 (2012): 254-261.

Cheng, Simon, and J. Scott Long. "Testing for IIA in the Multinomial Logit Model." *Sociological Methods & Research* 35, no. 4 (2007): 583-600.

Dixon, Bruce L., Bruce L. Ahrendsen, and Scott M. McCollum. *Models of FSA Guaranteed Loan Use Volume and Loss Claims Among Arkansas Commercial Banks*. Arkansas Agricultural Experiment Station, University of Arkansas Division of Agriculture, Fayetteville, Research Bulletin 962, (November 1999).

Dixon, Bruce L., Nan Ma, Bruce L. Ahrendsen, Latisha Settlage, and Jerome M. Stam. "Factors Affecting State-Level Chapter 12 Filing Rates: A Panel Date Model." *Emory Bankruptcy Developments Journal* 20, no. 2 (2004): 401-426.

Dixon, Bruce L., Bruce L. Ahrendsen, O. John Nwoha, Sandra J. Hamm, and Diana M. Danforth. "FSA Direct Farm Loan Program Graduation Rates and Reasons for Exiting." *Journal of Agricultural and Applied Economics* 39, no. 3 (2007): 471-487.

Earp, M.S., McCarthy, J. S., N. D. Shauer, and P. S. Kott. *Assessing the Effect of Calibration on Nonresponse Bias in the 2005 ARMS Phase III Sample Using Census 2002 Data*. US Department of Agriculture, National Agricultural Statistics Service, Research and Development Division, 2008.

Ellinger, Paul N., Bruce L. Ahrendsen, and Charles B. Moss. "Balance Sheet and Income Statement Issues in ARMS." *Agricultural Finance Review* 72, no. 2 (2012): 247-253.

Featherstone, Allen M., Timothy A. Park, and Jeremy G. Weber. "Keeping ARMS Relevant: Extracting Additional Information from ARMS." *Agricultural Finance Review* 72, no. 2 (2012): 233-246.

Farm Financial Standards Council. “Financial Guidelines for Agricultural Producers.” (2015). Available at: http://www.ffsc.org/index.php/guidelines/.

Gerling, Michael W., HoaiNam N. Tran, and Morgan S. Earp. "Categorizing Nonresponse in Phase III of the 2006 Agricultural Resource Management Survey in Washington State." *U.S. Department of Agriculture, National Agricultural Statistics Service, Research and Development Division* (September 2008): RDD Research Report Number RDD-08-08.

Greene, W.H. *Econometric Analysis, Seventh edition.* Upper Saddle River: Prentice Hall, Inc. 2011.

Hoppe, Robert A., and James M. MacDonald. *Updating the ERS Farm Typology*. EIB-110, U.S. Department of Agriculture, Economic Research Service (April 2013). Available at: http://www.ers.usda.gov/media/1070858/eib110.pdf. Last accessed October 19, 2015.

Katchova, Ani L. "Factors Affecting Farm Credit Use." *Agricultural Finance Review* 65, no. 2 (2005): 17-29.

Millar, Darcy, and Terry P. O’Connor. "Item Response Rates for the Agricultural Resource Management Survey III in 2006 and 2007." *U.S. Department of Agriculture, National Agricultural Statistics Service, Research and Development Division* (June 2012): RDD Research Report Number RDD-11-07.

Miller, Darcy, Michael Robbins, and Josh Habiger. "Examining the Challenges of Missing Data Analysis in Phase Three of the Agricultural Resource Management Survey." *JSM Proceedings*, Section on Survey Research Methods, Alexandria, VA: American Statistical Association (2010).

Nwoha, O. John, Bruce L. Ahrendsen, Bruce L. Dixon, Daniel M. Settlage, and Eddie C. Chavez. "FSA Direct Loan Targeting: Successful and Financially Necessary?" *Agricultural Finance Review* 67, no. 1 (2007): 35-53.

Settlage, Latisha A., Bruce L. Dixon, Bruce L. Ahrendsen, and Steven R. Koenig. *Models of Farm Service Agency Guaranteed Loan Loss Claim Rates in the U.S. for 1990-1998*. Arkansas Agricultural Experiment Station, Division of Agriculture, University of Arkansas, 2001.

U.S. Department of Agriculture, Economic Research Service. “ARMS Farm Financial and Crop Production Practices.” (April 23, 2015a). Available at: http://www.ers.usda.gov/data-products/arms-farm-financial-and-crop-production-practices.aspx. Last accessed: July 21, 2015.

U.S. Department of Agriculture, Economic Research Service. “Farm Income and Wealth Statistics: Farm Sector Balance Sheet, 2010-2015F.” (August 25, 2015b). Available at: http://www.ers.usda.gov/data-products/farm-income-and-wealth-statistics/balance-sheet.aspx. Last accessed: September 23, 2015).

U.S. Department of Agriculture, Farm Service Agency. “Programs.” (August 23, 2011). Available at: http://www.fsa.usda.gov/FSA/stateoffapp?mystate=wy&area=home&subject=prog&topic=landing. Last accessed July 21, 2015.

U.S. Department of Agriculture, Farm Service Agency. “Your Guide to FSA Farm Loans.” FSA-BR-01 (June 2012). Available at: http://www.fsa.usda.gov/Internet/FSA_File/fsa_br_01_web_booklet.pdf. Last accessed October 19, 2015.

U.S. Department of Agriculture, Farm Service Agency. “Farm Loan Programs.” (July 2015a). Available at: http://www.fsa.usda.gov/programs-and-services/farm-loan-programs/index. Last accessed July 21, 2015.

U.S. Department of Agriculture, Farm Service Agency. “FSA Handbook 1-FLP: General Program Administration.” (July 2015b). Available at: http://www.fsa.usda.gov/Internet/FSA_File/1-flp_r01_a119.pdf. Last accessed July 21, 2015.

U.S. Department of Agriculture, Farm Service Agency. “FSA Handbook 6-FLP: Special Programs.” (March 2015c). Available at: http://www.fsa.usda.gov/Internet/FSA_File/6-flp_r00_a22.pdf. Last accessed September 23, 2015.

U.S. Department of Agriculture, National Agricultural Statistics Service. “2001 Agricultural Resource Management Study Phase III Costs and Returns Report.” (2001). Available at: http://www.ers.usda.gov/datafiles/ARMS_Questionnaires__Manuals/2001/w_5e2001_5ecrr_5ephase3_20questionnaire_5eq_5efoh_1_.pdf. Last accessed: July 18, 2015.

U.S. Department of Agriculture, National Agricultural Statistics Service. “2002 Census Reveals an Increase in Women Principal Operators.” (March 2005). Available at: http://www.nass.usda.gov/Newsroom/2005/03_09_2005.asp. Last accessed: October 25, 2015.

U.S. Department of Agriculture, National Agricultural Statistics Service. “2009 Agricultural Resource Management Survey Phase III Survey Specifications.” Prepared by:

Environmental and Economic Surveys Section Survey Administration Branch. (October 2009a).

U.S. Department of Agriculture, National Agricultural Statistics Service. “2007 Census of Agriculture: United States Summary and State Data.” (February 2009b). Volume 1, Geographic Area Series, Part 51, AC-07-A-51. Available at: http://www.agcensus.usda.gov/Publications/2007/Full_Report/usv1.pdf. Last accessed: September 23, 2015.

Weber, Jeremy G., and Dawn Marie Clay. "Who Does Not Respond to the Agricultural Resource Management Survey and Does It Matter?" *American Journal of Agricultural Economics* 95, no. 3 (2013): 755-771.

Appendix A: Direct Binomial FLP Mean and Bootstrap Standard Error Summary Statistics								
Variables			**Outcomes**					
DEBT_PROX			**Yes/Ref**		**No**		**All**	
	Sample N	Weighted N	Mean	Btsp Std Err	Mean	Btsp Std Err	Mean	Btsp Std Err
Sample N			2,328 (86.3%)		368 (13.6%)		2,696 (100%)	
Weighted N			136,488 (87.1%)		20,204 (12.9%)		156,693 (100%)	
Direct FLP Variables								
FSADEBTTOTK ($1000)	2,696	156,693	133.157	4.659	80.256 d	8.550	126.336	4.149
INTRATE	2,696	156,693	0.047	0.001	0.046	0.002	0.047	0.001
PASTDUE_IND	2,696	156,693	0.049	0.009	0.095	0.051	0.055	0.011
FO_DIR	2,696	156,693	0.388	0.023	0.341	0.059	0.382	0.021
OL_DIR	2,696	156,693	0.213	0.018	0.276	0.057	0.221	0.018
EMEE_DIR	2,696	156,693	0.266	0.018	0.324	0.060	0.273	0.017
MULT_LN_DIR	2,696	156,693	0.133	0.015	0.059 d	0.018	0.123	0.013
MULT_PROG	2,696	156,693	0.164	0.014	0.078 d	0.020	0.153	0.012
Borrower Demographics								
OP_AGE	2,696	156,693	51.220	0.507	54.931 d	1.328	51.699	0.462
HS_EDUC	2,696	156,693	0.483	0.021	0.578	0.062	0.495	0.021
SC_EDUC	2,696	156,693	0.323	0.022	0.283	0.061	0.318	0.020
CGB_EDUC	2,696	156,693	0.194	0.016	0.139	0.031	0.187	0.014
OP_SDA_P	2,696	156,693	0.201	0.018	0.203	0.047	0.202	0.017
BF_ELIG	2,696	156,693	0.154	0.017	0.136	0.044	0.152	0.016
MARRIED	2,641	154,414	0.893	0.011	0.860	0.047	0.889	0.011

Appendix A: Direct Binomial FLP Mean and Bootstrap Standard Error Summary Statistics Cont.									
Variables	**Outcomes**								
DEBT_PROX			**Yes/Ref**		**No**			**All**	
	Sample N	Weighted N	Mean	Btsp Std Err	Mean		Btsp Std Err	Mean	Btsp Std Err
Sample N			2,328 (86.3%)		368 (13.6%)			2,696 (100%)	
Weighted N			136,488 (87.1%)		20,204 (12.9%)			156,693 (100%)	
Operation Characteristics									
HH_SIZE	2,612	153,689	3.053	0.069	2.631	d	0.126	2.998	0.062
RECORD_USE_MOT	1,775	121,302	0.667	0.025	0.378	d	0.080	0.633	0.025
RECORD_USE_SOT	1,775	121,302	0.162	0.020	0.199		0.060	0.166	0.019
RECORD_USE_NEV	1,775	121,302	0.171	0.019	0.423	d	0.084	0.200	0.020
RECORD_TP_YES	1,775	121,302	0.661	0.025	0.348	d	0.073	0.625	0.025
RECORD_TP_SOME	1,775	121,302	0.211	0.020	0.354	f	0.084	0.228	0.021
RECORD_TP_NONE	1,775	121,302	0.112	0.016	0.275	e	0.079	0.131	0.016
CROP_RATIO	2,696	156,693	0.463	0.020	0.542		0.054	0.473	0.018
Y2001	2,696	156,693	0.265	0.021	0.322		0.073	0.272	0.021
Y2004	2,696	156,693	0.324	0.021	0.291		0.055	0.320	0.020
Y2006	2,696	156,693	0.203	0.016	0.164		0.036	0.198	0.015
Y2007	2,696	156,693	0.209	0.017	0.223		0.044	0.210	0.016
Operation Financial Characteristics									
IGCFIK ($1000)	2,696	156,693	211.823	8.526	127.068	d	17.033	200.894	7.690
ETOTK ($1000)	2,696	156,693	163.365	6.440	98.822	d	12.707	155.042	5.812
INCFIK ($1000)	2,696	156,693	48.458	3.126	28.246	d	6.084	45.852	2.772
EARNEDK ($1000)	2,612	153,689	34.849	2.007	31.424		6.775	34.402	2.004
EFINTK ($1000)	2,696	156,693	16.035	0.656	5.215	d	1.155	14.640	0.592
INTFEE_REK ($1000)	2,696	156,693	10.515	0.511	4.105	d	1.105	9.688	0.470
INTFEE_NREK ($1000)	2,696	156,693	5.520	0.413	1.110	d	0.207	4.952	0.364

Appendix A: Direct Binomial FLP Mean and Bootstrap Standard Error Summary Statistics Cont.									
Variables	**Outcomes**								
DEBT_PROX			**Yes/Ref**		**No**			**All**	
	Sample N	Weighted N	Mean	Btsp Std Err	Mean		Btsp Std Err	Mean	Btsp Std Err
Sample N			2,328 (86.3%)		368 (13.6%)			2,696 (100%)	
Weighted N			136,488 (87.1%)		20,204 (12.9%)			156,693 (100%)	
Operation Financial Characteristics Cont.									
CAPEXP_TOTK ($1000)	1,775	121,302	5.121	0.675	2.802	e	0.779	4.851	0.604
NETWK ($1000)	2,696	156,693	707.861	27.505	702.269		63.133	707.140	25.816
ATOTK ($1000)	2,696	156,693	926.513	30.928	705.585	d	63.488	898.026	28.771
ACTOTK ($1000)	2,696	156,693	111.848	6.542	52.720	d	9.778	104.224	5.819
DTOTK ($1000)	2,696	156,693	218.651	8.218	3.316	d	0.524	190.886	7.284
LCTOTK ($1000)	2,696	156,693	66.601	3.253	3.316	d	0.524	58.441	2.853
Financial Ratio Variables									
Liquidity									
NWC_EXPENSE_RATIO (%)	2,696	156,693	35.538	0.093	63.624	d	0.112	39.160	0.083
CR	2,687	156,317	5.597	1.201	38.032	d	8.854	9.745	1.546
Solvency									
DAR (%)	2,696	156,693	29.060	1.176	0.709	d	0.150	25.405	1.078
Profitability									
ROA (%)	2,696	156,693	-0.758	0.609	-2.853		1.787	-1.028	0.590
OPM (%)	2,695	156,692	-34.851	5.809	-67.763		19.751	-39.094	6.001

Appendix A: Direct Binomial FLP Mean and Bootstrap Standard Error Summary Statistics Cont.									
Variables				**Outcomes**					
DEBT_PROX			**Yes/Ref**		**No**			**All**	
	Sample N	Weighted N	Mean	Btsp Std Err	Mean		Btsp Std Err	Mean	Btsp Std Err
Sample N			2,328 (86.3%)		368 (13.6%)			2,696 (100%)	
Weighted N			136,488 (87.1%)		20,204 (12.9%)			156,693 (100%)	
Financial Ratio Variables Cont.									
Debt Repayment									
DRCU (%)	2,696	156,693	4.433	1.505	0.050	d	0.285	3.868	1.307
Efficiency									
DEPER	2,695	156,692	0.197	0.027	0.131		0.034	0.188	0.023
OER (%)	2,695	156,692	123.522	10.788	111.663		7.676	121.993	9.454
Source: Merged FSA-ARMS data set (2001, 2004, 2006, and 2007)									
Footnote signifying significance levels for the difference in means. Yes/Refusal-No: d ($p < 0.01$); e ($p < 0.05$); f ($p < 0.10$)									

Appendix B: Guaranteed Binomial FLP Mean and Bootstrap Standard Error Summary Statistics									
Variables	**Outcomes**								
DEBT_PROX			**Yes/Ref**		**No**			**All**	
	Sample N	Weighted N	Mean	Btsp Std Err	Mean		Btsp Std Err	Mean	Btsp Std Err
Sample N			2,457 (90.5%)		257 (9.5%)			2,714 (100%)	
Weighted N			82,683 (90.1%)		9,087 (9.9%)			91,771 (100%)	
Guaranteed FLP Variables									
FSADEBTTOTK ($1000)	2,714	91,770	271.843	8.100	272.925		36.102	271.950	8.035
BORR_GUAR_INT_RATE (%)	2,714	91,770	7.281	0.111	7.679		0.277	7.320	0.101
FO_GTE	2,714	91,770	0.421	0.022	0.394		0.075	0.418	0.021
OL_GTE	2,714	91,770	0.122	0.015	0.069	f	0.024	0.117	0.014
OL_LOC_GTE	2,714	91,770	0.087	0.009	0.187	e	0.042	0.097	0.009
MULT_LN_GTE	2,714	91,770	0.369	0.020	0.347		0.081	0.367	0.019
MULT_PROG	2,714	91,770	0.277	0.020	0.173	e	0.047	0.267	0.019
Borrower Demographics									
OP_AGE	2,714	91,770	49.555	0.485	48.739		1.006	49.475	0.451
HS_EDUC	2,714	91,770	0.454	0.021	0.607	f	0.083	0.469	0.020
SC_EDUC	2,714	91,770	0.319	0.020	0.256		0.089	0.313	0.020
CGB_EDUC	2,714	91,770	0.225	0.018	0.136	e	0.038	0.216	0.017
OP_SDA_P	2,714	91,770	0.105	0.013	0.049	d	0.014	0.099	0.012
BF_ELIG	2,714	91,770	0.144	0.014	0.141		0.048	0.143	0.014
MARRIED	2,663	89,695	0.896	0.013	0.922		0.028	0.899	0.012

Appendix B: Guaranteed Binomial FLP Mean and Bootstrap Standard Error Summary Statistics Cont.								
Variables	**Outcomes**							
DEBT_PROX			**Yes/Ref**		**No**		**All**	
	Sample N	Weighted N	Mean	Btsp Std Err	Mean	Btsp Std Err	Mean	Btsp Std Err
Sample N			2,457 (90.5%)		257 (9.5%)		2,714 (100%)	
Weighted N			82,683 (90.1%)		9,087 (9.9%)		91,771 (100%)	
Operation Characteristics								
HH_SIZE	2,646	89,384	3.174	0.064	3.089	0.149	3.166	0.059
RECORD_USE_MOT	1,758	66,204	0.701	0.024	0.403 d	0.106	0.675	0.025
RECORD_USE_SOT	1,758	66,204	0.139	0.019	0.280	0.126	0.151	0.022
RECORD_USE_NEV	1,758	66,204	0.158	0.018	0.316	0.105	0.172	0.018
RECORD_TP_YES	1,758	66,204	0.711	0.023	0.687	0.098	0.709	0.022
RECORD_TP_SOME	1,758	66,204	0.165	0.017	0.100	0.047	0.159	0.017
RECORD_TP_NONE	1,758	66,204	0.106	0.015	0.173	0.075	0.112	0.015
CROP_RATIO	2,714	91,770	0.519	0.017	0.715 d	0.054	0.538	0.017
Y2001	2,714	91,770	0.206	0.020	0.090 e	0.045	0.195	0.019
Y2004	2,714	91,770	0.295	0.018	0.348	0.083	0.300	0.020
Y2006	2,714	91,770	0.256	0.018	0.280	0.064	0.258	0.017
Y2007	2,714	91,770	0.241	0.016	0.280	0.056	0.244	0.015
Operation Financial Characteristics								
IGCFIK ($1000)	2,714	91,770	365.661	14.612	242.727 d	28.996	353.487	14.038
ETOTK ($1000)	2,714	91,770	283.575	10.273	176.871 d	20.701	273.009	9.922
INCFIK ($1000)	2,714	91,770	82.085	6.338	65.855	13.194	80.478	5.979
EARNEDK ($1000)	2,646	89,384	32.889	2.173	38.305	6.724	33.437	2.067
EFINTK ($1000)	2,714	91,770	28.150	1.143	8.017	1.355	26.156	1.098
INTFEE_REK ($1000)	2,714	91,770	18.787	0.762	5.030 d	1.009	17.425	0.738
INTFEE_NREK ($1000)	2,714	91,770	9.362	0.731	2.987 d	0.618	8.731	0.670

Appendix B: Guaranteed Binomial FLP Mean and Bootstrap Standard Error Summary Statistics Cont.									
Variables	**Outcomes**								
DEBT_PROX			**Yes/Ref**		**No**			**All**	
	Sample N	Weighted N	Mean	Btsp Std Err	Mean		Btsp Std Err	Mean	Btsp Std Err
Sample N			2,457 (90.5%)		257 (9.5%)			2,714 (100%)	
Weighted N			82,683 (90.1%)		9,087 (9.9%)			91,771 (100%)	
Operation Financial Characteristics Cont.									
CAPEXP_TOTK ($1000)	1,758	66,204	9.824	1.306	2.340	d	0.869	9.178	1.200
NETWK ($1000)	2,714	91,770	872.961	35.696	825.610		104.783	868.272	33.786
ATOTK ($1000)	2,714	91,770	1247.24	41.524	831.992	d	105.134	1206.12	39.231
ACTOTK ($1000)	2,714	91,770	180.274	8.884	99.835	d	19.439	172.308	8.380
DTOTK ($1000)	2,714	91,770	374.282	13.221	6.382	d	0.992	337.851	12.858
LCTOTK ($1000)	2,714	91,770	121.078	5.973	6.382	d	0.992	109.720	5.494
Financial Ratio Variables									
Liquidity									
NWC_EXPENSE_RATIO (%)	2,714	91,770	15.911	0.055	95.013	d	0.165	23.744	0.052
CR	2,709	91,589	3.485	0.421	83.850	d	24.988	11.300	2.305
Solvency									
DAR (%)	2,714	91,770	35.417	0.986	1.925	d	0.549	32.100	1.023
Profitability									
ROA (%)	2,714	91,770	2.001	0.706	-6.549		7.452	1.154	0.951
OPM (%)	2,714	91,770	-14.418	3.733	-4.069		6.986	-13.393	3.384

Appendix B: Guaranteed Binomial FLP Mean and Bootstrap Standard Error Summary Statistics Cont.								
Variables	**Outcomes**							
DEBT_PROX			**Yes/Ref**		**No**		**All**	
	Sample N	Weighted N	Mean	Btsp Std Err	Mean	Btsp Std Err	Mean	Btsp Std Err
Sample N			2,457 (90.5%)		257 (9.5%)		2,714 (100%)	
Weighted N			82,683 (90.1%)		9,087 (9.9%)		91,771 (100%)	
Financial Ratio Variables Cont.								
Debt Repayment								
DRCU (%)	2,714	91,770	8.456	2.514	-0.542 d	0.607	7.565	2.282
Efficiency								
DEPER	2,714	91,770	0.137	0.012	0.070 d	0.016	0.130	0.011
OER (%)	2,714	91,770	91.989	3.323	108.120	19.028	93.586	3.538
Source: Merged FSA-ARMS data set (2001, 2004, 2006, and 2007)								
Footnote signifying significance levels for the difference in means. Yes/Refusal-No: d (p < 0.01); e (p < 0.05); f (p < 0.10)								

Appendix C: Average Weighted FSA Direct FLP Debt per Borrower			
FSA Debt Variables	**Level**		**All Borrowers**
Weighted N			172,789
Sample N			2,894
Total FSA DIR FO Debt	Borrower	Mean	$75,873
Total FSA DIR OL Debt	Borrower	Mean	$71,469
Total FSA DIR EM Debt	Borrower	Mean	$51,117
Total DIR EE Debt	Borrower	Mean	$22,257
Total FSA DIR Debt	Operation	Mean	$92,480
Source: Merged FSA-ARMS data set (2001, 2004, 2006, and 2007)			

Appendix D: Average Weighted FSA Guaranteed FLP Debt per Borrower			
FSA Debt Variables	**Level**		**All Borrowers**
N			99,177
Sample N			2,843
Total FSA GTE FO Debt	Borrower	Mean	$209,529
Total FSA GTE OL Debt	Borrower	Mean	$133,954
Total FSA GTE OL LOC Debt	Borrower	Mean	$140,346
Total FSA GTE Debt	Operation	Mean	$241,515
Source: Merged FSA-ARMS data set (2001, 2004, 2006, and 2007)			