United States
Department of
Agriculture

National
Agricultural
Statistics
Service

Research and
Development Division
Washington DC 20250

RDD Research Report
RDD-09-06
October 2009

Using Respondent Prediction Models to Improve Efficiency of Incentive Allocation

Morgan S. Earp
Jaki S. McCarthy

EXECUTIVE SUMMARY

With decreasing response rates, the use of incentives to encourage response is becoming more common; however, propensity for some survey units to respond may be unaffected by incentives. The variation in incentive utility results in wasted tax payer dollars.

One way to improve efficiency is to use auxiliary data to identify and screen out persons likely to respond without incentives. Historical data from cases where incentives were offered uniformly may be used to model respondent characteristics before and after incentives, and thus identify consistent respondent traits. Once consistent respondent characteristics are identified, future samples may be scored to screen out likely respondents (with or without an incentive) prior to incentive allocation.

This paper discusses the use of data mining to identify likely Agricultural Resource Management Survey (ARMS) respondents. In an effort to increase response rates, the National Agricultural Statistics Service (NASS) began experimenting with a monetary incentive in 2004. Follow-up assessments of the monetary incentive in 2005 demonstrated that ATM cash cards are beneficial in increasing ARMS Phase III (ARMS III) response rates and decreasing survey costs; however, it is unknown which sampled units would have responded without the incentive. A series of models were built using 2002 Census of Agriculture data to predict several years of ARMS III sample respondents before (2003-2004) incentives were introduced. These models were applied to the years after incentives were introduced (2005-2007) to confirm that they continued to identify likely respondents. This approach allows NASS to assess the consistency of respondent characteristics before and after incentives are introduced.

Currently, all ARMS III mail version sampled units receive a $20 Automatic Teller Machine (ATM) monetary incentive. By using respondent prediction models NASS may be able to flag persons likely to respond given no incentive and in turn use that money more effectively elsewhere.

RECOMMENDATIONS

1. Score the 2009 ARMS III Core sample using criteria specified in the five models. Records meeting the criteria specified on one or more models will be flagged as likely respondents.

2. Conduct a similar study using ARMS III Core 2005 training, 2006 validation, and ultimately 2007 test data to flag 2010 mail nonrespondents.

3. Contact flagged and confirmed ARMS III Core 2007 mail nonrespondents for cognitive interviews in order to identify alternative incentives for use in 2010.

4. Randomly divide flagged 2010 mail nonrespondents into three groups: 1) a control group receiving no incentive, 2) a treatment group receiving a $20 ATM card incentive, and 3) a treatment group receiving an alternative incentive identified via cognitive interviews, to determine if the identified alternative incentive is more effective for the given mail nonresponse group.

Using Respondent Prediction Models to Improve Efficiency of Incentive Allocation

Morgan Earp and Jaki McCarthy[1]

Abstract

In an effort to increase response rates, the National Agricultural Statistics Service (NASS) began experimenting with monetary incentives in 2004. Follow-up assessments of the monetary incentive in 2005 demonstrated that ATM cash cards are beneficial in increasing Agricultural Resource Management Survey Phase III (ARMS III) response rates and decreasing survey costs; however, it is unknown which sampled units would have responded without the incentive. This paper discusses the use of data mining to identify likely ARMS III respondents. A series of models were built using 2002 Census of Agriculture data to predict several years of ARMS III sample respondents before (2003-2004) incentives were introduced. These models were applied to the years after incentives were introduced (2005-2007) to confirm that they continued to identify likely respondents. The respondent prediction models discussed in this report enable NASS to flag persons likely to respond given no incentive. Providing incentives to these respondents requires substantial costs, but likely does not increase overall response rates. In addition, if providing them incentives does increase response rates, it may increase them in such a way that NASS estimates are further biased if only more of the same type of operations opt to respond.

Key Words: Nonresponse; response rate; bias; incentives.

[1] Morgan Earp is a survey and mathematical statistician in USDA's National Agricultural Statistics Service (NASS) – Research and Development Division (RDD), located in Room 305, 3251 Old Lee Highway, Fairfax, VA 22030. Jaki McCarthy, the Chief Cognitive Research Methodologist with USDA/NASS/RDD, provided assistance with this research.

1. INTRODUCTION

On February 2, 2006, a panel was formed within the National Research Council (NRC) at the recommendation of the Committee on National Statistics, Division of Behavioral and Social Sciences and Education to review the United States Department of Agriculture's (USDA) Agricultural Resource Management Survey (ARMS). Two years later, the NRC released a report entitled *Understanding American Agriculture; Challenges for the Agricultural Resource Management Survey* (2008). Section 6 of the NRC ARMS review specifically addresses nonresponse, imputation, and estimation. The review recommends that the USDA's National Agricultural Statistics Service's (NASS) Research and Development Division (RDD) explore characteristics of nonrespondents, as well as the relationship between incentives and nonresponse bias:

> Recommendation 6.3: The nature of ARMS nonresponse bias should be a key focus of the research and development program the panel recommends. This research and development program should focus, initially, on understanding the characteristics of nonrespondents.
>
> Recommendation 6.4: The research and development program should continue NASS's work on both public relations and incentives, and it should do so with a focus on nonresponse bias, not simple nonresponse rates.

The ARMS is conducted in three phases. Phase I screens for potential samples for Phases II and III. Phase II collects data on cropping practices and agricultural chemical usage, while Phase III collects detailed economic information about the agricultural operation, as well as the operator's household. ARMS data are used by farm organizations, commodity groups, agribusiness, Congress, State Departments of Agriculture, and the USDA. The USDA uses ARMS data to evaluate the financial performance of farms and ranches, which influence agricultural policy decisions. The Department also uses the ARMS Phase III (ARMS III) data for objective evaluation of critical issues related to agriculture and the rural economy; therefore, it is essential that measures be taken to minimize bias, especially for Phase III.

1.1 *Problem*

In an effort to increase response rates, the USDA's NASS began experimenting with monetary incentives in the 2004 ARMS III (Beckler, Ott, & Horvath, 2005). Follow-up assessments of the monetary incentive in the 2005 ARMS III, where operations were mailed a pre-survey letter with a pre-paid $20 ATM card prior to the survey, demonstrated that ATM cash cards are beneficial in increasing response rates and decreasing survey costs for ARMS questionnaires mailed to respondents (McCarthy, Beckler, & Ott, 2006); however, it is unknown how much the effectiveness of the monetary incentive varies across sampled entities. Are certain operations likely to respond regardless of incentives? Are certain operations more likely to respond via mail given a monetary incentive? Lastly, are there operations more likely to respond via mail given an alternative incentive? Without a basic understanding of operation characteristics, specifically those unique to ARMS III respondents versus nonrespondents, it is unclear whether incentives either vary in effectiveness or are distributed efficiently. Furthermore, offering incentives may increase response rates, but it does not necessarily decrease bias. There are four possible outcomes when giving incentives: 1) if persons are already more apt to respond and begin responding at a higher rate given incentives, we exacerbate response bias; 2) if persons previously responding stop responding, nonresponse bias may be increased; 3) if prior nonrespondents respond, we may reduce nonresponse bias; and 4) if prior nonrespondents continue not responding bias may continue (Figure 1). Note that only one of these outcomes results in a reduction of bias.

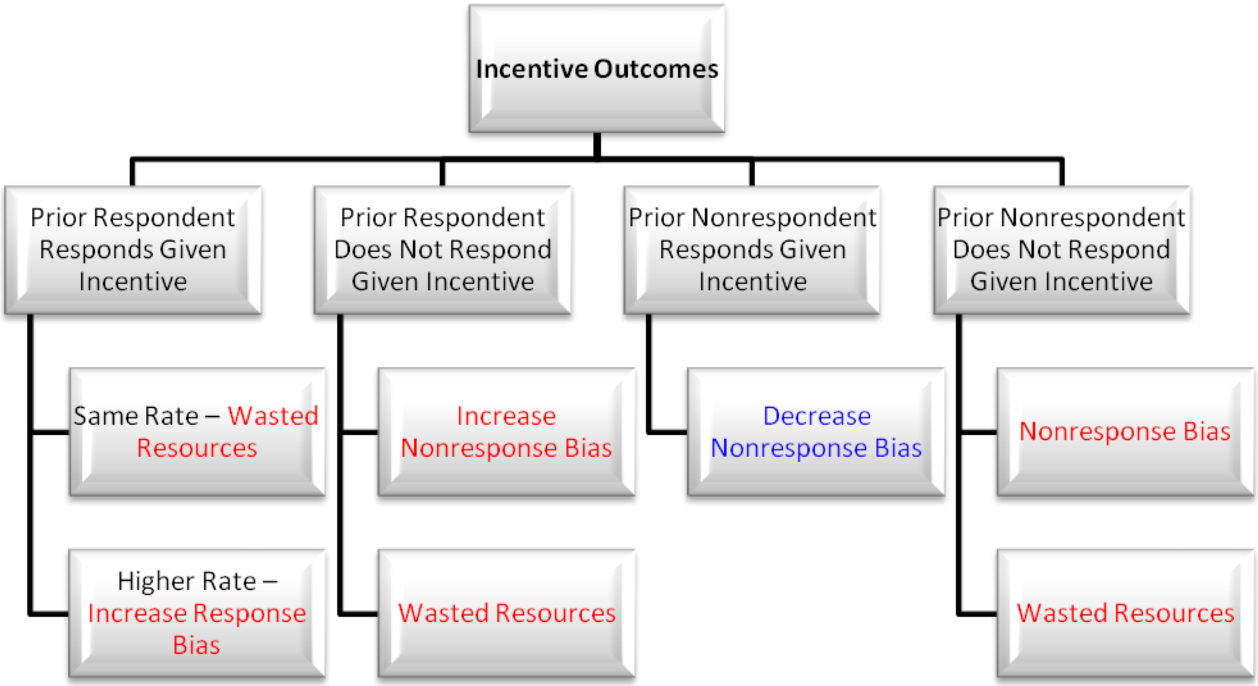

Figure 1. Incentive Outcomes

1.2 *Purpose*

This pilot study explores and identifies characteristics of 2003 through 2007 ARMS III respondents by testing whether certain operations are likely to respond regardless of incentives - the rationale being that the ATM monetary incentive may vary in effectiveness based on operation characteristics, and thus incentives may be unnecessary when persons are already apt to respond. By flagging persons likely to respond given no incentive, NASS may be able to decrease response bias and survey costs, and better allocate incentive funds toward those least likely to respond. Ultimately, this study aims to demonstrate a method for identifying likely respondents.

1.21 <u>Research Questions:</u>

What are the characteristics of operations that are likely to respond to the ARMS III Core Version regardless of incentives?

2. METHOD

In order to identify characteristics of ARMS III respondents, 2002 Census of Agriculture data were matched to sampled operations (both respondents and nonrespondents) in the 2003-2007 ARMS III Core Version. The research included data on various operation characteristics from the Census of Agriculture that were recommended by both NASS's Chief Cognitive Research

Methodologist and Chief Research Statistician. These operation characteristics were used to predict respondents in the 2003-2007 ARMS III using classification trees.

2.1 *Procedure*

Classification or decision trees (these terms are used interchangeably) were used to identify characteristics of ARMS III Core Version respondents. Classification trees model relationships with a categorical outcome (respondent or nonrespondent) using a tree-like structure.

In this type of analysis, the full data were comprised of the 2002 Census of Agriculture data for the 2003-2007 ARMS III Core form sample. The Core Version of the ARMS survey is the part of the sample that has included a $20 ATM Card mailing as an incentive and is the only part of the ARMS survey in which questionnaires are mailed to respondents. If a response is not received by mail, an enumerator will attempt to complete a face to face interview. The Core Version is only used in the 15 estimating states, which include the 15 leading cash receipts states (Arkansas, California, Florida, Georgia, Illinois, Indiana, Iowa, Kansas, Minnesota, Missouri, Nebraska, North Carolina, Texas, Washington, and Wisconsin). Maryland and Delaware were given special permission to use the Core Version in place of face-to-face interviews in 2003 and 2005 due to an avian influenza epidemic. Their data are also included.

The data were broken into subsets by year to be used as the training (2003), validation (2004), and test (2005-2007) sets. The training dataset was used to construct the initial tree model that identified subsets of records that responded at a higher rate than the overall sample. This model was applied to the validation dataset in order to prevent generating a model for the training data that would not fit other data or that would be unreliable (i.e. overfitted). The validation data were used when pruning the initial tree to generate the final model. Finally, the test data were used to evaluate the model's performance on independent data not used in the creation of the model. In this case, the initial tree was constructed using data from years immediately prior to the use of any incentives in the ARMS III survey. This model was applied to the test data, which consisted of the years following the introduction of incentives, to identify groups of records that responded consistently regardless of the use of incentives.

A decision tree model is constructed by segmenting the data through the application of a series of simple rules. Each rule assigns an observation to a subsegment based on the value of one input variable. One rule is applied after another, resulting in a hierarchy of segments within segments. The rules are chosen to maximally separate the subsegments with respect to the target variable. Thus, the rule selects both the variable and the best breakpoint to maximally separate the resulting subgroups. Variables may appear multiple times throughout the tree for further segmentation. The resulting hierarchy is called a tree, and each segment is called a node. The original segment contains the entire data set and is called the root node of the tree. A node with all its successors is termed a branch of the node that created it. The final nodes are called leaves. In our analysis, we are ultimately interested in the leaves that contain a higher proportion of records with the target (response).

Decision trees describe subsets of data and are constructed without any theoretical guidance. Variables are chosen to maximally separate the subsegments, so only one or a few similar correlated variables (which individually might be related to the target) may appear in the tree. There are several alternative methods for constructing decision trees. For the purposes of this report, trees were grown using the chi-square approach available in SAS Enterprise Miner, which is similar to the chi-square automatic interaction detection (CHAID) algorithm (deVille, 2006).

There are multiple stopping criteria used to decide how large to grow a decision tree. Generally, trees were pruned so leaves represented at least 500 records or when adding additional leaves did not markedly improve the overall misclassification rates of the tree as a whole. All trees had similar misclassification rates for the training and validation datasets used to grow the trees and for the test data used to verify reliability of the trees after construction.

For the purposes of this study, the target was ARMS III Core Version response. Operations responding to ARMS III were marked with a "1" and those not responding with a "0" in a new survey response target variable. A decision tree considers all input variables (independent variables) and grows branches using input variables that demonstrate significant relationships with the target, while also considering interaction effects between the various inputs. The classification trees described in this study explored the relationship between operation characteristics and survey response.

Trees were grown using the 2003 sample to train the models and identify significant splits. Trees were pruned and validated by assessing the average squared error of the model using data from the 2004 sample. Reliability of the trees was tested and compared using the 2005, 2006, and 2007 samples. It is assumed that characteristics consistently associated with significantly higher response rates from 2003 through 2007 are invariant to the effect of incentives, since no incentives were given in 2003 and 2004, but were in 2005, 2006, and 2007.

In a typical classification tree approach, the best initial splitting variable would be chosen and a single model built. Many models can be built using a single dataset, with increasing misclassification rates. Each model will identify different (but possibly overlapping) subgroups.

For this project, five separate models were built using each of the top five best initial splitting variables. Each of these five models was grown by forcing the primary split on a different one of the five potential splitting variables. All variables were available for each of the models and subsequent splits were determined automatically by the software. The groups of records with highest response rates were selected from each model. Each model identified unique subsets of respondents based on varying initial splits; furthermore, significance levels used to evaluate the initial splits were based solely on the training data. By creating several complementary models, we identified more respondents than we could have using a single model, and we were able to reevaluate the strength of the models in comparison to one another using the training, validation, and test data.

The significance of potential splitting variables was assessed using the LogWorth statistic, which measures how well a given input variable measures the target. All five decision trees were comparable, and thus, were explored for two reasons:

1) The LogWorth of initial split variables is calculated using only the training data (2003), so although it may be highly significant in the training phase, it may prove unreliable using the validation data (2004) or the test data (2005-2007). Therefore, competing models may in fact produce better results when tested over time.

2) The characteristics identified in a given tree vary given the variable used in the initial split; therefore, each tree is capable of identifying unique subsets of respondents. Predicted response probabilities generated using the five models are available for scoring of future ARMS III samples.

2.2 *Data*

Data from the 2002 Census of Agriculture were matched to both respondents and nonrespondents in the ARMS III Core Version sample. Associated characteristics of respondents were identified prior to incentives being used in the ARMS III (2003-2004). There were 28,372 records with available 2002 Census data in this set. These models were used to flag likely respondents after incentives were used in the ARMS III Core Version (2005-2007). There were 40,487 records with matching census data in this second data set.

In order to ensure reliability of results, data were partitioned into three groups: training, validation, and test. Training data were used to grow trees. Validation data were used to prune trees when classification became unreliable. Test data were used to compare trees (models) in terms of gain rates and reliability over time.

The respondent characteristic models data for the ARMS III Core Version sample were identified using the available 2002 census data. The respondent characteristic models were trained using the matched 2003 sample (n = 14,193), validated using the matched 2004 sample (n = 14,179), and tested using the matched 2005 (n = 14,027), 2006 (n = 13,614), and 2007 (n = 12,846) samples. Census data were matched and available for most records. See Appendix A for a comparison of match rates by year, version, and respondents versus nonrespondents.

2.3 *Variables*

Eighty-four variables from the 2002 Census were selected and used to explore respondent characteristics. The variables included descriptive information about the operation such as its size, the type of commodities produced, its location, etc. as well as information about the principal operator, such as the operator's race, gender, number of days worked off the farm, etc. The full list of variables used is shown in Table 1.

Table 1: 2002 Census Operational Characteristic Variables

Varname	Description
Census 2002 Variables	
K43	Acres of Land Owned
K44	Acres Rented - Census
K45	Acres of Land Rented to Others
K46	Total Acres Operated - Census
K121	Tenth-Acres of Fruits and Nuts
K395	Tenth-Acres of Cantaloupe Harvested
K423	Tenth-Acres of Honeydew Harvested
K473	Tenth-Acres of Watermelon Harvested
K475	Tenth-Acres of Other Vegetables
K684	Government Payments - Census
K787	Acres of Cropland Harvested
K788	Acres of Cropland Used for Pasture
K790	Acres of Cropland for Which All Acres Failed
K791	Acres of Cropland in Summer Fallow
K794	Acres of Woodland Pasture
K795	Acres of Woodland Not in Pasture
K796	Acres of Permanent Pasture & Rangeland
K797	Acres of All Other Land
K798	Total Acres - Reported
K803	Total Cattle and Calf Inventory
K815	Total Hog and Pig Inventory
K923	Principal Operator- Residence on Place
K925	Operator's Age - Census
K926	Principal Operator- Sex
K927	Principal Operator- Spanish Origin
K928	Principal Operator- Principal Occupation
K929	Principal Operator Days Worked off Farm
K930	Principal Operator- Year Began Operation
K941	Hired Workers Less Than 150 days
K942	Hired Workers Greater Than or Equal to 150 days
K943	Machinery and Equipment Value - Census
K1021	Acres of all Hay and Forage Harvested
K1022	Acres of all Irrigated Hay and Forage Harvested
K1043	Sum of Tenth-Acres of Berries
K1050	Agriculture on Indian Reservations Y/N
K1062	Acres of Cropland Idle or Used for Cover Crops
K1069	Acres of Certified Organic Farming
K1080	Possible Duplicate Y/N
K1086	Have Other Farm Y/N
K1314	Total $ - Under Production Contract
K1347	Total Sales – Not Under Production Contract
K1367	Total Sales – Under Production Contract
K1501	Fertilizer Expenses - Census
K1502	Chemical Purchases - Census
K1503	Seed Expenses - Census
K1504	Operator's (+LL) Breeding Livestock Purchased
K1505	Operator's (+LL) All Other Livestock Purchased
K1567	Partnership Registered under State Law? Y/N
K1568	Any Fertilizer or Chemicals Y/N
K1569	Acres on Which Manure Was Applied
K1573	Any Migrant Workers Y/N
K1574	Number of Women Operators
K1575	Number of Operators
K1576	Any Hired Manager? Y/N
K1577	Principal Operator- Number of Persons Living in Household
K1578	Percent of Household Income from Operation
K1602	Computer Used Y/N
K1603	Internet Access Y/N
K1608	HHs Sharing in Net Farm Income
K1701	Principal Operator – Race, White
K1702	Principal Operator – Race, Black
K1703	Principal Operator – Race, American Indian
K1704	Principal Operator – Race, Native Hawaiian or Pacific Islander
K1705	Principal Operator – Race, Asian
NASS_STATE_	State
NGFS	Nursery Indicator
OAQ	Aquaculture Indicator
OVMHA	Tenth-Acres of (Other Vegetables + Honeydew+ Watermelons + Cantaloupe) Harvested
RCROP	Sum of Cropland harvested
REXP	Reported sum of expenditures
RPLTINV	Sum of poultry Inventory Data
RSUMFA	Sum of all Reported Fruit Acres
SHEP	Sheep and Lamb indicator
TCL	Cropland Acres - Census
TCTA	Total Citrus Acres
TENURE	Operation Farm Tenure (1=full owner, 2=part owner, 3=tenant)
TFPE	Total Production Expenses - Census
TOTOTLVK	Matching Variable for Other Livestock Animals
TVPG	Total Value of Products Sold + Government Payments
VEGA	Sum Acres of Vegetables
operatorrace_census	Operator's Race - Census
cropexp_census	Crop Expenses - Census
totalsales	Total Sales - Census
livestpurch_census	Livestock Purchases - Census

[2]

[2] See the *PRISM II Code Book* (United States.Department of Agriculture, 2008) for variable descriptions.

3. RESULTS

Enterprise Miner 5.2 identified five top competing models (initial tree splits). The five initial split variables used to build the tree models were:

1) Acres of Cropland Harvested[3] less than 211 acres versus Acres of Cropland Harvested equal to or greater than 211 acres (Figure 2);
2) Acres of Cropland less than 354 acres versus Acres of Cropland equal to or greater than 354 acres (Figure 3);
3) Total Sales Not Under Production Contract (NUPC) less than $43,551 versus greater than or equal to $43,551 (Figure 4);
4) Sum of Cropland Harvested[4] less than 211 acres versus greater than or equal to 211 acres (Figure 5); and
5) State: California, Delaware, Florida, Illinois, Indiana, Iowa, Kansas, Minnesota, Missouri, Nebraska, Washington, and Wisconsin versus Arkansas, Georgia, North Carolina, and Texas (Figure 6).

3.1 *Models*

In order to reduce the risk of future misclassification, only the subgroups that demonstrated substantial gains (response rates > 80 percent), and thus minimal misclassification rates in 2003 and 2004, were selected to design future scoring criteria. This approach resulted in one subgroup being selected from each model.

3.1.1 Model One: Acres of Cropland Harvested

The first model split used Acres of Cropland Harvested less than 211 acres versus Acres of Cropland Harvested equal to or greater than 211 acres (Figure 2). This model identified the 13,919 operations with less than 211 Acres of Cropland Harvested in Arkansas, Georgia, Illinois, North Carolina, and Texas, and Total Sales Not under Production Contract less than $522,250, exhibiting response rates between 78 percent and 83 percent (2003-2007).

3.1.2 Model Two: Acres of Cropland

The second model split using Acres of Cropland less than 354 acres versus Acres of Cropland greater than or equal to 354 acres (Figure 3). This model identified the 10,910 operations with less than 354 Acres of Cropland in Arkansas, Georgia, Illinois, North Carolina, and Texas with Total Sales Not under Production Contract less than $38,074, exhibiting response rates between 80 percent and 84 percent (2003 - 2007).

3.1.3 Model Three: Total Sales Not Under Production Contract (NUPC)

The third model split using Total Sales Not under Production Contract (NUPC) less than $43,551 versus greater than or equal to $43,551 (Figure 4). This model identified the 15,206 operations with Total Sales Not under Production Contract less than $43,551 in Arkansas, Florida, Georgia, Illinois, Indiana, North Carolina, and Texas, exhibiting response rates between 80 percent and 84 percent (2003 - 2007).

[3] Acres of Cropland Harvested are reported by the respondent.
[4] Sum of Cropland Harvested is calculated by summing reported individual crop harvested acres.

3.1.4 Model Four: Sum of Cropland Harvested

The fourth model split using Sum of Cropland Harvested less than 211 acres versus greater than or equal to 211 acres (Figure 5). This model identified the 14,678 operations with Sum of Cropland Harvested less than 211 acres, in Arkansas, Georgia, Illinois, North Carolina, and Texas, exhibiting response rates between 77 percent and 82 percent (2003-2007).

3.1.5 Model Five: State

The fifth model split using two state groupings: 1) California, Delaware, Florida, Illinois, Indiana, Iowa, Kansas, Maryland, Minnesota, Missouri, Nebraska, Washington, and Wisconsin; and 2) Arkansas, Georgia, North Carolina, and Texas (Figure 6). This model identified the 17,181 operations in Arkansas, Georgia, North Carolina, and Texas with Total Acres Operated less than 201 acres, exhibiting responses rates between 75 percent and 81 percent (2003-2007).

3.1.6 All Models: Model One, Model Two, Model Three, Model Four, and Model Five

The "All Models" indicator (applied to any record flagged as a potential respondent by Model One, Model Two, Model Three, Model Four, and Model Five) identified the 9,272 operations that appeared in each of the model nodes above. These operations had response rates between 81 percent and 85 percent (2003-2007).

3.1.7 Any Model: Model One, Model Two, Model Three, Model Four, or Model Five

The "Any Model" indicator (applied to any record flagged as a potential respondent by Model One, Model Two, Model Three, Model Four, or Model Five) identified 22,603 operations that appeared in the likely respondent nodes of at least one of the above models. These operations exhibited response rates between 76 percent and 81 percent.

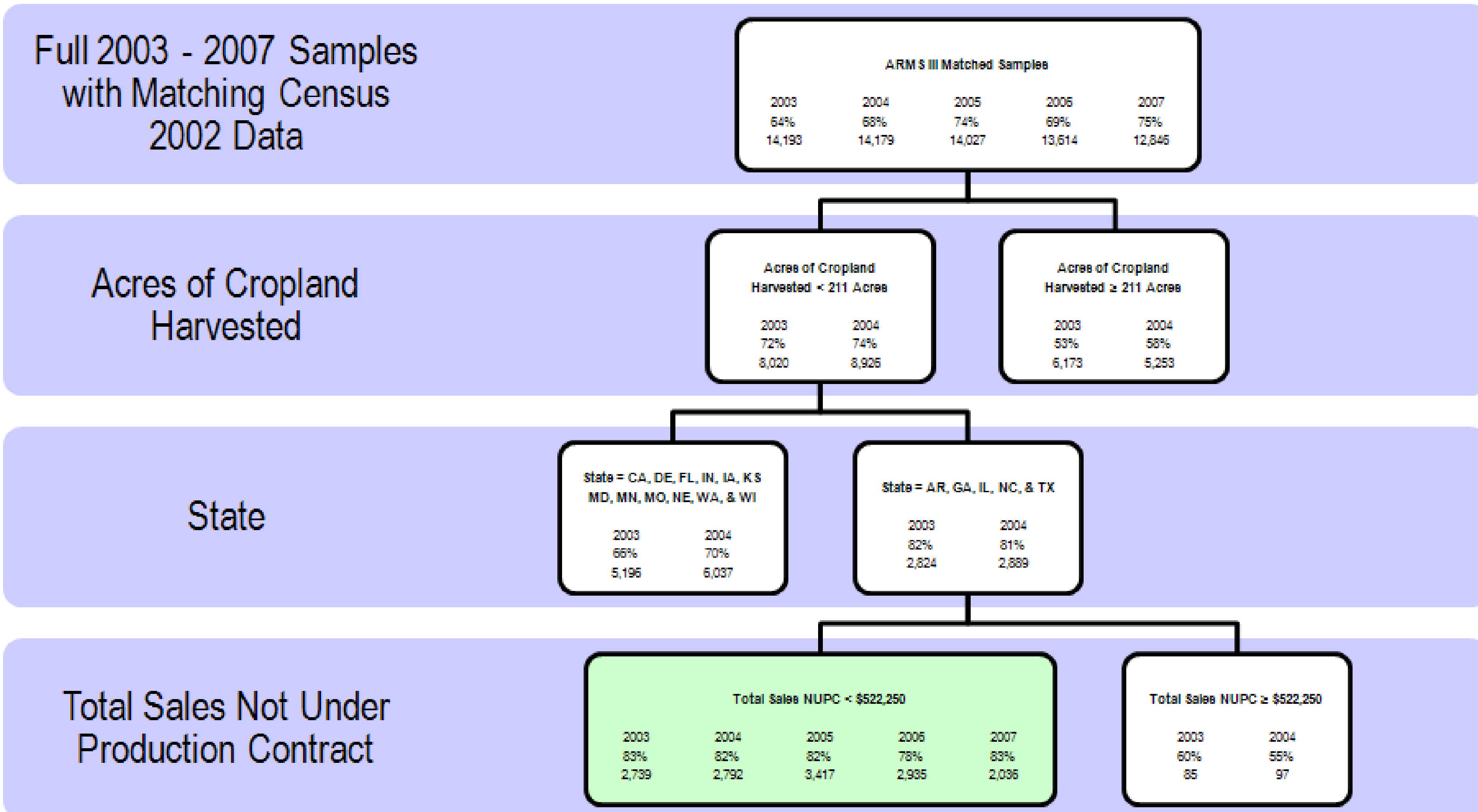

Figure 2. Model One - Acres of Cropland Harvested (LogWorth = 43.9, n = 14,193, p < .20)

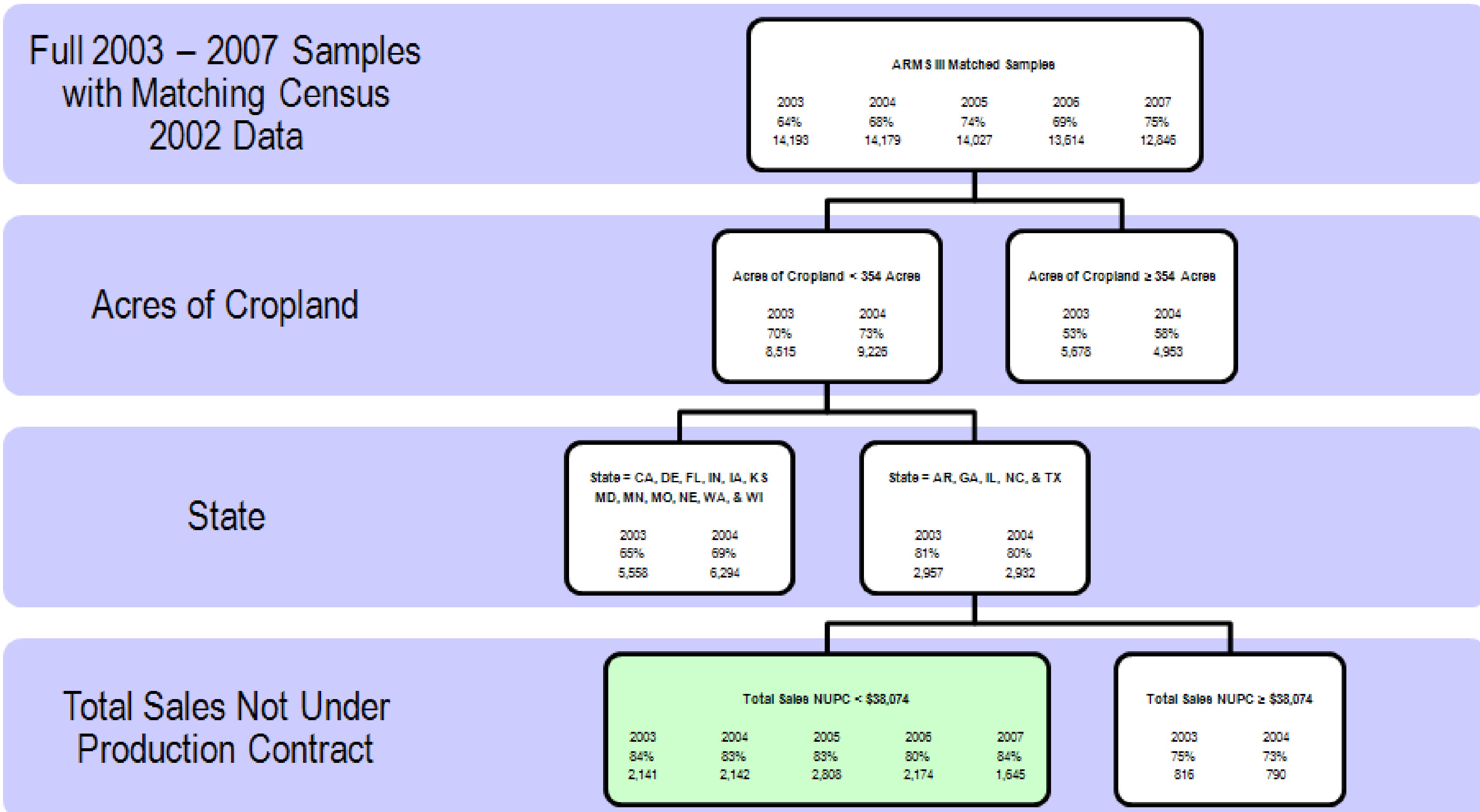

Figure 3. Model Two - Acres of Cropland (LogWorth = 43.0, n = 14,193, $p < .20$)

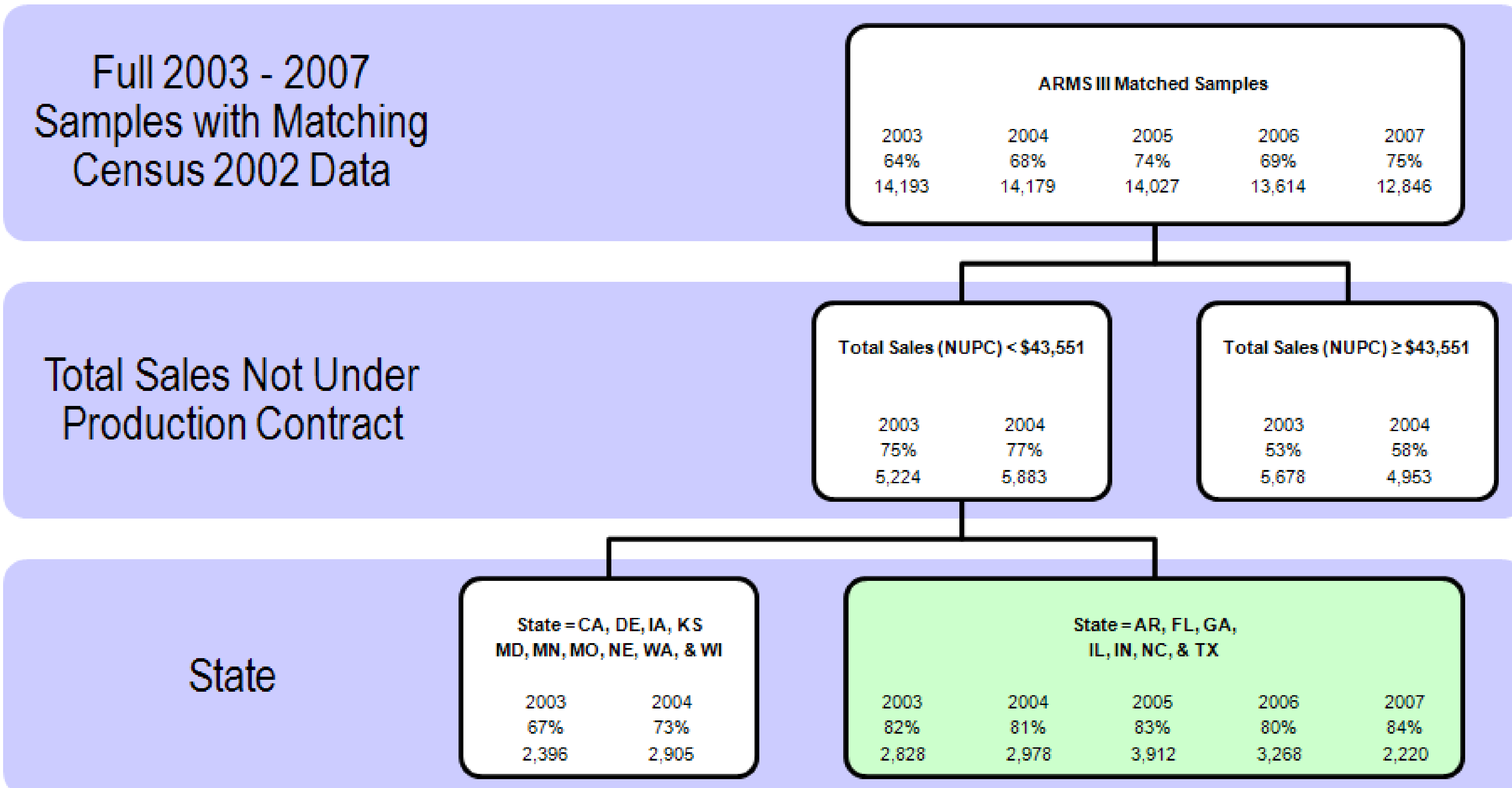

Figure 4. Total Acres under Production Contract (LogWorth = 40.5, n = 14,193, p < .20)

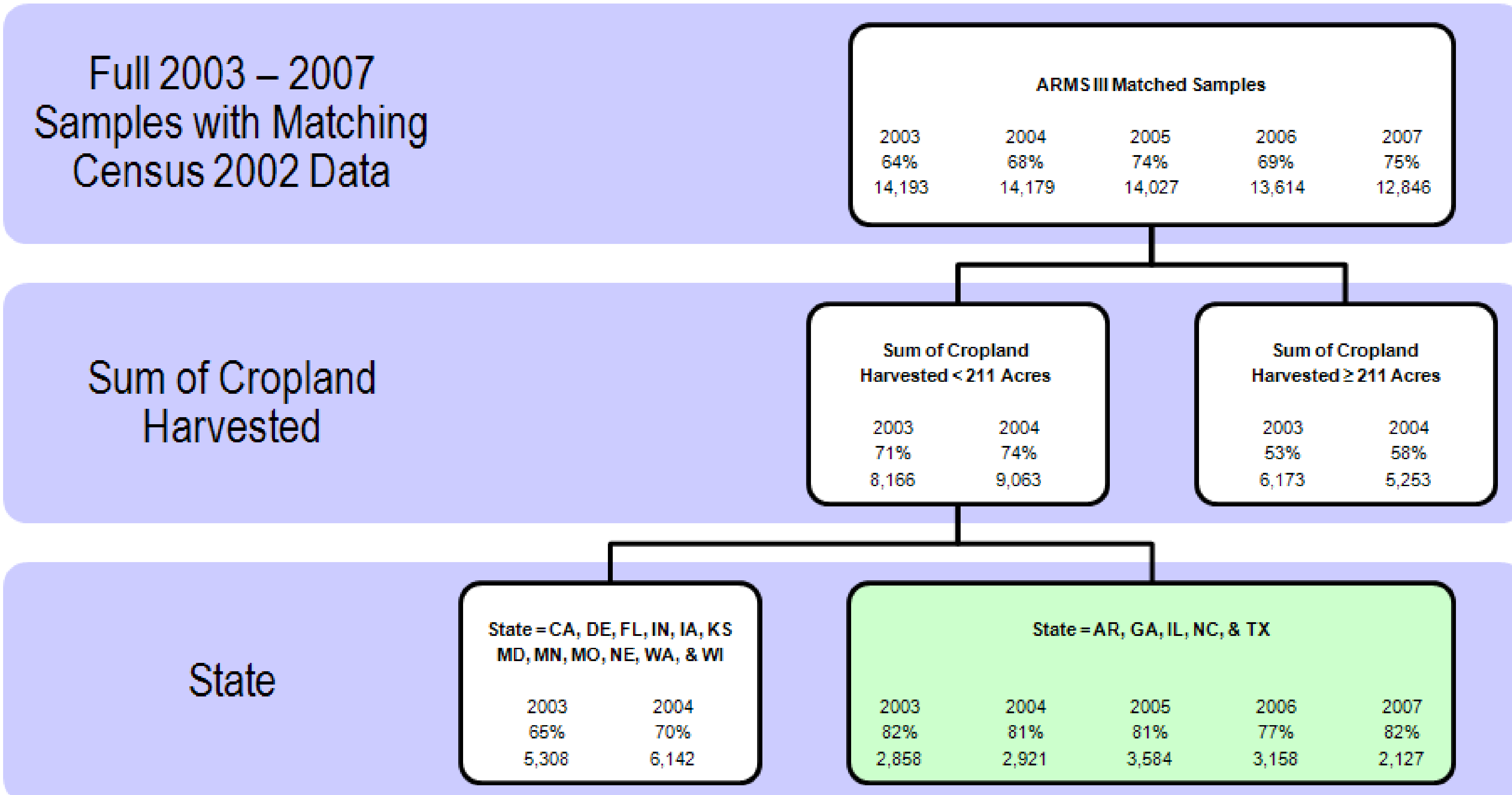

Figure 5. Sum of Cropland Acres (LogWorth = 39.2, n = 14,193, $p < .20$)

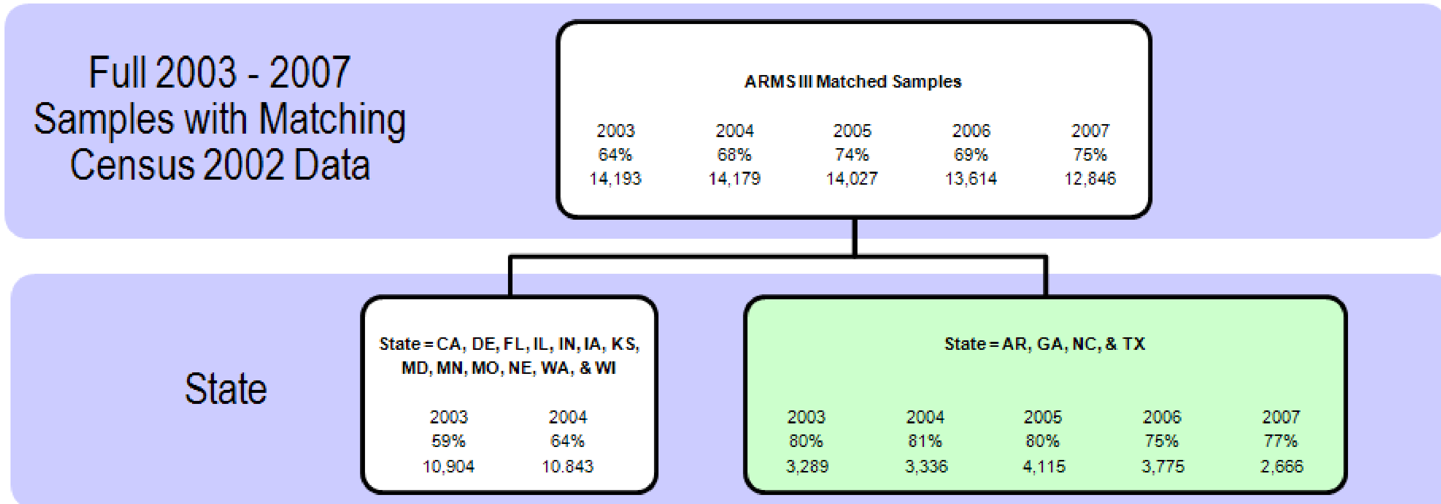

Figure 6. State (LogWorth = 29.7, n = 14,193, p < .20)

3.2 *Model Assessment*

3.2.1 Accuracy Assessments

The identified subgroups of likely respondents maintained consistent rates of response over time (2003-2007) even after incentives were introduced (Figure 7).

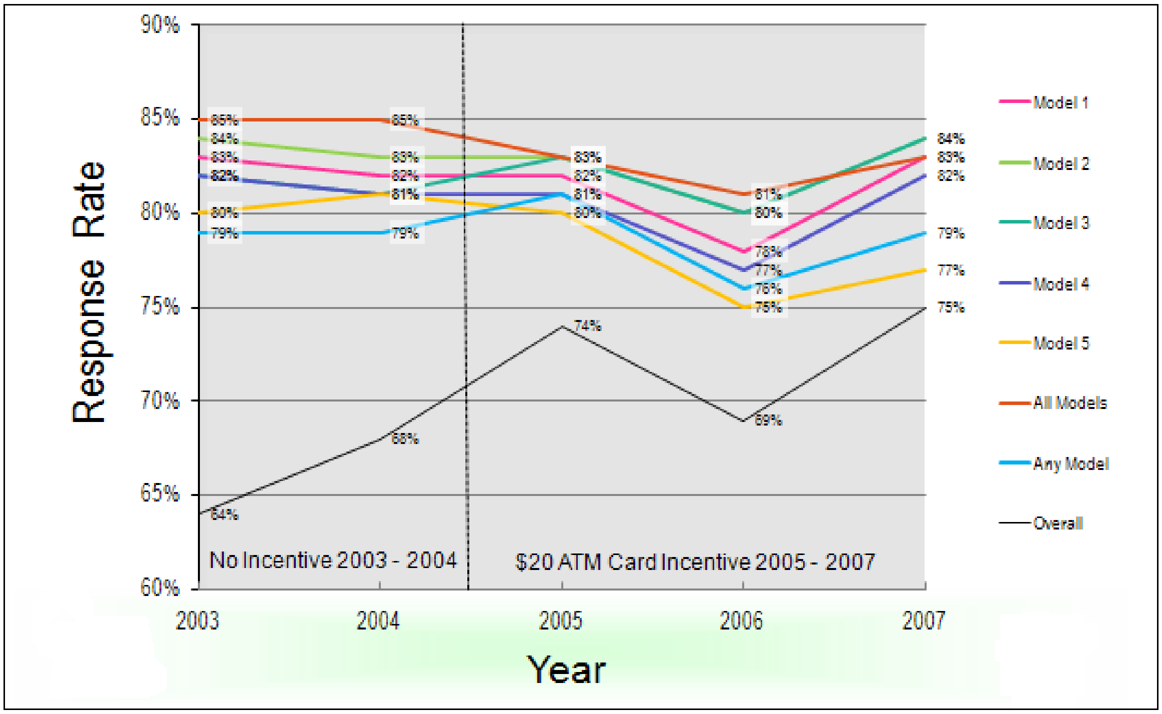

Figure 7. Model Accuracy over Time

In other words, these subgroups of operations responded at a higher rate than the sample as a whole, both with and without incentives. Data from 2003-2007 were used to compare the five respondent characteristic subgroups and determine average response rate gain and variability over time.

- **Model One** – Operations with Acres of Cropland Harvested < 211 acres, located in Arkansas, Georgia, Illinois, North Carolina, or Texas, with Total Sales NUPC < $522,250 ($n$ = 13,919) responded at an average rate of 81.60 percent (s = 2.07 percent), averaging 11.60 percent (s = 4.83 percent) above the overall sample from 2003 through 2007.

- **Model Two** – Operations with Acres of Cropland < 354 acres, located in Arkansas, Georgia, Illinois, North Carolina, or Texas, with Total Sales Not Under Production Contract less than $38,074 ($n$ = 10,910) responded at an average rate of 82.80 percent (s = 1.64 percent), averaging 12.80 percent above the overall sample from 2003 through 2007.

- **Model Three** – Operations with Total Sales NUPC < $43,551 located in Arkansas, Florida, Georgia, Illinois, Indiana, North Carolina, and Texas (n = 15,206) responded at an average rate of 82.00 percent (s = 1.58 percent), averaging 12.00 percent (s =3.74 percent) above the overall sample from 2003 through 2007.

- **Model Four -** Operations with Sum of Cropland Harvested < 211 acres located in Arkansas, Georgia, Illinois, North Carolina, and Texas (n = 14,648) responded at an average rate of 80.60 percent (s = 2.07 percent), averaging 10.60 percent (s = 4.83 percent) above the overall sample from 2003 through 2007.

- **Model Five -** Operations located in Arkansas, Georgia, North Carolina, and Texas (n = 17,181) responded at an average rate of 78.60 percent (s =2.51 percent), averaging 8.60 percent (s = 5.73 percent) above the overall sample from 2003 through 2007.

- **All Models** – Operations identified in Model One, Model Two, Model Three, Model Four, and Model Five responded at an average rate of 83.40 percent (s = 1.67 percent), averaging 13.40 percent (s = 5.50 percent) above the overall sample from 2003 through 2007.

- **Any Model** – Operations identified in Model One, Model Two, Model Three, Model Four, or Model Five responded at an average rate of 78.80 percent (s = 1.79 percent), averaging 8.80 percent (s = 4.27 percent) above the overall sample from 2003 through 2007.

Using the "All Models" indicator to identify likely respondents will result in the greatest respondent classification accuracy; however, it also identifies the smallest group of respondents (Figure 8). Although there is a 4.60 percent drop in prediction accuracy when using the "Any Model" indicator versus the "All Models" indicator, the "Any Models" indicator correctly identified over twice as many respondents; therefore, using the "Any Models" indicator provides the potential for saving over twice the resources, that may be better allocated toward converting likely mail nonrespondents not enticed by the $20 ATM Card currently offered.

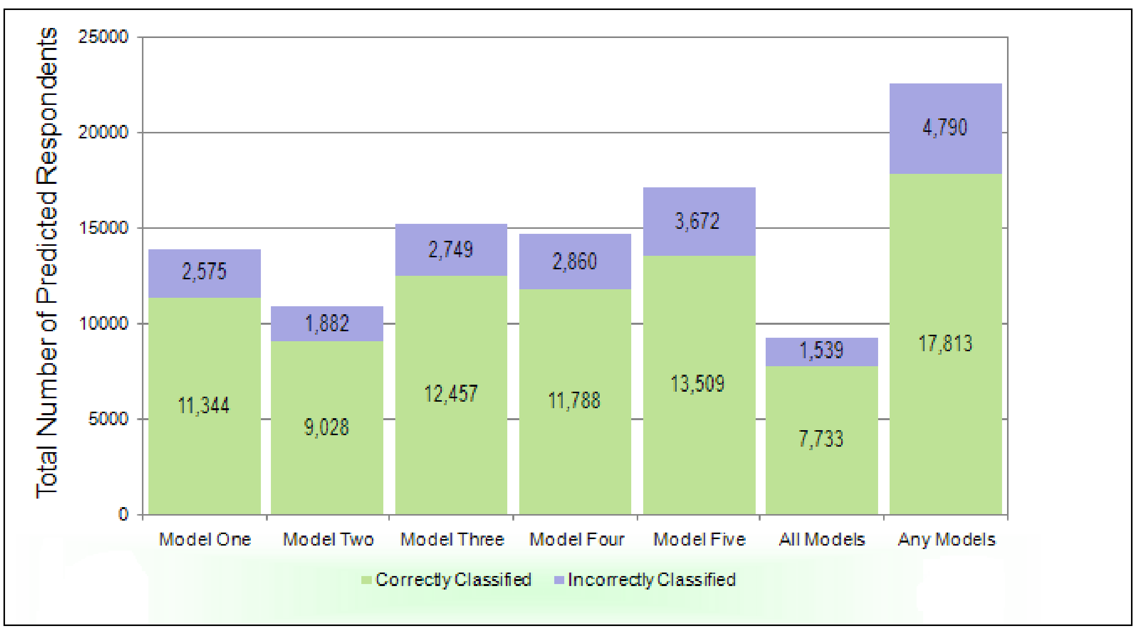

Figure 8. Number of Respondents Correctly Identified by Model

3.2.2 Cost Assessments

The identified subgroups of likely respondents maintained consistent rates of response over time (2003-2007) even after incentives were introduced in Models One, Two, Three, and Four, but not in Model Five (Figure 7). Therefore, incentives are neither increasing nor decreasing response rates for these subgroups, and are unnecessary for them. These resources may be better allocated elsewhere. Each of the above models was reassessed in terms of cost as opposed to accuracy in order to determine the most efficient way to reallocate current incentive resources. According to Figure 8, the "All Models" indicator may exhibit the lowest misclassification rate, but it is clear that the "Any Model" indicator correctly identifies more respondents overall. The average minimum and maximum projected cost savings within the five respondent characteristic subgroupings identified in Models 1-5 were estimated using 2002 Census data for the 2005-2007 ARMS III sample to calculate the average annual cost of administering the Mail Version of ARMS III, both with and without the $20 ATM Card incentive (Figure 9 & 9).

Table 2: ARMS III Mail Version Administration and Incentive Costs per Sample Unit

Cost Factors	Given $20 ATM Card				Given No $20 ATM Card	
	Minimum Cost in Dollars [5]		Maximum Cost in Dollars [6]		Cost in Dollars	
	Respondent	Nonrespondent	Respondent	Nonrespondent	Respondent	Nonrespondent
First Class Mail	0.305	0.305	0.305	0.305	0.305	0.305
Pre-Survey Letter	0.045	0.045	0.045	0.045	0.045	0.045
ATM Card Instructions	0.045	0.045	0.045	0.045	0.000	0.000
ATM Card Stuffing Fee	0.440	0.440	0.440	0.440	0.000	0.000
ATM Administrative Fee	0.850	0.850	0.850	0.850	0.000	0.000
ATM Card Value	0.000	0.000	20.000	20.000	0.000	0.000
ATM Bank Fees	0.000	0.000	4.000	4.000	0.000	0.000
Follow Up Enumerator Visit	0.000	128.000	0.000	128.000	0.000	128.000
Total Cost in Dollars	1.6850	129.685	25.685	153.685	0.350	128.350

[5] Cost if card is not cashed and no bank fees are used.
[6] Cost if card is cashed in full and all bank fees are used.

Table 3: Average Annual Minimum and Maximum Cost Estimates of Incentives from 2005 to 2007 [7 & 8]

Models	Average Annual Minimum Cost in Dollars			Average Annual Maximum Cost in Dollars		
	Respondent ($1.685)	Nonrespondent ($129.685)	Total	Respondent ($25.685)	Nonrespondent ($153.685)	Total
Model 1	3,808.73	69,462.74	73,271.47	58,057.69	82,317.78	140,375.47
Model 2	3,062.00	50,808.85	53,870.86	46,675.12	60,211.73	106,886.86
Model 3	4,339.53	72,357.31	76,696.84	66,148.81	85,748.03	151,896.84
Model 4	3,975.95	77,385.63	81,361.58	60,606.67	91,706.91	152,313.58
Model 5	4,592.23	102,880.41	107,472.63	70,000.79	121,919.85	191,920.63
All Models	2,589.60	42,728.18	45,317.79	39,474.16	50,635.62	90,109.79
Any Model	6,247.46	130,200.28	136,447.75	95,232.10	154,295.64	249,527.75

Table 4: Projected Average Annual Reduced Cost for Likely Respondent Groups

Models	Average Annual Cost of Monetary Incentives and Savings having Eliminated Monetary Incentives for Likely Respondents in Dollars						
	Average Annual Reduced Cost			Average Annual Minimum Savings		Average Annual Maximum Savings	
	Respondent ($0.35)	Nonrespondent ($128.35)	Total	Total	Percent	Total	Percent
Model 1	$791.13	$68,747.68	$69,538.81	$3,732.66	5.09%	$70,836.66	50.46%
Model 2	$636.02	$50,285.82	$50,921.84	$2,949.02	5.47%	$55,965.02	52.36%
Model 3	$901.39	$71,612.45	$72,513.84	$4,183.00	5.45%	$79,383.00	52.26%
Model 4	$825.86	$76,589.01	$77,414.88	$3,946.71	4.85%	$74,898.71	49.17%
Model 5	$953.87	$101,821.34	$102,775.21	$4,697.42	4.37%	$89,145.42	46.45%
All Models	$537.90	$42,288.33	$42,826.23	$2,491.56	5.50%	$47,283.56	52.47%
Any Model	$1,297.69	$128,859.98	$130,157.67	$6,290.08	4.61%	$119,370.08	47.84%

- **Model One** – Operations with Acres of Cropland Harvested < 211 acres, located in Arkansas, Georgia, Illinois, North Carolina, or Texas, with Total Sales NUPC < $522,250 ($n$ = 8,388) responded at an average rate of 81.00 percent from 2005 to 2007 (s = 2.65 percent) resulting in an average annual minimum cost of $73,271.47 and an

[7] Assumes no one cashes the ATM card or uses funds allocated for fees:
[**Minimum Cost = Respondent Minimum Cost** ($0.305 First Class Mail + $0.045 Pre-Survey Letter + $0.045 ATM Card Instructions + $0.440 ATM Card Stuffing Fee + $0.850 ATM Administrative Fee = **$1.685**) **+ Nonrespondent Minimum Cost** ($0.305 First Class Mail + $0.045 Pre-Survey Letter + $0.045 ATM Card Instructions + $0.440 ATM Card Stuffing Fee + $0.850 ATM Administrative Fee + $128.00 Follow-Up Enumerator Visit = **$129.685**)]

[8] Assumes everyone cashes the ATM card and uses all funds allocated for fees:
[**Maximum Cost = Respondent Maximum Cost** ($0.305 First Class Mail + $0.045 Pre-Survey Letter + $0.045 ATM Card Instructions + $0.440 ATM Card Stuffing Fee + $0.850 ATM Administrative Fee + $20.00 ATM Card Value + $4.00 ATM Bank Fees = **$25.685**) **+ Nonrespondent Maximum Cost** ($0.305 First Class Mail + $0.045 Pre-Survey Letter + $0.045 ATM Card Instructions + $0.440 ATM Card Stuffing Fee + $0.850 ATM Administrative Fee + $20.00 ATM Card Value + $4.00 ATM Bank Fees + $128.00 Follow-Up Enumerator Visit = **$153.685**)]

annual average maximum cost of $140,375.47 per year. By eliminating such operations from incentive allocation, NASS is capable of reducing costs within the above group to $69,538.81, saving between $70,836.66 (50.46 percent) and $3,732.66 (5.09 percent) annually.

- **Model Two** – Operations with Acres of Cropland < 354 acres, located in Arkansas, Georgia, Illinois, North Carolina, or Texas (*n* = 6,627), and Total Sales Not Under Production Contract less than $38,074 responded at an average rate of 82.33 percent from 2005 to 2007 (*s* = 2.08 percent) resulting in an average annual minimum cost of $53,870.86 and an average annual maximum cost of $106,886.86 per year. By eliminating such operations from incentive allocation, NASS is capable of reducing costs within the above group to $50,921.84, saving between $55,965.02 (52.36 percent) and $2,949.02 (5.47 percent) annually.

- **Model Three** – Operations with Total Sales NUPC < $43,551 located in Arkansas, Florida, Georgia, Illinois, Indiana, North Carolina, and Texas (*n* = 9,440) responded at an average rate of 82.33 percent from 2005 to 2007 (*s* = 2.08 percent) resulting in an average annual minimum cost of $76,696.84 and an average annual maximum cost of $151,896.84 per year. By eliminating such operations from incentive allocation, NASS is capable of reducing costs within the above group to $72,513.84, saving between $79,383.00 (52.26 percent) and $4,183.00 (5.45 percent) annually.

- **Model Four –** Operations with Sum of Cropland Harvested < 211 acres located in Arkansas, Georgia, Illinois, North Carolina, and Texas (*n* = 8,869) responded at an average rate of 80.00 percent from 2005 to 2007 (*s* = 2.65 percent) resulting in an average annual minimum cost of $81,361.58 and an average annual maximum cost of $152,313.58 per year. By eliminating such operations from incentive allocation, NASS is capable of reducing costs within the above group to $77,414.88, saving between $74,898.71 (49.17 percent) and $3,946.71 (4.85 percent) annually.

- **Model Five** – Operations located in Arkansas, Georgia, North Carolina, and Texas (*n* = 10,556) responded at an average rate of 77.33 percent from 2005 to 2007 (*s* = 2.52 percent) resulting in an average annual minimum cost of $107,472.63 and an average annual maximum cost of $191,920.63 per year. By eliminating such operations from incentive allocation, NASS is capable of reducing costs within the above group to $102,775.21, saving between $89,145.42 (46.45 percent) and $4,697.42 (4.37 percent) annually.

- **All Models –** Operations identified in Model One, Model Two, Model Three, Model Four, and Model Five (*n* = 5,599) responded at an average rate of 82.33 percent (*s* = 1.15 percent) resulting in an average minimum cost of $45,317.79 and an average maximum cost of $90,109.79 per year. By eliminating such operations from incentive allocation, NASS is capable of reducing costs within the above group to $42,826.23, saving between $47,283.56 (52.47 percent) and $2,491.56 (5.50 percent) annually.

- **Any Model** – Operations identified in Model One, Model Two, Model Three, Model Four, or Model Five (*n* = 14,135) responded at an average rate of 78.67 percent (*s* = 2.52 percent) resulting in an average minimum cost of $136,447.75 and an average maximum cost of $249,527.75 per year. By eliminating such operations from incentive allocation, NASS is capable of reducing costs within the above group to $130,157.67, saving between $119,370.08 (47.84 percent) and $6,290.08 (4.61 percent) annually.

Using "Any Model" to identify likely respondents and reallocate monetary incentive funds will result in the greatest annual savings between $6,290.08 and $119,370.08 and will provide NASS with the ability to reallocate funds earmarked for those sample units to entice likely mail nonrespondents currently not enticed by the monetary incentive.

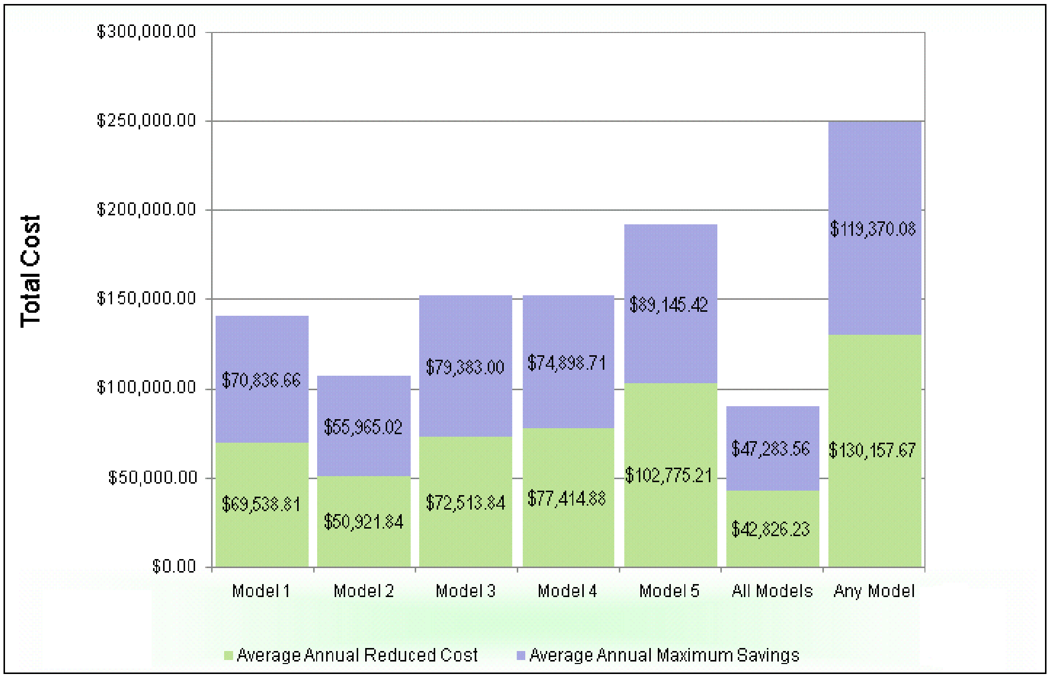

Figure 9. Average Annual Maximum Savings

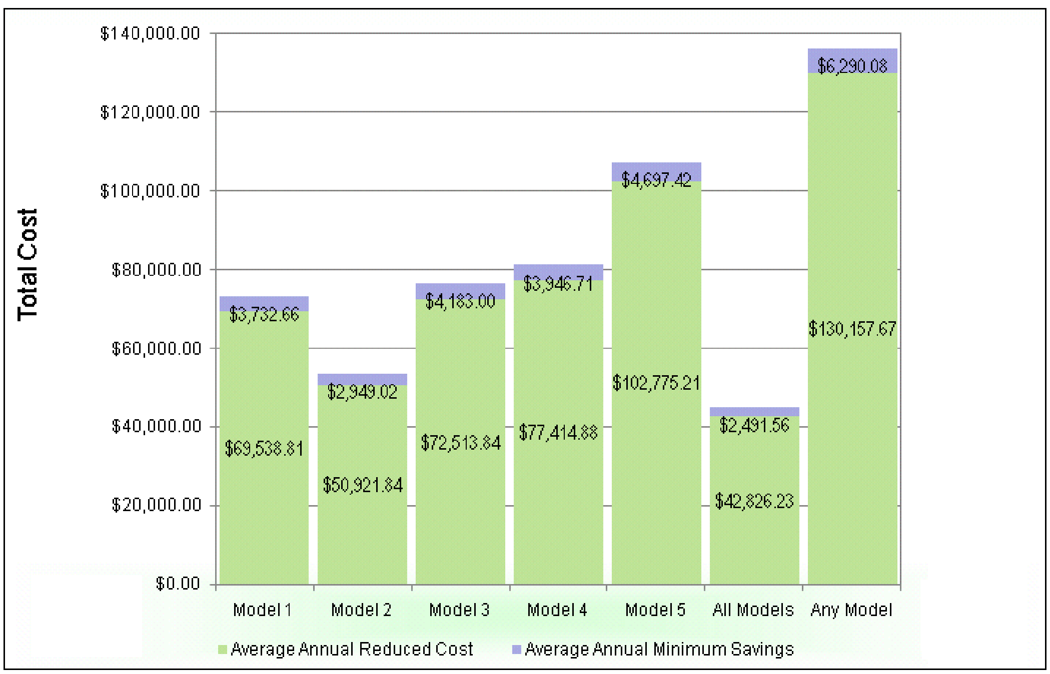

Figure 20. Average Annual Minimum Savings

4. DISCUSSION

The models discussed in this report identify consistent respondent characteristics with or without the use of monetary incentives. In order to determine the operations for which incentives are not necessary (those already responding at higher rates than other operations), criteria identified by any of the five models should be used to score future ARMS III Core samples, starting with 2009. In fact, to minimize bias, once these operations are identified, their previously allocated incentive funds could be redirected towards exploring alternative incentives for consistent mail nonrespondents currently unresponsive to monetary incentives. The identification of those non-respondent subgroups will be the subject of a follow-up research report to this one. Redirecting these funds is not only cost effective, but also works toward reducing both response and nonresponse bias. Given that the above groups are already more apt to respond relative to the rest of the sample, efforts should focus on soliciting responses from under-represented operations, not necessarily those already represented, if the goal is ultimately to reduce nonresponse bias.

The follow-up report will identify mail nonrespondents using 2002 Census data for ARMS III Core 2005-2007 samples. Cognitive interviews should be conducted with a sample of operations identified as mail nonrespondents in 2007, in order to determine alternative incentives or data collection strategies, since the current $20 ATM monetary incentive appears ineffective. Based on cognitive interviews, alternative incentives will be identified for specific groups of mail nonrespondents and recommended for use with the 2010 ARMS III Core sample. Ultimately, it is expected that this two-phase study will result in more efficient incentive allocation methods, and reductions in survey costs, non-response, and bias.

5. LIMITATIONS

Although the above research aims to improve data quality and reduce the waste of taxpayer funds, implementation depends on approval from the Office of Management and Budget (OMB) which currently has only approved equitable distribution of incentives. The *Guidance on Agency Survey and Statistical Information Collections* report states, "Agencies should treat all respondents equally with regard to incentives. OMB generally does not approve agency plans to give incentives solely to convert refusals, or treat specific subgroups differently, unless the plan is part of an experimental design for further investigation into the effects of incentives" (2006, p.25).

If such an incentive allocation is approved for an experimental design study, the question becomes whether differential allocation may continue beyond the experiment. If it is not approved, alternative uses of the research may include eliminating incentives and reallocating funds towards oversampling likely mail nonrespondent groups. However, although this will likely reduce bias, it will increase the overall likelihood of nonresponse within the sample, almost ensuring that response rates will remain well below the OMB's standard of 80 percent.

Another potential alternative could be elimination of incentives and reallocation of funds towards rewarding enumerators for obtaining good responses from operations identified as likely nonrespondents. However, this too has a downside in that encouraging refusal conversions may not actually improve data quality. It is possible that enumerators will feel financially pressured to convert a refusal regardless of reporting capability or accuracy.

6. RECOMMENDATIONS

1. Score the 2009 ARMS III Core sample using criteria specified in the five models. Records meeting the criteria specified on one or more models will be flagged as likely respondents.

2. Conduct a similar study using ARMS III Core 2005 training, 2006 validation, and ultimately 2007 test data to flag 2010 mail nonrespondents.

3. Contact flagged and confirmed ARMS III Core 2007 mail nonrespondents for cognitive interviews in order to identify alternative incentives for use in 2010.

4. Randomly divide flagged 2010 mail nonrespondents into three groups: 1) A control group receiving no incentive, 2) a treatment group receiving a $20 ATM card incentive, and 3) a treatment group receiving an alternative incentive identified via cognitive interviews to determine if the identified alternative incentive is more effective for the given mail nonresponse group.

7. REFERENCES

Beckler, D.G., Ott, K., and Horvath, P. (2005), *Indirect Monetary Incentives for the 2004 ARMS Phase III Core*. Research and Development Division Staff Report RDD-05-05, United States Department of Agriculture, National Agricultural Statistics Service.

deVille, B. (2006). Decision Trees for Business Intelligence and Data Mining using SAS Enterprise Miner. Carey, NC:SAS Institute, Inc.

McCarthy, J.S., Beckler, D.G., and Ott, K. (2006). *The Effect of Incentives on Response in 2005 ARMS Phase III Interviews*. Research and Development Division Staff Report RDD-06-07, United States Department of Agriculture, National Agricultural Statistics Service.

National Research Council. (2008), *Understanding American Agriculture: Challenges for the Agricultural Resource Management Survey.* Panel to Review USDA's Agricultural Resource Management Survey. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington DC: The National Academic Press.

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

United States. Department of Agriculture. (2008). *PRISM II Code Book.*. Washington, DC: U.S. Department of Agriculture.

United States. Executive Office of the President. (2006). *Guidance on Agency Survey and Statistical Information Collections*. Washington, DC: U.S. Executive Office of the President.

8. APPENDIX A

Table A1: ARMS III Sample Sizes and Census 2002 Operation Match Rates

ARMS III Sample Year	ARMS III Sample Size (n)	ARMS III & Census 2002 Operation Match Frequency	ARMS III & Census 2002 Operation Match Percent
2003	33,861	27,052	79.89%
2004	33,908	28,192	83.14%
2005	34,937	28,336	81.11%
2006	34,203	27,408	80.13%
2007	31,924	23,946	75.01%
Total	168,333	134,934	79.92%

Table A2: ARMS III Sample Sizes and Census 2002 Operation Match Rates for Respondents versus Nonrespondents [9]

ARMS III Sample Year	ARMS III Respondents (n_r)	Respondent Match Frequency	Respondent Match Percent	ARMS III Nonrespondents (n_n)	Nonrespondent Match Frequency	Nonrespondent Match Percent
2003	21,282	17,262	81.11%	12,579	9,790	77.83%
2004	22,966	19,436	84.63%	10,942	8,756	80.02%
2005	24,684	20,521	83.13%	10,253	7,815	76.22%
2006	23,227	18,996	81.78%	10,965	8,412	76.72%
2007	22,288	17,218	77.25%	9,623	6,728	69.92%
Total	114,447	93,433	81.64%	54,362	41,501	76.34%

[9] The 2002 Census match rate was significantly higher for ARMS III respondents than nonrespondents:
$Z_{2003} = 7.28, \quad p < .05$
$Z_{2004} = 10.59, p < .05$
$Z_{2005} = 15.03, p < .05$
$Z_{2006} = 10.97, p < .05$
$Z_{2007} = 13.87, p < .05$
$Z_{Total} = 25.40, p < .05$

Table A3: ARMS III Sample Sizes and Census 2002 Operation Match Rates for ARMS Core Version versus the Noncore Versions [10]

ARMS III Sample Year	ARMS III Core Version Recipients (n_{cv})	Core Version Match Frequency	Core Version Match Percent	ARMS III Noncore Version Recipients (n_{ncv})	Noncore Version Match Frequency	Noncore Version Match Percent
2003	16,954	14,193	83.71%	16,907	12,859	76.06%
2004	15,900	14,179	89.18%	18,008	14,013	77.82%
2005	16,499	14,027	85.02%	18,438	14,309	77.61%
2006	16,489	13,614	82.56%	17,703	13,794	77.92%
2007	16,493	12,846	77.89%	15,418	11,100	71.99%
Total	82,335	68,859	83.63%	86,474	66,075	76.41%

[10] The 2002 Census match rate was significantly higher for the ARMS III Core Version than the Noncore Versions:
$Z_{2003} = 17.56$, $p < .05$
$Z_{2004} = 27.88$, $p < .05$
$Z_{2005} = 17.66$, $p < .05$
$Z_{2006} = 10.75$, $p < .05$
$Z_{2007} = 12.16$, $p < .05$
$Z_{Total} = 37.02$, $p < .05$