 doi:10.20944/preprints201804.0022.v1

Article

Heterogeneity of Cooperative Membership: Implications for Cooperative Sustainability

Matthew Elliott [1], Lisa Elliott [2] and Evert Van der Sluis[2,*]

[1] South Dakota State University; Matthew.Elliott@sdstate.edu
[2] South Dakota State University; Lisa.Elliott@sdstate.edu
[3] South Dakota State University; Evert.VanDerSluis@sdstate.edu

* Correspondence: Matthew.Elliott@sdstate.edu; Tel.: +01-605-688-4855

Abstract: The effects of heterogeneity of cooperative membership on cooperative and collective action sustainability has been previously discussed. However, despite the importance of membership heterogeneity in recent theoretical frameworks, empirical examinations have been limited. We determine the effect of changes to cooperative member heterogeneity on cooperative sustainability and discuss changes to heterogeneity overtime that can advance our understanding to cooperative sustainability long-term. This study uses USDA Agricultural Management Resource Survey data, coupled with USDA-Rural Development cooperative financial data at the state level, to quantify effects of cooperative member heterogeneity to sustainability of U.S. farmer cooperatives. We use random forest regression to interpret the significance of heterogeneity with cooperative sustainability at an aggregate level. The findings of this empirical study narrowly reconciles the theoretical understanding of the emergence of intra-cooperative issues while providing consistent empirical evidence and expectations for the sustainability of cooperatives in the near term.

Keywords: Cooperatives, Membership Heterogeneity, Random Forest, Collective Action

---

1. Introduction

Recent cooperative theoretical literature has focused on membership heterogeneity in understanding cooperative sustainability long-term [1]. Despite the significance of membership heterogeneity in advancing the theoretical understanding of cooperative sustainability, the empirical attention has not been comparable. This is in part due to lack of data on cooperative membership heterogeneity and/or limited empirical methods that can advance our understanding. An ideal dataset to conduct an empirical examination of cooperative membership heterogeneity on cooperative performance is not presently available. However, demographic and financial information on farm producers who reported they received a cooperative patronage or had equity in a cooperative has been collected in the USDA Agricultural Resource Management Survey (ARMS) since 1996. Additionally, the ARMS survey collects a number of other variables that have been largely posited to be associated with cooperative membership heterogeneity that would affect cooperative sustainability [2]. Moreover, recent developments in machine learning methods have enabled researchers to examine high dimensional data, and make accurate statistical inferences concerning variables with low explanatory power due to aggregation issues [3-4]. Specifically, random forest ensemble methods have been found to be substantially more powerful in understanding treatment effect heterogeneity compared to classical methods [5-6]. We aim to use random forest methods to understand the treatment effect to cooperative membership heterogeneity

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 April 2018 doi:10.20944/preprints201804.0022.v1

on cooperative sustainability if membership heterogeneity was more or less. Using random forests, we can better account for the interactions, non-linearities, and hidden effects that membership heterogeneity may have in predicting cooperative sustainability long-term -- particularly in the presence of irrelevant and strongly relevant covariates.

We used annual aggregated data at the state level due to the limited availability of more detailed data on cooperatives and their membership makeup. We combined multiple farm and socioeconomic variables of cooperative membership and non-cooperative membership heterogeneity at the state level from the ARMS data with the long historical series of farmer cooperative membership, gross business volume, and number of cooperatives headquartered in a state that is maintained by the USDA-Rural Development [7]. We employed the random forest method to draw statistical inferences of variables associated with membership heterogeneity in order to advance our understanding of the effects to cooperative sustainability at an aggregated level. We aim to understand: 1) the effects to cooperative sustainability due to heterogeneity of cooperative membership, 2) changes to cooperative members and participation due to membership heterogeneity, and 3) assess the effect of membership heterogeneity on long-term cooperative sustainability. We included other variables that may have an equal impact on cooperative sustainability long-term. These variables include changes in prices for farm products, food and feed products, consumer products, and data on the amount of value added by industry derived by the Bureau of Economic Analysis (BEA) in estimating U.S. state GDP [8].

The findings of our study in general show that membership heterogeneity plays a lesser role compared to regional variables in understanding cooperative gross business volume at the state level. This may be in part due to cooperatives have established themselves in industries and regions for providing goods and services at competitive prices. Furthermore, we find that value addition at the farm is correlated with greater cooperative business volume at the state level. However, we draw no conclusions on causality in this analysis. The correlation is consistent with the theoretical notion that cooperatives are able to return greater value to the farm than would be expected in oligopsony or oligopoly markets. The finding is also consistent with the notion that cooperatives are sustainable long-term despite the potential for intra-cooperative issues.

We do find stronger evidence that membership heterogeneity plays an important role in understanding the future makeup and number of cooperatives long-term. For example, the model that was generated indicates that farmer cooperative member mean socioeconomic (SES) status is inversely related to the number of cooperative members and the number of cooperatives headquartered in a state. Indeed, states where cooperative members have greater mean socioeconomic status we observe greater degrees of cooperative consolidation and cooperative member declines. Moreover, cooperative members SES diversity was found to be inversely related to the number of cooperative members and cooperatives headquartered in the state. Additionally, the coefficient of variation in cooperative membership age class was found to also be inversely related to the number of cooperatives headquartered in the state and the number of cooperative members. Contrary to the notion that diversity makes collective action more difficult, the random forest model that was generated showed expectations that there would be more cooperatives and cooperative members as age and socioeconomic diversity of cooperative members increased. This finding may be counterintuitive to the theoretical notion that cooperative membership heterogeneity decreases the likelihood of cooperative action. However, the finding may be consistent in explaining that in the presence of greater cooperative membership heterogeneity, consolidation and acquisition of cooperatives has been less rapid.

The findings of this study seem to suggest that intra-cooperative issues associated with membership heterogeneity expectedly play a role in cooperative consolidation and cooperative extent. At the same time, intra-cooperative issues may not be as important in understanding the long-term sustainability of cooperatives business volume and market share of cooperatives in the agri-food industry. The findings of this study can (i) further our understanding of emerging cooperative issues, (ii) draw implications on the long-term sustainability of cooperatives due to changes to cooperative member heterogeneity, and (iii) reconcile the recent theoretical focus on intra-

 doi:10.20944/preprints201804.0022.v1

cooperative issues that is consistent with observations of cooperative survival over long periods of time. Our empirical study supports a continued focus in understanding intra-cooperative issues for advancing our understanding of cooperative sustainability, membership makeup, and cooperative extent at a micro level. At the same time, the study provides a reason to why a focus on intra-cooperative issues due to member heterogeneity does not necessarily imply that long-term cooperative sustainability is in peril. Indeed, the results of our analysis indicate that in the near-term the cooperative gross business volume will remain stable, if not increase, regardless of membership heterogeneity and associated intra-cooperative issues.

2. Conceptual Framework: Defining Cooperative Sustainability in the Presence of Membership Heterogeneity

Cook, Chaddad and Illopoulos [1], discuss three theoretical approaches that have advanced cooperative understanding. The first approach addressed how membership heterogeneity may not necessarily lead to the instability of the cooperative. The supposition was that the cooperative can be sustainable by providing socially desirable efficiency outcomes in imperfect markets. The theoretical works in this approach view the cooperative as extensions of the firm in a neo-classical framework. Theoretical deductions provide proof that cooperatives could successfully provide heterogeneous cooperative members surplus in multipurpose cooperatives [9-10], and over wider spatial dimensions, and in the presence of competition from investor owned firms (IOFs) [11]. These frameworks largely assume an oligopsony or oligopoly market, and/or degrees of asset specificity (site, temporal, physical, human) in segments of the supply chain. The frameworks largely view cooperative sustainability as dependent on providing socially desirable efficiency improvements by correcting market failures resulting from IOFs exerting market power.

The second approach describes cooperative sustainability as maintaining a coalition with a common interest, yet potentially diverse incentives. Much of the literature in this approach implicitly concedes the need for the cooperative, and/or collective action institution in general, to address market inefficiencies from externalities. The second approach is largely framed in the *tragedy of the commons* context--where cooperatives can provide a second-best contractual solution to market failures [12]. However, the second approach highlights the inevitable influence problem and horizon problems that can result, and lead to a similar inefficiencies that exists without a collective action solution. Horizon and influence problems can be exacerbated by membership heterogeneity that would diminish the common interest of the coalition and create greater instability of cooperative equilibriums long-term. The frameworks rely primarily on game theory and draw largely from public choice [13] and collective action frameworks [14]. The second approach defines sustainability of the cooperative as making optimal governance decisions despite membership heterogeneity. Cooperatives can be sustainable by altering cooperative size and governance rules to more effectively economize on the bargaining costs and decision-making costs that result in sustaining collective action.

The third approach defines cooperative sustainability as controlling agency issues by designing optimal incentive contracts and reducing exhaustive bargaining in contractual holdup. This approach largely uses incomplete contract theory and ex-ante/ex-post asset specificity in understanding the sustainability of cooperatives. Cooperatives are unique in that the typical principal-agent frameworks used for understanding optimal contracts, and incentive mechanisms in incomplete contract theory, do accurately represent most cooperative relationships. This is because cooperative members can be both the principal and agent in contractual relationships with the managers of the cooperative. For example, a cooperative member/owner is a principal that is allowed to ratify manager (agent) proposals, but cooperative members can also be an agent when supplying products or services to the cooperative. The dual objectives of the managers and cooperative member owners complicates the modeling for optimal incentive schemes to maintain desirable agent action when action is hidden. Thus, the design of optimal cooperative contracts are mired in complexities that are not well understood or are expected to remain stable at an optimal equilibrium. Thus, sustainability of cooperative in the third approach is defined at how effective cooperatives can control agency issues in exhaustive bargaining in incomplete contract relationships, particularly as member and manager heterogeneity increases where incentives may not be aligned and complexity of strategic interaction increases.

In all three approaches, membership heterogeneity plays an important role in determining the sustainability of cooperatives overtime.

3. Materials and Methods

To analyze the relationship of cooperative membership heterogeneity on cooperative sustainability we obtained annual cross sectional data from the USDA-ARMS survey on U.S. farm producers and data from the USDA-Rural Business Development on cooperatives. We combined these datasets by aggregating the ARMS data to the state level and joining the data with USDA-Rural Business development data on farmer cooperative members that is made publically available at the state level. Since 1996, the USDA-ARMS data has annually surveyed a sub sample of farm producers (approximately 10,000 to 30,000 a year) in each state, across all types of farms, to collect financial and demographic information on farm businesses and households to maintain suitable farm benchmarks. In each survey, they ask the respondent to indicate the amount of cooperative patronage and cooperative equity they have received in that year or possess as a part of their financial and asset information. We coded each producer who reported receiving patronage or possessing equity as a cooperative member and ones who didn't as a non-cooperative member. We then estimated the mean and variance of cooperative members in each state compared to non-cooperative members over a number of other variables that are in the ARMS dataset. The variables chosen have all been previously identified in the cooperative literature as being important for understanding membership heterogeneity issues with cooperatives.

Hohler and Kuhl [2] quantified the number of publications that identify different dimensions of membership heterogeneity as being important in understanding impact to cooperative sustainability. The most cited dimensions they found were farm size, type of product, age, location, and education in that order. In our study, we estimate farm size and diversity by the mean farm asset value (C1_ATOT_mean) and coefficient of variation of asset value by cooperative members (C1_CV_ATOT). We also estimate farm size by the mean (C1_Acres_mean) and coefficient of variation acres (C1_CV_Acres) that is operated. We estimate diversity of farm type by two survey questions, one whether the farm is reported to be primarily grain or livestock (C1_Farmtype_stddev), and then a more detailed question concerning what type of product represents the largest portion of the operation's gross income. In this question there are 16 enterprise types that range from grains and oilseed production to equine and aquaculture (C1_typefarm_stddev). We estimate location of the cooperative member's state by the sub region (Sub_region), where there is 9 sub regions in the U.S., and by the longitude (Centroid_x) and latitude (Centroid_y) of the center of the state. Finally, we score the diversity and mean level of cooperative member education (C1_cv_education), as well as income (C1_dv_finci), age (C1_cv_agecls_stddev), and amount of other business income (C1_cv_busi) in a socioeconomic variable using factor analysis. The factor incorporates the means (SES) and coefficients of variations (SES_diversity) of age, income, and education variables. We also include producer and consumer price indexes that are derived by the Bureau of Labor and Statistics (BLS). Values we include are the farm producer price index (farm_ppi), the food and feed producer price index (food_feed_ppi), the urban consumer price index for the U.S. (cpi), and the commodity producer price index (comdty_ppi). All producer price indexes have a base year of 1982. In addition, we joined value added estimates and quantity indexes at the state level for industries that are derived by the BEA to estimate U.S. and State GDP. The BEA's industry identificaiton for the Agricultural, Forestry, Farming and Fishing Industry is 3, thus our coded variable for value added in this industry in the dataset is "Ind_va_3". Values for the farms alone (Ind_va_4), and food and kindered products (Ind_va_20) are available as well. We included the quantity indexes estimated by BEA as "Ind_q_" and value added per quantity as "Ind_va_q_".

To calculate the variables means and variances, and to measure the socioeconomic factor we used SAS 9.4. We then used the *randomForest* package in R (see Appendix for R code) (Liaw and

 doi:10.20944/preprints201804.0022.v1

Wiener, 2002) to generate the random forest model, assess variable importance, and plot the marginal effects of membership heterogeneity on the response variables that indicate cooperative sustainability. We derived variable importance plots and main and interaction plots using Welling et al.'s [4] R package *forestFloor*.

The target variables (dependent) that we associate with cooperative sustainability and we predict in the random forest regressions consists is the data available from the USDA Rural business development on cooperative businesses at the state level. The data includes the sum of cooperative gross business volume by cooperatives headquartered in a state (gbv), we deflate using the commodity producer price index (gbv_dfl), cooperative net business volume in the state (net_gbv), deflated (net_gbv_dfl), the sum of the number of cooperative members (members), and sum of cooperatives that are headquartered in the state (coops_num).

The decision trees that are randomly generated that make up a random forest model are distinct rule sets that comprehensively describe how to interpret attributes of inputs to make a prediction for a target value. In this case, an ensemble of decision trees create rules to examine attributes of cooperative members in a state, industries in the state, the state and/or region itself, and predict the amount of cooperative business volume, number of cooperatives, and the number of cooperative members. Each decision-tree is expected to be biased, but when included in an ensemble of decision trees and aggregated is expected to enhance predictive accuracy. The training data to develop the rule sets are the observed patterns given the input values. Accuracy assessments are done by cross validating against observations not included in the training set. The decision trees are uncorrelated because of preselected parameters that constrain what inputs variables can be used at each decision node split, and through a random selection of the training data to allow different slices of the data to inform the prediction. For example, a hypothetical decision tree for cooperatives headquartered in a state, for one random draw of the training data, may be a simple, single decision rule such as: "if the state is in the East North Central Region then then the number of cooperatives will be x, else the number of cooperatives will be y". Another random sample of the training data could result in a distinct decision tree with a similar rule but with a different cutoff for region (e.g. if state is in the West North Central Region or East North Central Region then x, else y), or in combination with another variable (e.g. if in the East North Central Region and if the number of farms in the state is greater than 20,000 then x, else y). A third tree may not have randomly selected the region variable to be used in a node split, thus the tree may consist of a split based on the number of farms alone. In random forest prediction, the final prediction is the mode of the predictions of many distinct, decision trees generated, and the confidence in the prediction is indicated by the frequency the set of decision trees result in the same prediction relative to the total number of decision trees generated.

After we omitted observations from our dataset with missing data, the number of observations we were left in the analysis was approximately 580 cross sectional observations of cooperative numbers, business volume, and attributes related to cooperative membership heterogeneity, and other control variables. To generate a number of uncorrelated decision trees, but also maintain stability in our importance and main effects of variable rankings, we generated 800 trees to make up the random forest model. The number of observations from our data that we randomly drew from at each tree to train the regression trees was 200. We restricted the number of variables that were randomly selected to optimally split a node in the tree to be 4. The selections of the number of trees, training data, and number of variables to try at a node split is somewhat arbitrary, though there are general rules of thumb given the number of variables and number of observations being analyzed. As the number of trees that are generated increase, the predictions and accuracy is expected to converge. Thus importance rankings and marginal effects are not expected to be sensitive to the preliminary selections, but may affect values of the measures used to rank the variables including the percent of accuracy increase in the mean square error and the gini node purity index [4-5].

 doi:10.20944/preprints201804.0022.v1

3. Results

3.1. Prediction of Number of Cooperatives Headquarted in the State

The number of farmer cooperatives headquarted in each state has been in decline across all regions of the U.S. This is in part due to mergers, consolidations, acquisitions, and exits. However, some regions have expierienced greater rates of decline (See Figure 1). For example, the West North Central region and the East North Central Region have historically had the most cooperatives headquartered in the state, but have also observed the greatest rates of declines since 1979 as indicated by the loess slope in Figure 1.

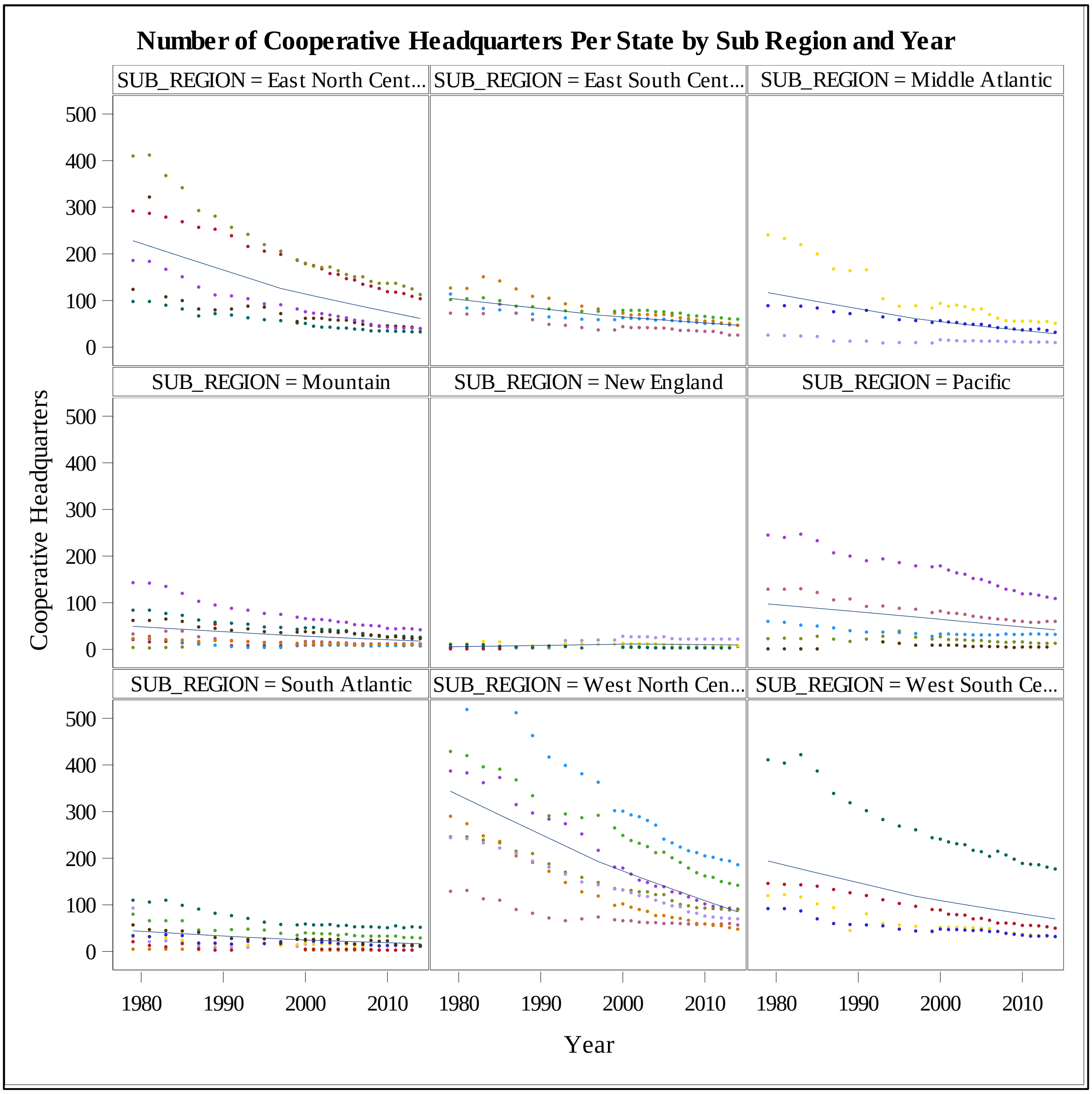

Figure 1. Source: USDA Rural Business Cooperative Statistics

When we analyze the variables that were most important in explaining the number of cooperatives headquartered in the state over this period, regional variables, as well as the number of farms and the number of acres, or milk production tended to be ranked the highest (See Figure 2). What is of interest in this study, however, is the expected effect of the lesser variables related to membership heterogeneity that help in providing explanatory power unrelated to the effects of the more important variables. Specifically, do indicators of membership heterogeneity provide a significant main or interaction effect to cooperative numbers in a state? In the case of number of cooperatives headquartered in a state we find that socioeconomic status (SES) of cooperative members and the coefficient of variation of the age class of cooperative members appears to be relatively important. Perhaps even more important than the diversity of farm size—measured by acres—of cooperative members (c1_cv_acres), for example.

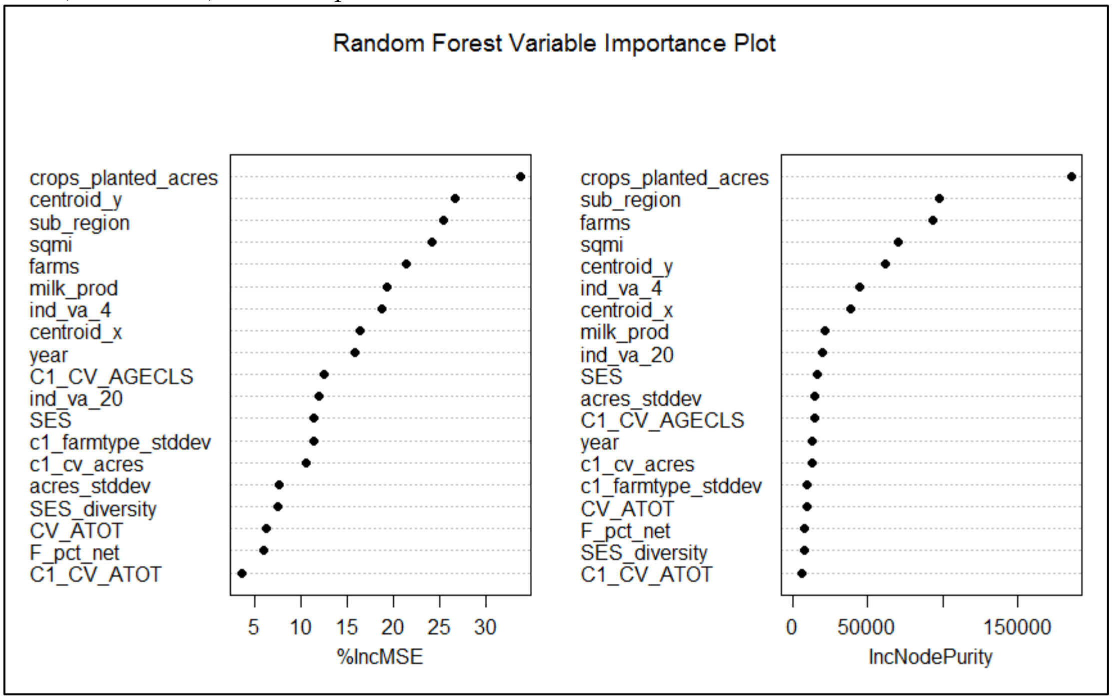

Figure 2. Random Forest Variable Importance Plot for Number of Cooperatives Headquartered in a State

To further examine the main effects of membership heterogeneity we can plot the main effects of the number of cooperatives headquartered in a state using the *forestFloor* package. With the forestFloor package, we can also observe what interactions effect may occur with other variables in the analysis by examining the color gradient. For example, the main effect of cooperative member's age class coefficient of variation on number of cooperatives indicates that states where cooperative members have greater age diversity (.3 to .4 on the x axis), the random forest model predicts slightly more cooperatives (y-axis) *cetaris paribus*. Specifically, the forest predicts up to 5 more cooperatives in a state when the member coefficient of variation of age class is .35 (See Figure 3). The number of cooperatives predicted is indicated on the y-axis in the model. This value is additive to other cooperative number predictions in the model. Thus the total number of cooperatives predicted by the random forest model would include a sum of all the marginal effect values. For example, when cooperative member coefficient of variation score is .2, indicating less age diversity, the random forest model expected no change in the number of headquarters in a state compared to the other variables in the model. This is illustrated by gray main effects line for certain values of the coefficient of variation of age class in Figure 3 that is near 0.

The color gradient in Figures 3 and 4 was determined by what sub region the observation was in. Blue in Figure 3, for example, indicates the observation was in the South Atlantic, West North Central, or West South Central regions. It appears in these regions the diversity of age effect was more pronounced than in other regions of the U.S. Thus when the age coefficient of variation was increased, coupled with being in the South Atlantic, West North Central, or West South Central Regions, the model predicted more cooperatives headquartered in the state. The 12 most important main effects are plotted in the same fashion as Figure 3 in Figure 4. In the lower right panel of Figure 4 (4c), the plot on the x-axis is the standard deviation of cooperative member's farm type, for example. On the y-axis is the number of cooperative headquarters predicted in the state. The random forest model predicts that states where cooperative members have greater variations in farm types there

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 April 2018 doi:10.20944/preprints201804.0022.v1

was a slight expectation of more cooperatives headquartered in the state *cetaris paribus*. Though this effect was not very pronounced and could be largely disregarded as nearly meaningless.

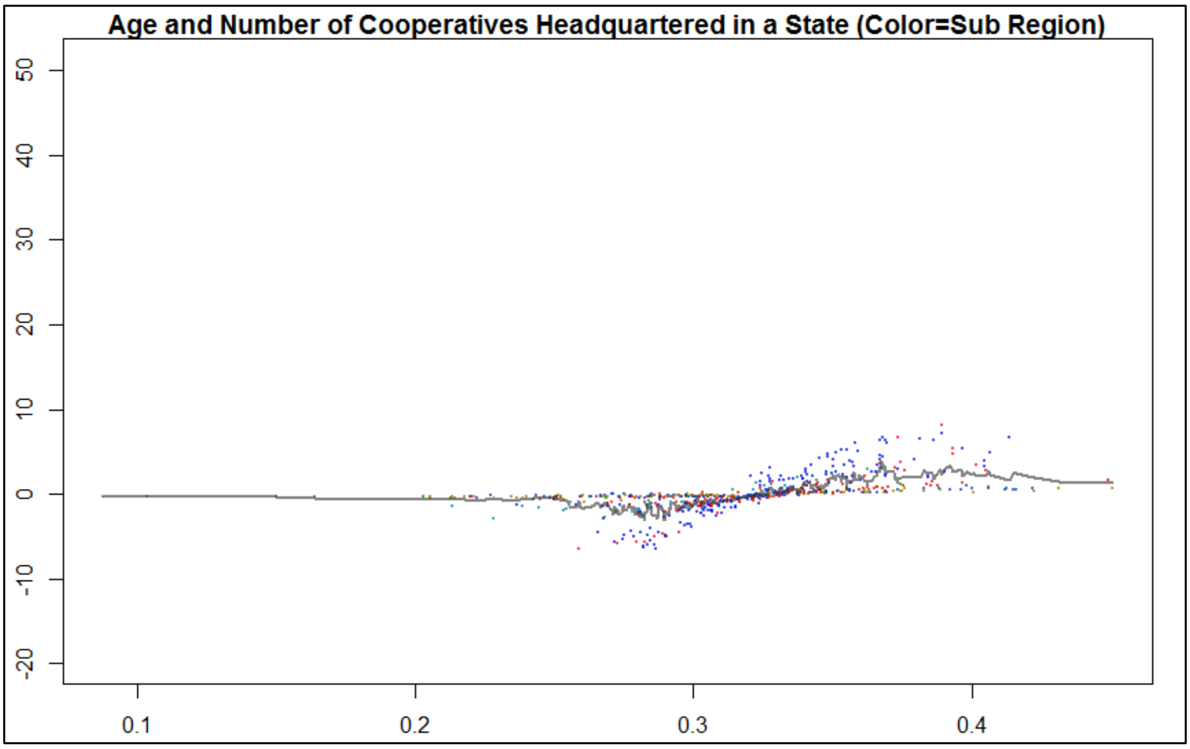

Figure 3. Coefficient of Variation of Farmer Cooperative Member Age Class (x-axis) and the Predicted Number of Cooperatives Headquartered in a State (y-axis).

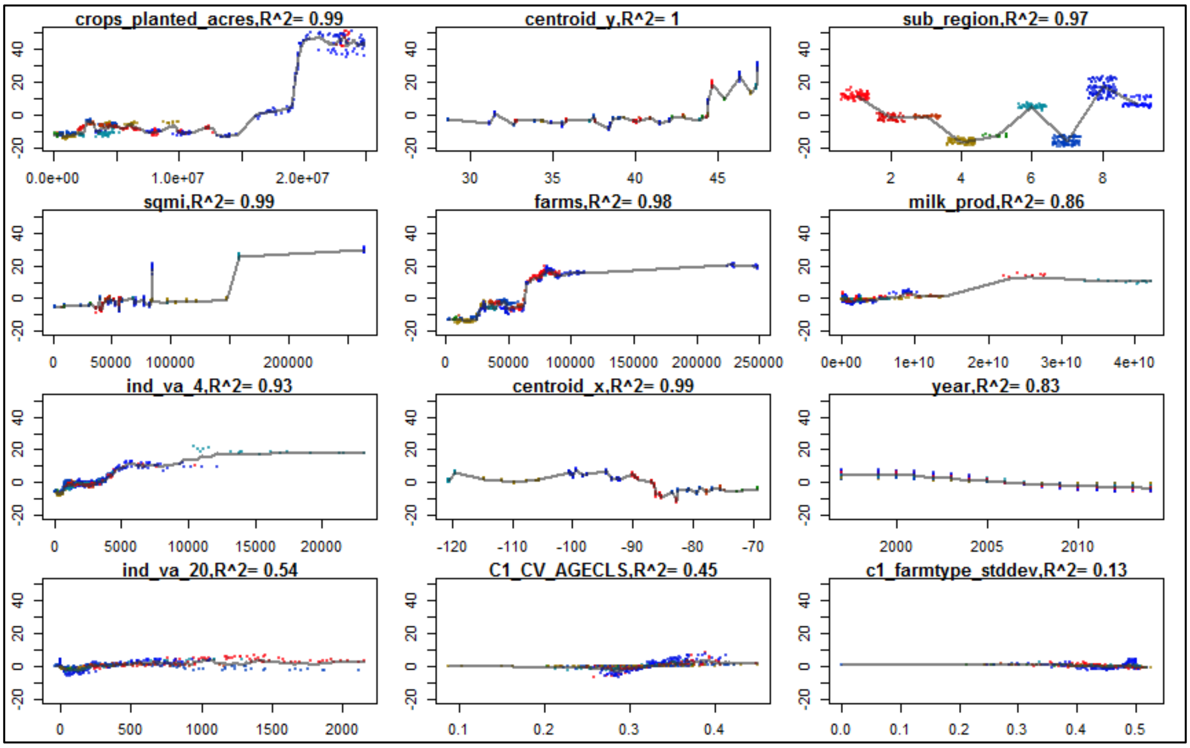

Figure 4.

Panel (1a). The number of acres planted to field crops in a state (x-axis) and cooperative headquarters predicted (y-axis).

Panel (1b). The latitude of the center of a state (x-axis) and cooperative headquarters predicted (y-axis)

Panel (1c). The sub region the state is in and the number of cooperative headquarters predicted. X-axis values correspond to `1=East North Central, 2=East South Central, 3=Middle Atlantic, 4=Mountain, 5=New England, 6=Pacific, 7=South Atlantic, 8=West North Central, and 9= West South Central.

Panel (2a). The square miles of a state (x-axis) and the number of cooperative headquarters predicted (y-axis).

Panel (2b). The number of farms (x-axis) in a state and the number of cooperative headquarters predicted (y-axis).

Panel (2c). The amount of fluid milk produced (x-axis) in a state and the number of cooperative headquarters predicted (y-axis).

Panel (3a). The amount of value added (millions of US Dollars) by farms (x-axis) in a state and the number of cooperative headquarters predicted (y-axis). Source: U.S. Bureau of Economic Analysis.

Panel (3b). The longitude of the center of state (x-axis) and the number of cooperative headquarters predicted (y-axis).

Panel (3c). The year (x-axis) and the number of cooperative headquarters predicted (y-axis).

Panel (4a). The value added by the food and kindered products industry (millions of U.S. Dollars) (x-axis) and the number of cooperative headquarters predicted in a state (y-axis). Source: U.S. Bureau of Economic Analysis.

Panel (4b). The coefficient of variation of cooperative member age class (x-axis) and the number of cooperative headquarters predicted (y-axis).

Panel (4c). Farm type diversity (x-axis) and the Predicted Number of Cooperatives Headquartered in a State (y-axis).

3.2. Prediction of the Number of Cooperatives Members in a State

The number of farmer cooperative members have also decreased in a similar fasion to the number of coopertives that are headquaretered in a state (See Figure 5). The decreasing trends in cooperative members have closely followed trends of larger, more consolidated farms and greater labor productivity from technical advances in agricultural production. However, in some regions cooperative members have been more resistent to the declining trend. For example, in the New England region, cooperative membersship has not declined at the rapid rate observed in the Mountain region.

When we predicted cooperative membership by state in the random forest model we found the most important variables were similar to the variable importance rankings when we predicted cooperative headquarters. Theese included included regional variables and the number of farms in the state overall (See Figure 6). Again, the variables associated with membership heterogeneity that improved prediciton performance the most in the random forest model was the mean socioeconomic (SES) status of the cooperative members in the state, and coefficient of variation of age class (C1_cv_agecls). One variable that was was more predictive in cooperative membership than was predictive in the headquarters model was the standard deviation of acres of non-cooperative members (acres_stddev).

Similar to the number of cooperatives headquartered in a state, when socioeconomic status of cooperative members was below average, the model expected more cooperative members. Specifically, a mean cooperative member socioeconomic status of negative one was expected to increase state cooperative membership by approximately 8,000 members *cetaris paribus* (See Figure 7). Other important variables for predicitng cooperative membership in the state can be observed in Figure 8. Notably, cooperative membership is expected to particularly high in the North Central regions of the U.S., and when ratio of net income to gross cash sales is particularly low.

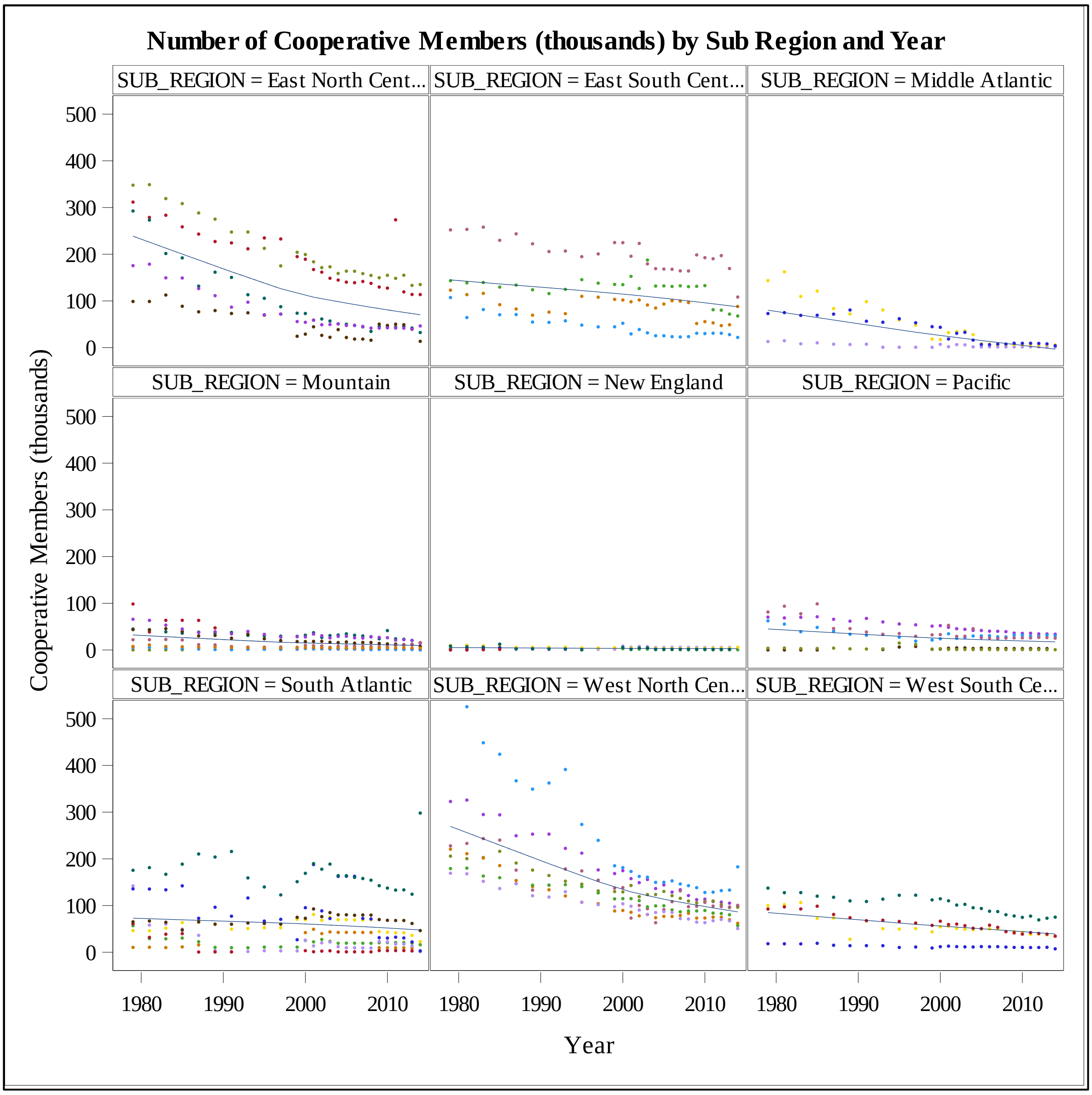

Figure 5. Number of Farmer Cooperative Members (thousands) by Sub Region and Year. Source: USDA Rural Business Cooperative Statistics

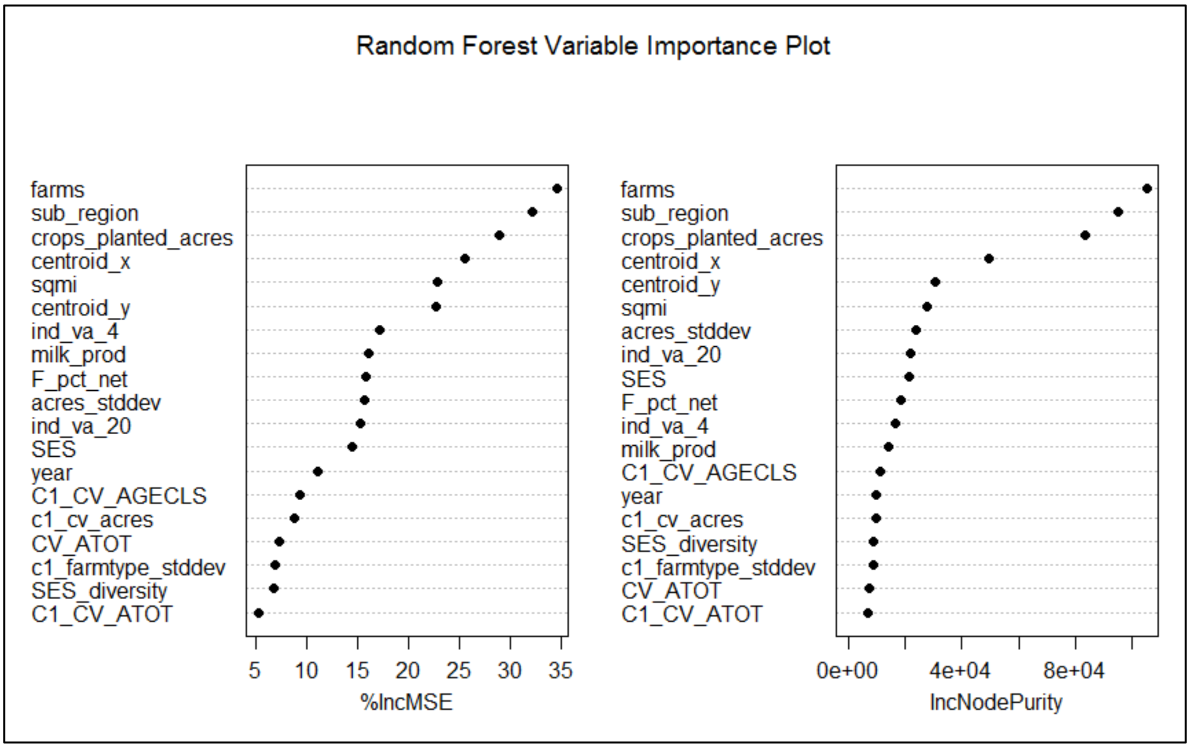

Figure 6. Random Forest Variable Importance Plot for Number of Cooperatives Members in a State

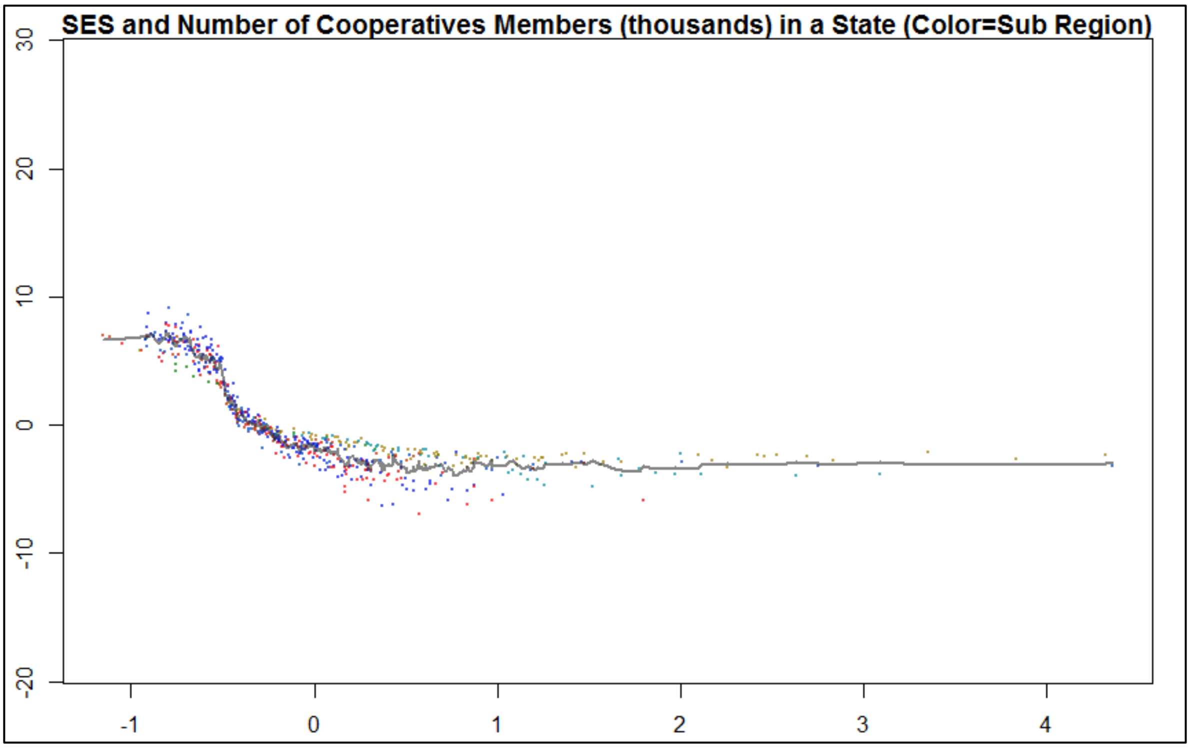

Figure 7. Socioeconomic status of Farmer Cooperative Members (x-axis) and the Predicted Number of Cooperatives Members in a State (y-axis).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 April 2018 doi:10.20944/preprints201804.0022.v1

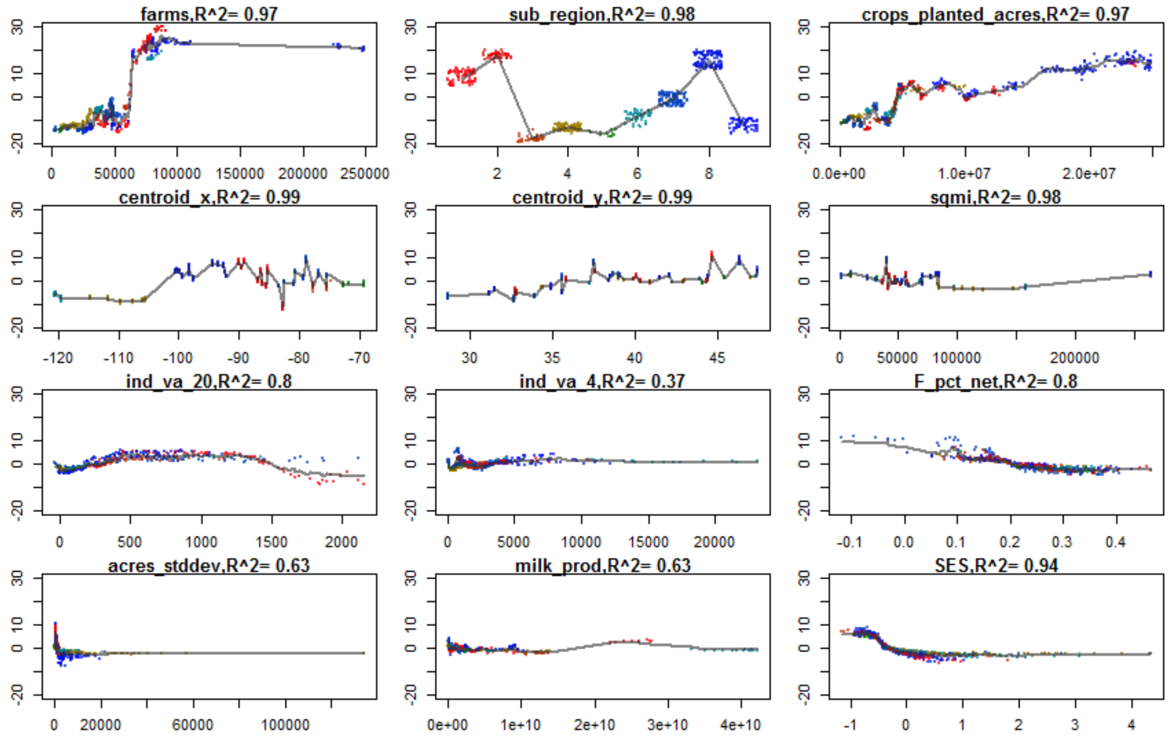

Figure 8.

Panel (1a). The number of farms in a state (x-axis) and cooperative members predicted (y-axis).

Panel (1b). The sub region the state is in and the number of cooperative members predicted. X-axis values correspond to,1=East North Central, 2=East South Central, 3=Middle Atlantic, 4=Mountain, 5=New England, 6=Pacific, 7=South Atlantic, 8=West North Central, and 9= West South Central.

Panel (1c). The number of field crop acres planted (x axis) and predicted cooperative membership (y-axis).

Panel (2a). The longitude of the center of state (x-axis) and the number of cooperative members predicted (y-axis).

Panel (2b). The latitude of the center of a state (x-axis) and cooperative members predicted (y-axis)

Panel (2c). The square miles of a state (x-axis) and the number of cooperative members predicted (y-axis).

Panel (3a). The amount of value added (millions of US Dollars) by the food and kindered products industry (x-axis) in a state and the number of cooperative members predicted (y-axis). Source: U.S. Bureau of Economic Analysis.

Panel (3b). The amount of value added (millions of US Dollars) by farms (x-axis) in a state and the number of cooperative members predicted (y-axis). Source: U.S. Bureau of Economic Analysis.

Panel (3c). The ratio of net farm income to gross cash farm sales in a state and the number of cooperative members predicted (y-axis). Source: Bureau of Economic Analysis.

Panel (4a). The standard deviation of farm acres of non-cooperative members (x-axis) and the number of cooperative members predicted (y-axis).

Panel (4b). The amount of fluid milk produced (x-axis) in a state and the number of cooperative members predicted (y-axis).

Panel (4c). Socioeconomic status of Farmer Cooperative Members (x-axis) and the Number of Cooperatives members predicted in a State (y-axis).

3.3. Predicted Deflated Cooperative Gross Business Volume in a State

Unlike the decreasing trends seen in cooperative headquarters and cooperative members per state, the deflated cooperative gross business volume has seen a revival in recent years. Specifically, in the East North Central Region, the West North Central Region, the West South Central Region, the Mountain Region, and the Pacific region have shown a"U" shaped pattern to gross business volume (See Figure 9).

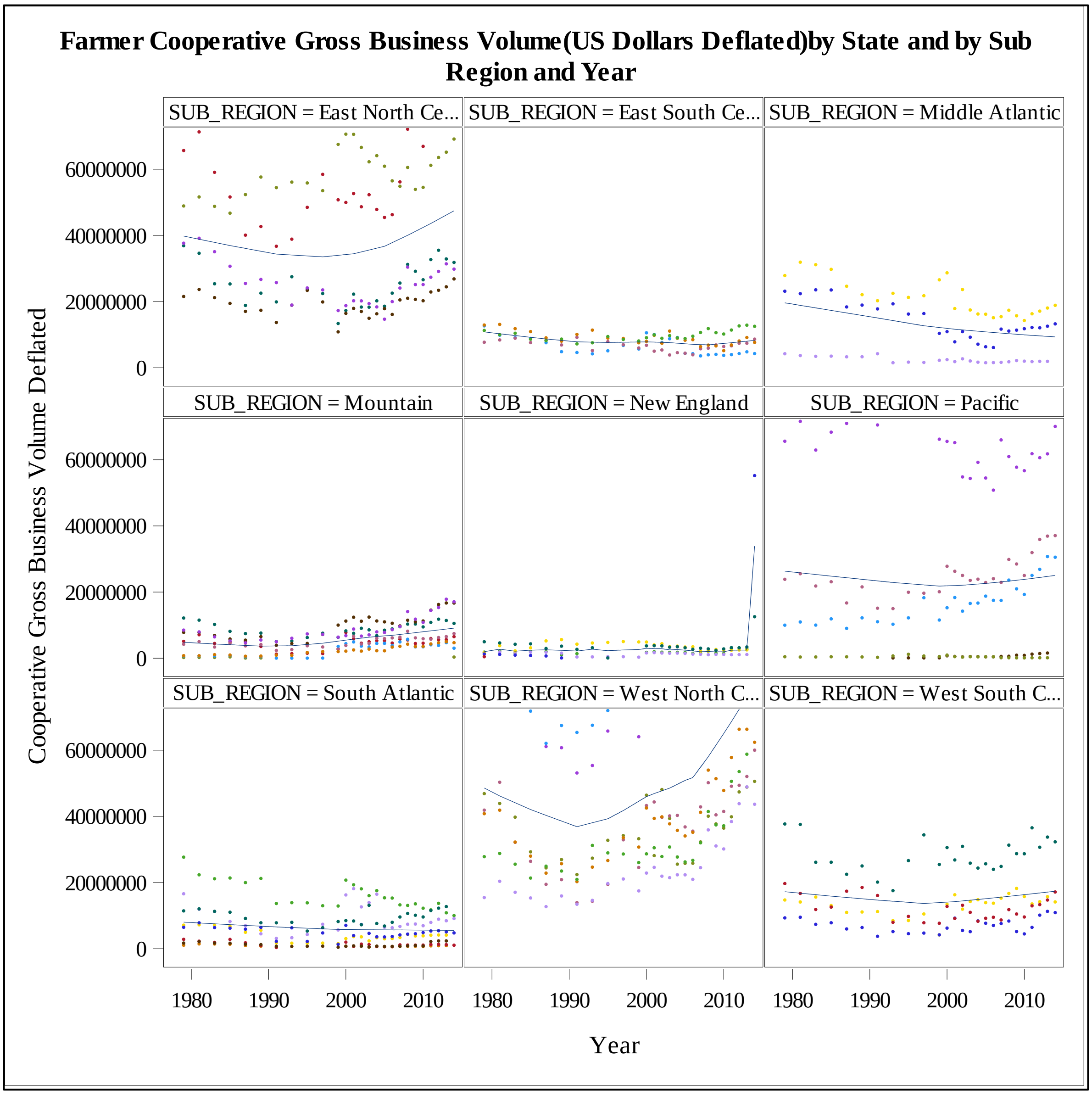

Figure 9. Deflated Cooperative Gross Business Volume (Base Year=1982). Source: USDA Rural Business Cooperative Statistics

Accordingly, the random forest model that was generated found the sub region to be the most important variable in predicting deflated gross business volume at the state level. Variables of membership heterogeneity were found to be less important in predicting deflated gross business volume in the state relative to their importance in predicting the cooperative numbers and members (See Figure 10).

Perhaps the most interesting finding in predicting gross business volume was the importance of the amount of value added dollars by the farm industry (Ind_va_4). Specifically, value added by farms was ranked third in importance under both measures (mse and node purity) in improving the understanding of cooperative gross business volume at the state level (See Figure 10). Indeed, the amount of value added the farm industry contributes can have a strongly significant impact on the amount of cooperative business volume that exists according to the random forest model. Specifically, as the farm industry expects to add value to the state GDP in the order of 5 billion current U.S. dollars, the corresponding expected value of deflated (1982 based) cooperative gross volume at the state level increases to 10 million dollars (See Figure 11).

The variables mostly associated with membership heterogeneity appeared to play a lesser role in understanding cooperative gross business volume at the state level overall. Though in certain regions, the mean level of socioeconomic status (SES) of cooperative members appears to play a slight role in increasing gross business volume (See Figure 12) .

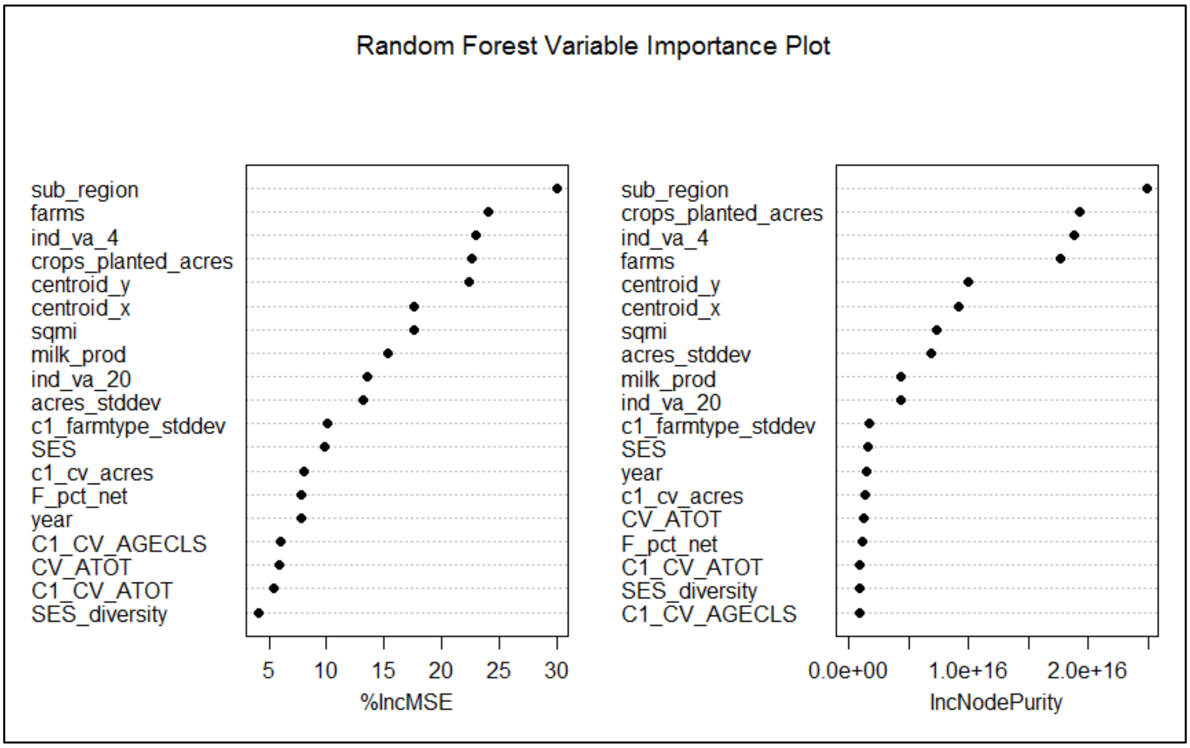

Figure 10. Random Forest Variable Importance Plot for Deflated Gross Business Volume of Cooperatives in a State

 doi:10.20944/preprints201804.0022.v1

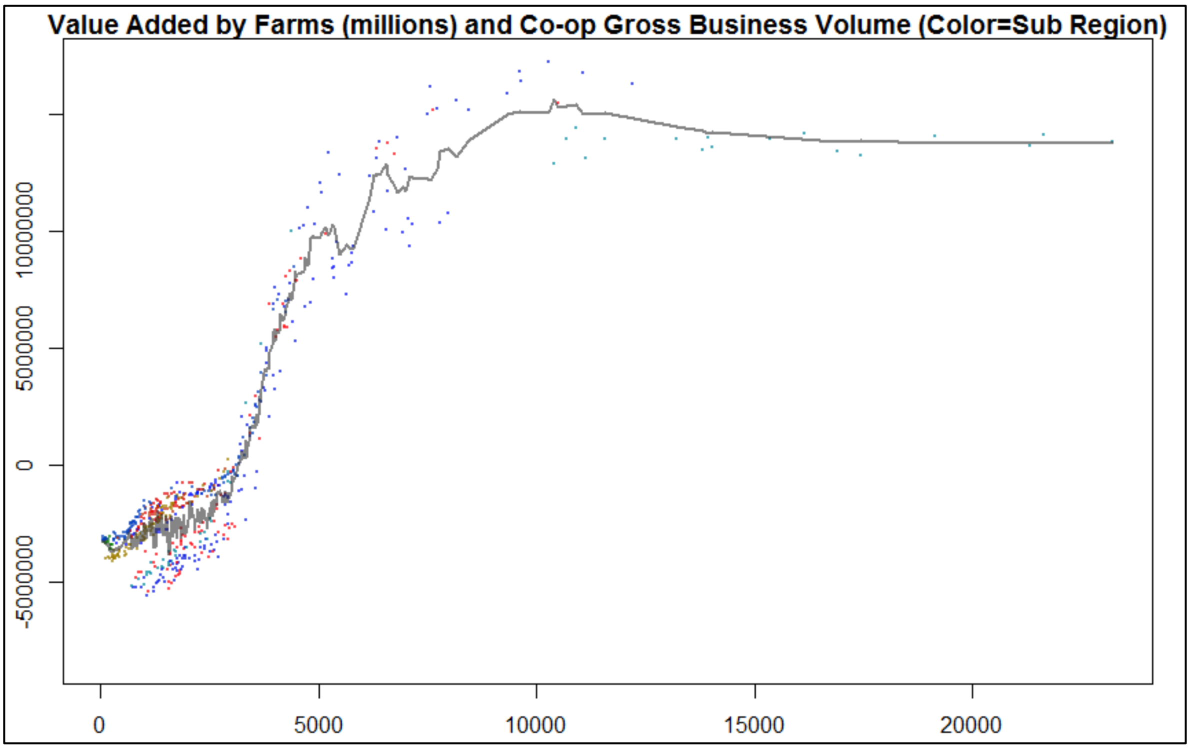

Figure 11. State value added by Farms in millions of current U.S. dollars (x-axis) and deflated cooperative business volume in 1982 dollars.

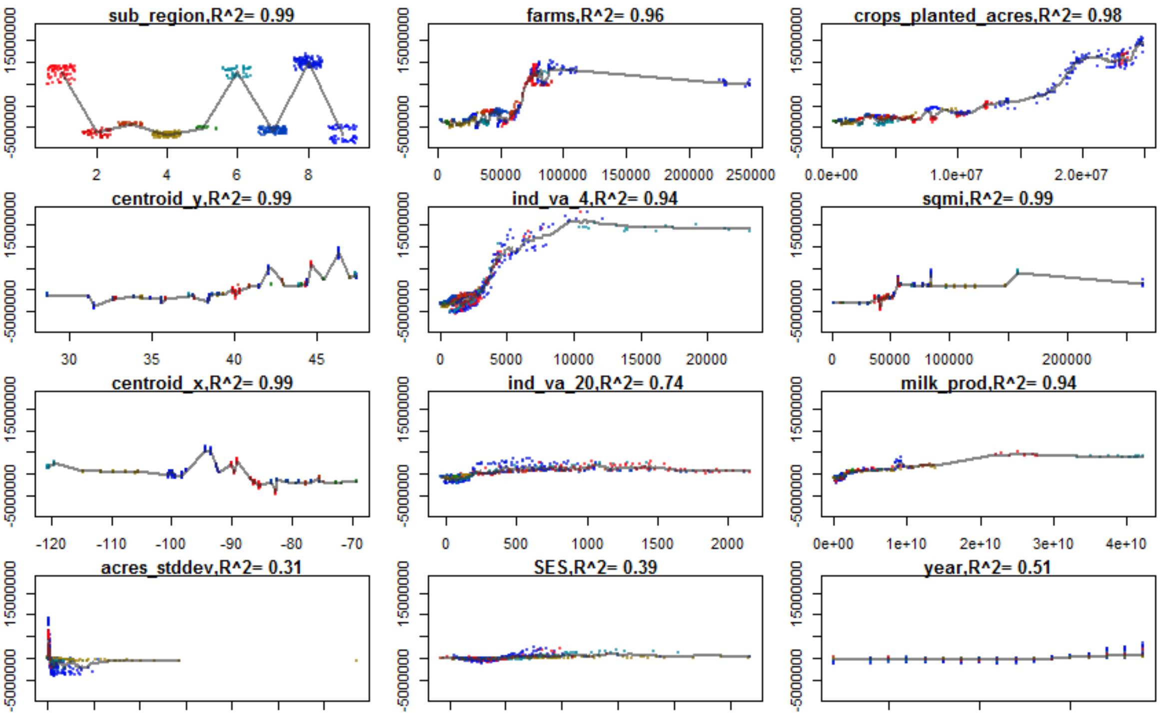

Figure 12.

Panel (1a). The sub region the state is in and the deflated cooperative business volume predicted. X-axis values correspond to,1=East North Central, 2=East South Central, 3=Middle Atlantic, 4=Mountain, 5=New England, 6=Pacific, 7=South Atlantic, 8=West North Central, and 9= West South Central.

Panel (1b). The number of farms in a state (x-axis) and cooperative gross business volume predicted (y-axis).

Panel (1c). The number of field crop acres planted (x axis) and the cooperative gross business volume predicted (y-axis).

Panel (2a). The latitude of the center of a state (x-axis) and the cooperative gross business volume predicted (y-axis)

Panel (2b). The amount of value added (millions of US Dollars) by farms (x-axis) in a state and the cooperative gross business volume predicted (y-axis). Source: U.S. Bureau of Economic Analysis.

Panel (2c). The square miles of a state (x-axis) and cooperative gross business volume predicted (y-axis).

Panel (3a). The longitude of the center of state (x-axis) and cooperative gross business volume predicted ((y-axis).

Panel (3b). The amount of value added (millions of US Dollars) by the food and kindered products industry (x-axis) in a state and the cooperative gross business volume predicted (y-axis). Source: U.S. Bureau of Economic Analysis.

Panel (4a). The standard deviation of farm acres of non-cooperative members (x-axis) and cooperative gross business volume predicted (y-axis).

Panel (4b). Socioeconomic status of Farmer Cooperative Members (x-axis) and the cooperative gross business volume predicted in a State (y-axis).

Panel (4c). The year (x-axis and cooperative gross business volume predicted in a State (y-axis).

4. Discussion

In this study, we empirically examined the effects of cooperative membership heterogeneity on cooperative sustainability in U.S. farmer cooperatives. We found that membership heterogeneity was expected to affect the number of cooperatives headquartered in a state and the number of cooperative members. However, membership heterogeneity was found to be less important in understanding cooperative gross business volume at the state level. The findings of this empirical work reconciles the notion that cooperatives can be sustainable in the long-term despite much of the recent focus of cooperative literature on intra-cooperative issues that may arise due to cooperative membership heterogeneity. Moreover, we find that cooperative member heterogeneity may play an important role in decreasing the rate of consolidation and acquisition of cooperatives in the U.S. This finding raises new questions about what is efficient, and more sustainable long-term, and whether cooperative membership numbers and number of cooperatives should be the objective measure of cooperative sustainability.

Future research should continue to pursue a greater theoretical and empirical understanding to intra-cooperative issues related to membership heterogeneity. More detailed data sets and new empirical methods may allow us to parse the effect of cooperative membership heterogeneity on cooperative sustainability more precisely. This study provides an initial expectation and understanding of how membership heterogeneity may affect cooperative sustainability long-term.

Acknowledgments: This study was partly funded by a USDA NIFA Hatch grant, accession No. 1006890.

References

1. Cook, Michael Lee, Fabio Ribas Chaddad, and Constantine Iliopoulos. "Advances in cooperative theory since 1990: A review of agricultural economics literature." Erasmus University Rotterdam, Rotterdam School of Management, 2004.
2. Höhler, Julia, and Rainer KÜHL. "POSITION AND PERFORMANCE OF FARMER COOPERATIVES IN THE FOOD SUPPLY CHAIN OF THE EU-27." *Annals of Public and Cooperative Economics* 85.4 (2014): 579-595.
3. Breiman, Leo. "Random forests." *Machine learning* 45.1 (2001): 5-32.
4. Liaw, Andy, and Matthew Wiener. "Classification and regression by randomForest." *R news* 2.3 (2002): 18-22.
5. Caruana, Rich, and Alexandru Niculescu-Mizil. "An empirical comparison of supervised learning algorithms." *Proceedings of the 23rd international conference on Machine learning*. ACM, 2006.
6. Wager, Stefan, and Susan Athey. "Estimation and inference of heterogeneous treatment effects using random forests." *Journal of the American Statistical Association* just-accepted (2017).
7. United States Department of Agriculture, Rural Development. Annual Farmer, Rancher, and Fishery Cooperative Statistical Data. United States Department of Agriculture: Washington, DC, USA, 2017; Available online: https://www.rd.usda.gov/programs-services/all-programs/cooperative-programs (accessed on 29 March 2018).
8. U.S. Bureau of Economic Analysis, "NAICS Gross Operating Surplus by State" https://www.bea.gov/regional/downloadzip.cfm (accessed March 21st, 2018).
9. Feinerman, Eli, and Meira S. Falkovitz. "An agricultural multipurpose service cooperative: Pareto optimality, price-tax solution, and stability." *Journal of Comparative Economics*15.1 (1991): 95-114.
10. Choi, E. Kwan, and Eli Feinerman. "Regulation of nitrogen pollution: taxes versus quotas." *Journal of Agricultural and Resource Economics* (1995): 122-134.
11. Sexton, Richard J. "Imperfect competition in agricultural markets and the role of cooperatives: a spatial analysis." *American Journal of Agricultural Economics* 72.3 (1990): 709-720.
12. Zusman, Pinhas. "Constitutional selection of collective-choice rules in a cooperative enterprise." *Journal of Economic Behavior & Organization* 17.3 (1992): 353-362.
13. Buchanan, James M., and Gordon Tullock. *The calculus of consent*. Vol. 3. Ann Arbor: University of Michigan Press, 1962
14. Olson, Mancur. *Logic of collective action: Public goods and the theory of groups (Harvard economic studies. v. 124)*. Harvard University Press, 1965.