Determinants of Land Allocation in a Multi-Crop Farming System: An Application of the Fractional Multinomial Logit Model to Agricultural Households in Mali

James E. Allen IV, M.S.
Community and Economic Development Initiative of Kentucky
Department of Agricultural Economics
University of Kentucky
Email: james.allen4@uky.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2014 AAEA Annual Meeting, Minneapolis, MN, July 28, 2014.

ABSTRACT

Effective food security work in developing countries, such as Mali, relies on a thorough understanding of the rural farming system. A common approach is to study land allocation decisions to specific crops. In accomplishing this, one challenge is to model all production outcomes in a multi-crop system. This paper attempts to overcome this challenge to study the determinants of household allocation to cotton, maize, sorghum, millet, and secondary crops. First, a reduced form of the agricultural household model helps to identify factors that explain land allocation to various crops. This framework is applied to survey data from six villages in Mali's Koutiala Cercle. A fractional multinomial logit econometric model is used to estimate the effect of household and production attributes on shares of cotton, maize, sorghum, millet, and secondary crops simultaneously, the results of which are presented as average marginal effects. Among other results, the analysis shows that ethnic groups not native to the Koutiala Cercle are associated with significantly smaller shares of maize, and that villages with better market access are correlated with much higher shares of secondary crops and smaller shares of cotton. These results provide insights for policymakers on the role of cotton in farming system, the need to promote and develop better markets for coarse grains and secondary crops, and the importance of understanding the dynamic farming system in Mali's Koutiala Cercle.

I. INTRODUCTION

Effective food security work relies on a thorough understanding of the local farming system. A common approach in farming systems research is to study determinants of land allocation to specific crops, but this is difficult in a multi-crop system where farmers choose more than two crops. When faced with this problem, some studies simply take data from a single- or dual-crop production system, ignore land allocation to other crops, or include other crops as explanatory variables, which risks simultaneity bias by assuming that farmers have a sequential decision-making process when allocating their land to various crops. Aware of the simultaneity problem, other studies simplify crop diversity into two categories, typically cash and staple crops (De Boer and Chandra 2001), or use a mix of econometric and linear programming models that rely on various assumptions (Ahn, Singh, and Squire 1981). Often, farming system research cannot predict crop-specific supply response in a multi-crop system, though production of three of more crops is observed throughout the developing world.

Comprehending farmers' land allocation decisions is an important step in advancing food security in Mali's cotton-dependent Koutiala Cercle, where households are largely subsistent, face greater market obstacles, and often grow cotton, maize, sorghum, millet, and other field crops. Indeed, the region has long struggled with food insecurity despite being in "the breadbasket of Mali" (USAID 2010). Moreover, many households in the Koutiala Cercle are reliant on Mali's cotton industry, which serves as the main source of income, inputs, and financial credit for up to 96 percent of households in some villages. Certainly, in a country where the majority of the population relies heavily on rain-fed agriculture for their livelihoods, it is critical that policymakers understand how and why farmers devote certain shares of their land to growing cotton, coarse grains (i.e., maize, sorghum, and millet), or other field crops. Therefore, this paper answers: what are the main determinants of land allocation to cotton, maize, sorghum, millet, and secondary crops for households in Mali's Koutiala Cercle?

Current work in Mali is being carried out by the Project to Mobilize Food Security Initiatives in Mali (PROMISAM), a joint USAID/Mali, Gates foundation, and Michigan State University program aimed at assisting Mali's Food Security Commission. Collectively, this group has produced the Mali Agricultural Sector Assessment (Staatz et. al. 2011) and an extensive paper about Mali's cereal markets since 1990 (Kelly et. al. 2012), both of which take a comprehensive approach to their analysis of Mali's agricultural sector. Specifically, one research activity under PROMISAM aims to analyze the relationship between Mali's cotton and coarse grain subsectors and particularly how participation in the cotton value chain affects coarse grain productivity and food security. The hypothesis that the cotton industry may affect coarse grain yields arises from the fact that farmers have improved access to credit and fertilizer through Mali's singular cotton parastatal the Compagnie Malienne de Developpement des Textiles (CMDT). Yet, field work by Boughton and Dembelé (2010) has noted that Mali is in "a process of transition from a cotton-cereal production system, where the cotton subsector facilitates access to fertilizer for cereals as well as cotton, to a cereal-based production system." These studies highlight the importance of continued farming systems research in Mali.

This article develops a model for studying determinants of land allocation in a multi-crop system and uses it to study the determinants of how much of a Malian household's cultivated land it allocates to cotton, maize, sorghum, millet, and secondary crops. First, the article will apply the agricultural household model to survey data from six villages in Mali's Koutiala Cercle. Then, a fractional multinomial logit econometric model is used to estimate the effect of household and production attributes on shares of cotton, maize, sorghum, millet, and secondary crops simultaneously, the results of which are presented as average marginal effects. As Mali's southern-most Sikasso Region—which contains the Koutiala Cercle—has long struggled with food insecurity with poverty rates 30 percent higher than any other region (USAID 2010), understanding how certain factors affect crop allocation of coarse grains and other crops can lead policymaker toward effective, evidence-based strategies for improving household food security.

II. AGRICULTURAL HOUSEHOLD MODEL

Factors affecting Koutiala farmers' land allocation decisions are modeled using a reduced form of the agricultural household model, which has the primary working hypothesis that while semi-subsistent households are rational, they do not necessarily aim to maximize profit. This is because these farms are not traditional businesses, but as partial consumers of what they produce, they seek to maximize household utility (Singh, Squire, and Strauss 1986). According to Dorward (2011), the main contributions of this model are representing the interactions between consumption and production decisions (i.e., non-separability) present in household decision-making among the rural poor and, as a result of later work, identifying the effect of market failures for labor, variable inputs, credit, and staple crops. In particular, the agricultural household model has been adopted and adapted to study a wide-scope of issues, including transactions costs and market participation (Omamo 1998; Barrett 2008; Goetz 1992), missing markets (de Janvry, Fafchamps, and Sadoulet 1991; Van Dusen and Taylor 2003), risk aversion (Fafchamps 1992; Hazell 1982), labor availability (Benjamin 1992), and credit access (Dorward 2011), to list a few.

A summary of the agricultural household model as adapted for use in this study of Mali's Koutiala production zone presented below:

$\text{Max } U = f(X, Y, H)$	*utility maximization, s.t.*
$P_xX + TC_{bx} + P_yY + TC_{by} \leq P_xQ - TC_{sx} - P_vV - TC_{bv} + \pi_y$	*budget constraint*
$Q = f(F, V, A, K, \sigma)$	*production constraint*
$H + F + O = T$	*time constraint*
$X, Y, Q \geq 0$	

As rational agents, poor farmers strive to maximize their expectation of future household utility, which is a function of leisure H, consumption of agricultural goods X, and consumption of non-agricultural goods Y. Utility is maximized subject to a budget constraint in which the cost of consumption, determined by prices P_x and P_y and transaction costs for buying TC_b, is less than or equal to profits from non-agricultural activities π_y and profits from crop production, which is equal to this value of production (output Q times P_x) minus the cost of inputs (V times P_v), the transaction costs for selling agricultural goods TC_{sx}, and buying agricultural inputs TC_{bv}. Also present is a production constraint, for which output is a function of farm labor F, other variable inputs V, land area A, capital K, and risk σ. Household choices are further restricted by a time constraint, which limits the sum of farm labor F, leisure H, and off-farm labor O to be less than the household's total time endowment T. Note that all labor is assumed to come exclusively from the household, and so does not factor into the budget constraint. Finally, we assume that X, Y, and Q are non-negative values.

The research question asks which factors affect land allocation of standard Malian crops in the country's cotton-growing region. Land allocation is studied as a main determinant of agricultural production, represented by the variable Q. While other household choices (e.g., fertilizer usage) affect total production, most supply response models use land area as a proxy for total production, since the two are highly correlated and because agricultural production is difficult to measure (Askari and Cummings 1977). Thus, here, Q will represent a vector of the portfolio of crops chosen and the share of land devoted to those crops.

Within the model, four different types of variables influence the choice of Q. First, variables that may lead households to specialize in the production of a particular crop include the availability of farm labor F, other variable inputs V, land area A, capital K, and price P. Second, variables that may discourage specialization and lead to diversification include risk σ and transaction costs TC, because these factors reveal how the reality of missing markets affects household expected utility. Third, as a result of missing markets, the agricultural household model also predicts that household characteristics affecting tastes and preferences—such as ethnic group and family structure—will directly influence land allocation decisions. Finally, a fourth set of variables includes profits from other activities π_y, such as livestock raising or carpentry, that represent alternative ways for the household to earn profit and seek more utility. Since the decision to engage in other income-generating activities requires an assessment of household inputs and future consumption needs, profits from other activities π_y may represent other outcome variables in a model seeking to maximize utility. However, if the decision to participate in other income-generating activities is made in conjunction with the choice to grow agricultural products, then they are, by definition, simultaneous outcomes. Recognizing a possible simultaneity bias, I will not represent other income-generating activities in the empirical model due to data limitations, though I recognize it here in the conceptual model as a possible influence on land allocation decisions.

To conclude, according to the agricultural household model, crop choice and land allocation at the time of planting is a function of production factors, transaction costs, household characteristics, and alternative income-generating activities. Together, these have been reduced from the five equations of the agricultural household model into the following expression:

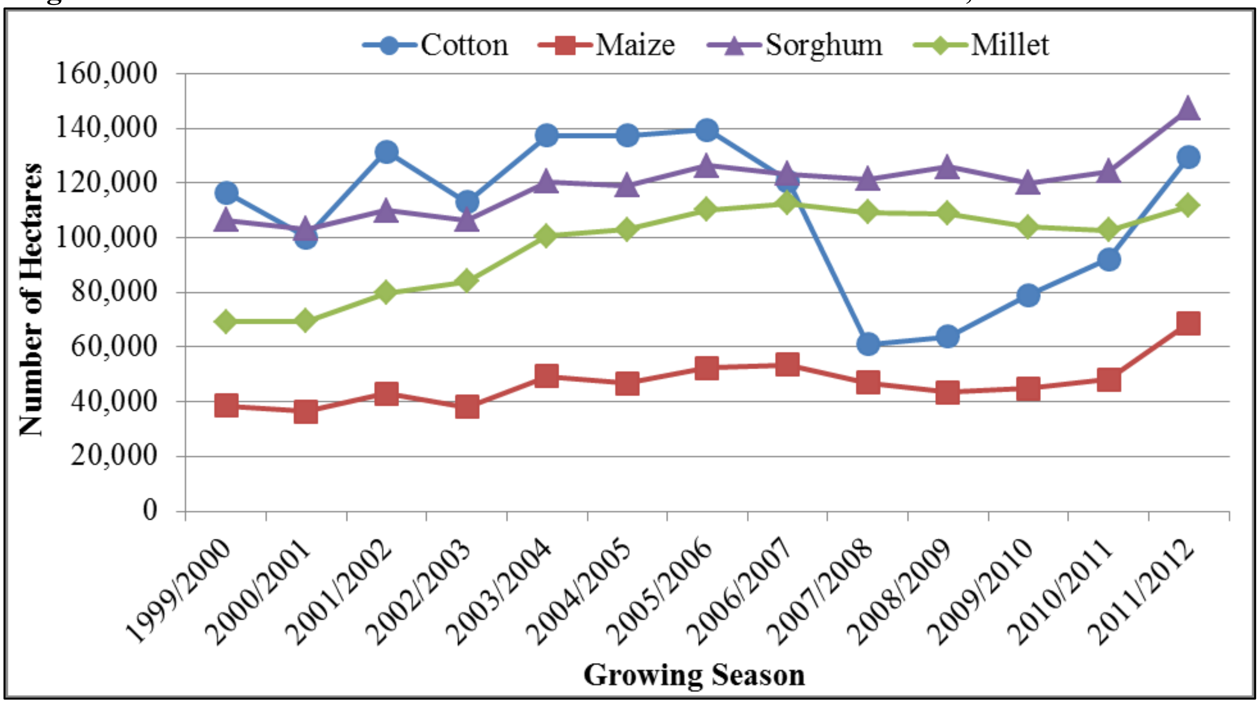
$Q = f(P, F, V, A, K, \sigma, TC, Hh, \pi y)$ *reduced-form land allocation function*

III. DATA DESCRIPTION

This study uses data from Mali's Koutiala Cercle, which is in the northeast of Mali's southern-most Sikasso region and shares a border with Burkina Faso. The Cercle's capital is the city of Koutiala, which is Mali's third most populous city and a regional trade hub. The Koutiala Cercle is known as the "Old Cotton Basin" and is home to the Minianka ethnic group.

In the Koutiala Cercle, the standard crop portfolio includes cotton, maize, sorghum, millet, and some secondary crops, such as peanuts, cowpeas, sweet potatoes, or vegetables. Using data collected annually from CMDT, which neglects secondary crops, Figure 1 presents the total cultivated hectares for cotton, maize, sorghum, and millet over the past decade in the Koutiala and neighboring Yorosso Cercle. The significant decrease in cotton was due to a cotton strike in 2007/08 that protested low cotton prices and late payments to farmers. Overall, millet had the largest growth over the decade in absolute terms, adding over 30,000 hectares.

Figure 1: Hectares of Cotton and Coarse Grains in the Koutiala Zone, 1999/00 to 2010/11



Source: Author's manipulation of data from CMDT, Koutiala Regional Office

To analyze crop shares at the household level, I will draw on three hundred observations that come from two rounds of surveys in the Koutiala Cercle that gathered information on the same households for the 2008/09 and 2009/10 farming seasons. Data were collected by IER (*Institut d'Economie Rurale du Mali*), CIRAD (*Centre de cooperation internationale en recherche agronomique pour le développement*), and Michigan State University with funding from USAID and the Gates foundation.

Within the Koutiala Cercle, six villages were selected purposively to ensure representation of intra-cercle diversity in a region known for its economic, social, and ecological homogeneity. The criteria for village selection were primarily based on two factors: population and access to markets. The latter criterion simply distinguished whether market access was easy or difficult; three villages are described by each category. Table 1 below lists the six villages, their 1998 population (Samake 2008), ease of market access, location of and distance to the village's primary market (Kelly et. al. 2012), average cultivated hectares per household, and the average number of members present in the household. Below it, Figure 2 displays a map of the Koutiala Cercle and the sampled villages (Kelly et. al. 2012).

Table 1: Description of Six Villages in Sample from Koutiala Cercle

No.	Village	Pop. 1998	Market Access	Primary Market	Distance to Primary Market	Avg. Cultivated Hectares per Household	Avg. Size of Household
1	Nampala II	982	Difficult	Zangasso	30 km	10.6	14.9
2	Tonon	286	Difficult	Zangasso	22 km	7.6	14.1
3	Kaniko	1735	Easy	Koutiala	15 km	10.9	20.2
4	Try I	864	Easy	Koutiala	13 km	8.5	16.1
5	Signe	1005	Easy	Koutiala	15 km	9.9	18.3
6	Gantiesso	3219	Difficult	Mpessoba	22 km	10.7	18.5

Figure 2: Map of Koutiala Cercle and Villages in Sample

Source: Kelly et. al. 2012, map designed by Steve Longabaugh

Approximately 25 households were randomly selected and surveyed in the six villages and the dataset contains two observations for each household. Therefore, after data cleaning, the sample consists of 284 observations. Because outcomes are likely to be highly correlated over time within the same household, observations will be "clustered" by a household identification number to control for potential correlation in the econometric model.

Surveys were also conducted in the local language of Bambara. The data contain detailed information on household demographics and acquired assets, including farming equipment and livestock. Further, the dataset describes the household's agricultural production in full, including land allotment and some other measures not needed for this study, such as production levels and input costs.

One main limitation of the Koutiala data is susceptibility to human error. All estimates of hectares planted as well as total production are based on household memory. Thus, efforts were made by IER to survey groups of farmers within each household, instead of just the household head, who is sometimes removed from current farming activity due to old age. Additionally, all survey rounds were conducted between February and May of the production cycle, at least three to four months after harvest and nine months since planting. While any delay in survey questioning raises the probability of human error, surveys were implemented during these months because they are Mali's inactive, hot season. Since planting typically begins in May and the harvest usually ends in November, this was the optimal time to interview a large number of households.

IV. ECONOMETRIC MODEL

Fractional Multinomial Logit Model

The fractional multinomial logit was developed by Sivakumar and Bhat (2002), and has been described and applied by a few others (Ye and Pendyala 2005; Mullahy 2010; Koch 2010). The technique combines two variations on the standard logit model: the fractional logit and the multinomial logit. The consequence is a model where the explained variable y is able to represent the different shares of various types of y, all of which sum to one, much like the various categories in a pie chart. For this reason, the model is in the family of multivariate fractional logit models (e.g., Murteira and Ramalho 2012), because it is measuring the changes in shares of multiple variables simultaneously as a result of some explanatory variables. In other words, it allows one to ask how the slices of a pie chart change between observations as a result of differences in a certain set of related factors. In this case, the whole pie chart is a household's total number of cultivated hectares, meaning that the fractional multinomial logit model can help to see how changes in market and household characteristics affect the share of land devoted to cotton, maize, sorghum, millet, or secondary crops.

Combining some main elements of the fractional logit and the multinomial logit models to come up with the fractional multinomial logit model is fairly straightforward. The fractional logit model differs from the standard logit model as it treats the dependent variable as an expected value defined by an interval rather than a response probability (Papke and Wooldridge 1996). Similarly, the fractional multinomial logit model must ensure that the expected share of any outcome j lies between parameters A and B and that the sum of shares for all outcomes sums to unity. Mathematically,

$$A \leq E(S_j \mid x) \leq B, \quad j = 0,\ldots, J, \text{ where } A=0 \text{ and } B=1, \tag{1}$$

$$\sum_{j=0}^{J} E(S_j \mid x) = 1 \tag{2}$$

This technique permits the evaluation of shares of total farm land instead of the probability of whether or not a crop was cultivated.

The multinomial logit describes a technique for comparing the response probabilities for several categorical variables through use of a pivot outcome, which is the difference between one and the sum of expected shares for all other outcomes. Likewise, the fractional multinomial logit model defines a pivot outcome as well, but again, its dependent variables are fractional outcomes (i.e., crop shares), not response probabilities. Defining j = 0 as the pivot outcome, the fractional multinomial model also must establish expressions for every outcome within the logit framework.

$$E(S_j \mid x) = G(\beta_0 + \beta_k x_k) = G(z) = \exp(z)/[1 + \sum \exp(z)], \qquad j = 1, 2, \ldots, J. \tag{3}$$

$$E(S_0 \mid x) = G(\beta_0 + \beta_k x_k) = G(z) = 1 / [1 + \sum \exp(z)], \qquad j = 0 \tag{4}$$

Use of the pivot outcome equation (4) to estimate multiple outcomes makes it possible to evaluate the effect of explanatory variables on several crops simultaneously. Therefore, when joined together, the fractional multinomial logit model estimates coefficients which predict the expected share of several categorical outcomes within a defined interval, such as the share of cultivated land that a Malian household devotes to various crops.

By embedding the fractional logit function into the multinomial logit quasi-likelihood function, the econometric model can measure shares of outcomes—not probabilities—in what is a simplified form of the log likelihood function (Mullahy 2010). This new function, as a member of the linear exponential family, uses a quasi-maximum likelihood estimator (QMLE) and is efficient and consistently normally distributed provided the fractional logit function holds true (Ye and Pendyala 2005). The QMLE approach will maximize this new function and, with the assistance of a fractional multinomial logit Stata package (Buis 2008, updated 2012), run until it has converged and is able to predict crop shares.

However, because the multinomial logit estimator requires some normalization, these QMLE estimates will correspond to the coefficients in the multinomial shares model (Mullahy 2011). Thus, it produces coefficients that may be difficult to interpret. For this reason, using the coefficients predicted from an estimation of the fractional multinomial logit model, I calculate average marginal effects for every explanatory variable on each crop outcome, taking into account the coefficients for interaction terms when applicable.

Dependent Variables

The dependent variables are the crop shares for the portfolio of crops chosen by a household. The portfolio of crops for the Koutiala production zone consists primarily of cotton, maize, sorghum, and millet. Households may also cultivate peanuts, beans, or other crops, all of which will be grouped together under the category of secondary crops. The shares of a household's total cultivated hectares devoted to each of these crops are represented by *Cotton Share*, *Maize Share*, *Sorghum Share*, *Millet Share*, and *Secondary Share*, respectively. This makes five dependent variables—the percentage of total cultivated land devoted to cotton, maize, sorghum, millet, and secondary crops—the sum of which represents all cultivated land on a farm.

Table 2: Summary Statistics for Dependent Variables

Variable	Mean	Std. Dev.	Percent w/o Share	Minimum (after 0)	Maximum
Cotton Share	13.5%	12.4%	37.1%	5.6%	46.2%
Maize Share	10.5%	7.9%	16.0%	2.2%	60.0%
Sorghum Share	34.9%	17.7%	1.1%	3.1%	100.0%
Millet Share	24.0%	16.1%	11.4%	4.4%	92.3%
Secondary Share	17.1%	15.1%	18.9%	0.1%	69.6%

Table 2 summarizes some basic descriptive information about the dependent variables. The standard deviations show that crop shares are heterogeneous between households. The percent of households without any share (i.e., not cultivating the crop) reveals that 37% of observations did not cultivate any cotton. Therefore, more so than with the other crops, results that suggest a decrease in the expected share of cotton due to a change in the explanatory variable may explain either a marginal decrease in the share of land devoted to cotton or, perhaps for some households, dropping cotton altogether. Readers are asked to bear this mind when interpreting the results.

Independent Variables

The independent variables selected to predict crop shares were informed by the reduced-form land allocation function and the agricultural household model. Collectively, the independent variables describe "Labor and Land," "Capital," "Transaction Costs," "Households Characteristics," and "Time and Location," all of which are likely determinants of crop shares in an agricultural household. Under these five categories, Table 3 describes the specific variables chosen for the model, including household members per hectare, possession of farming and transportation equipment, literacy, and dummies for ethnicity, year, and village.

Under Labor and Land, the model starts by including multiple variables for each relevant age and gender categories—adult men, adult women, youth (ages 6-14), and infants (ages 0-5)—per number of the household's total cultivated hectares. Together, these four variables represent a household's labor supply and consumption demand relative to the household's available supply of land. I have inserted an interaction term between *WomenPerHct* and *InfantsPerHct* to control for an expected negative effect of children ages five and under on their mother's (or other caregiver's) labor productivity. I have also included the variables %*MenInactive* and %*WomenInactive*, and their respective interaction terms, to measure the percentage of household adults who were "inactive" during the farming season.

Capital is represented by possession of common farming equipment: pesticide sprayers used mostly for cotton production, draft-powered plows, and draft animals. An interaction term is included between *Plows* and *Oxen*, because I expect to see a high correlation between the number of draft animals and draft-powered plows owned by the household. By including capital in the model, there is a risk of simultaneity bias. However, I assume that at the time of planting, land allocation decisions are not decided by the household's long-term goals, but by their current capital constraints (among other things), and that most households do not have excess cash or credit during the planting season to purchase more equipment.

Table 3: Summary of Independent Variables

	Variable	Description	Mean	Std Dev
LABOR AND LAND: Household Members per Hectare and Inactivity				
1	MenPerHct	# of adult males (age ≥15) per hectare	0.42	0.23
2	WomenPerHct	# of adult females (age ≥15) per hectare	0.50	0.30
3	YouthPerHct	# of children (6 ≤ age ≤ 14) per hectare	0.57	0.59
4	InfantsPerHct	# of children (age ≤ 5) per hectare	0.33	0.32
5	WomenPerHct* InfantsPerHct	Interaction term between (2) and (4)	0.22	0.57
6	%MenInactive	% of adult males who are inactive	0.05	0.13
7	MenPerHct* %MenInactive	Interaction term between (1) and (10)	0.02	0.06
8	%WomenInactive	% of adult females who are inactive	0.12	0.17
9	WomenPerHct* %WomenInactive	Interaction term between (2) and (12)	0.06	0.09
CAPITAL: Farming Equipment				
10	Sprayers	# of sprayers owned	1.17	0.87
11	Plows	# of plows owned	1.75	1.86
12	Oxen	# of draft oxen owned	2.95	2.72
13	Plows*Oxen	Interaction term between (15) and (16)	7.35	17.87
TRANSACTION COSTS: Transportation				
14	Motorcycles	# of motorcycles owned	0.63	0.79
15	Bicycles	# of bikes owned	2.11	1.49
16	Carts	# of draft animal carts owned	0.99	0.54
HOUSEHOLD CHARACTERISTICS: Literacy and Ethnic Identity				
22	%MenLiterate	% of adult males who are literate	0.64	0.40
23	MenPerHct* %MenLiterate	Interaction term between (1) and (6)	0.27	0.25
24	%WomenLiterate	% of adult females who are literate	0.50	0.45
25	WomenPerHct* %WomenLiterate	Interaction term between (2) and (8)	0.26	0.32
26	Bambara	Dummy: 1 if Bambara ethnicity, 0 if not	0.13	0.34
27	Senoufo	Dummy: 1 if Senoufo ethnicity, 0 if not	0.09	0.28
28	Peulh	Dummy: 1 if Peulh ethnicity, 0 if not	0.06	0.23
29	OtherEthnic	Dummy: 1 if other ethnicity, 0 if not	0.03	0.18
TIME AND LOCATION: Year and Village				
30	Year_2010	Dummy: 1 if from year 2009/10, 0 if not	0.43	0.50
31	Village_2	Dummy: 1 if from Tonon, 0 if not	0.14	0.34
32	Village_3_{M*}	Dummy: 1 if from Kaniko, 0 if not	0.14	0.35
33	Village_4_{M}	Dummy: 1 if from Try I, 0 if not	0.14	0.34
34	Village_5_{M}	Dummy: 1 if from Signe, 0 if not	0.14	0.35
35	Village_6	Dummy: 1 if from Gantiesso, 0 if not	0.15	0.36

**Subscript M indicates those villages where market access is "easy"*

Transaction Costs are represented by possession of three modes of private transportation: motorcycles, bicycles, and carts, all of which can reduce transactions costs to a distant market. Carts can be a production factor too, as they also help households transport cotton and coarse grains to storage after the harvest. Another important measure for transaction costs—distance to the regional Koutiala market—are incorporated through the village dummy variables, discussed below.

Household characteristics are represented firstly by literacy in the model. For both men and women, literacy can reduce transaction costs at market and also serves as a proxy for other factors that are difficult to measure, such as intelligence and personal motivation, which may affect crop share decisions in the long-term. While many development studies only consider the educational level of the household head, I include literacy for all household adults after observing that while household decisions are authorized or approved by the household head, most adult males participate in the decision-making process for maintenance of household lands and many women maintain small fields or gardens.

Household characteristics are also represented by ethnicity. Since the vast majority of households in the Koutiala production zone identify themselves as Minianka, this ethnic group is the omitted category to which other ethnic groups are compared. The variable *OtherEthnic* is comprised of five ethnic groups, the Soninke and Malinke people from the west and the Bozo, Bobo, and Dogon from north-eastern Mali. Together, they total only twelve households, which is why they are grouped together into one variable, despite the fact that they represent very diverse cultures. The purpose of dummy variables for ethnic identity is that ethnicity may hint at each household's tastes and preferences. If the theory behind the agricultural household model is correct, then in the presence of transaction costs or other barriers to market transactions, a household's tastes and preferences may factor directly into its land allocation choices.

Finally, time and location dummy variables are included representing the year and village to which the observation is associated. First, both the year and village are intended to control for differences in environment, such as rainfall and soil quality. While environmental characteristics can vary within villages in a given year, these variables are the best available proxies. Secondly, these variables help control for differences in expected prices, which assumes that households from the same village and in the same year have very similar price expectations for coarse grains (a pan-territorial producer price for cotton is fixed annually). Finally, location defines a household's market access (shown in Table 1), which is correlated with transaction costs. The three villages defined as "easy" are approximately twenty kilometers or less from the regional capital Koutiala, which has a bustling market, while the three villages labeled "difficult" are further away. While weekly markets exist in other towns and villages, access to trade in Koutiala likely reduces transactions costs for market vendors.

Mostly due to a lack of sufficient data and possible simultaneity biases, some of the determinants shown in the reduced-form land allocation function are not represented in the model. These are profits generated by off-farm employment (π_y), variable inputs (V), and a household's risk preference (σ). Profits generated by off-farm employment, such as livestock or migratory work in a city, were not included because the data were unavailable and may have led to a simultaneity bias. Variable inputs were also excluded due to issues of simultaneity, since fertilizer use is required for cotton production and fertilizer procurement often occurs after planting. Moreover, a household's risk preference is not included as it was not directly measured by the survey.

V. RESULTS & DISCUSSION

Results as Average Marginal Effects

Drawing from 284 observations, the fractional multinomial model converged on a log pseudo-likelihood of -404.70 with a Wald chi-squared of 2809.19. To control for potential correlation over time within the same household, observations were "clustered" by a household identification number to ensure that standard errors were estimated robustly.

Table 4 presents the average marginal effects of the independent variables on crop shares. Average marginal effects that are statistically different from zero at the 10%, 5%, and 1% levels are indicated with one, two, or three asterisks, respectively; coefficients that are not statistically different from zero at the 10% level or below receive no asterisk. Of the model's 120 coefficients for average marginal effects, 24 are significant at the 10% level.

A few other points must be made about the interpretation of the coefficients in Table 4. For continuous variables, the coefficients represent the mean of the change in crop shares as a result of a marginal change in the explanatory variables for all observations. So for example, the first coefficient for *MenPerHct* under the outcome *Cotton Share* is -0.0114, which suggests that a one-unit increase in *MenPerHct*, all else equal, is associated with an average decrease of 1.14% for land allocated to cotton across all households, though in statistical terms, this is no different than zero. For binary variables, the coefficients represent the average change in crop shares resulting from a shift in the variables' minimum to its maximum, across all households. Thus, the coefficient for *Bambara* under *Cotton Share* is 0.0148, which suggests that—relative to a Minianka household—a Bambara household has an average of 1.48% more land devoted to cotton, though in statistical terms, this is no different than zero. The upcoming discussion will highlight coefficients deemed to have economic and statistical relevance in explaining difference in crop shares across all households in the sample.

Furthermore, because crop shares must always sum to one—as they are defined by a finite amount of total cultivated hectares—the sum of the average marginal effects for any one independent variable is zero; in other words, what an independent variable takes away from one crop's share, it gives to others' shares. Additionally, the average marginal effects of interaction terms are not presented as such a measure does not exist; under the assumption of "all else equal," an interaction term has no marginal effect as it is only the product of two other explanatory variables. However, the coefficients estimated for the interaction terms by the fractional multinomial logit model were incorporated into the calculation of the average marginal effects for those explanatory variables involved in the interaction term. In short, while the interaction terms have no measure of average marginal effect for themselves, the consequences of the interactions in the fractional multinomial logit are present in the average marginal effects in Table 4.

Table 4: Average Marginal Effects of the Independent Variables (Derived from Results of Fractional Multinomial Logit)

Obs: 284 | **Wald Chi^2: 2809.2**

Log pseudolikelihood = -404.70322 | **Prob > Chi^2: 0.0000**

	Cotton Share			Maize Share			Sorghum Share			Millet Share			Secondary Share		
	Coef.	Rbst S.E.	Sig.	Coef.	Rbst S.E.	Sig.	Coef.	Rbst S.E.	Sig.	Coef.	Rbst S.E.	Sig.	Coef.	Rbst S.E.	Sig.
MenPerHct	-0.0114	0.0623		0.0448	0.0290		-0.0690	0.0603		0.0022	0.0639		0.0334	0.0746	
WomenPerHct	-0.0573	0.0453		-0.0007	0.0362		0.0466	0.0736		0.1303	0.0554	**	-0.1188	0.0584	**
YouthPerHct	0.0252	0.0249		0.0101	0.0206		0.0564	0.0400		-0.1214	0.0432	***	0.0297	0.0503	
InfantsPerHct	0.0790	0.0379	**	0.0100	0.0348		0.0039	0.0632		-0.0488	0.0561		-0.0441	0.0752	
%MenInactive	-0.0845	0.0675		0.0491	0.0379		0.1360	0.1184		-0.1024	0.0881		0.0017	0.0614	
%WomenInactive	-0.0047	0.0544		0.0091	0.0481		0.1236	0.0700	*	-0.0978	0.0695		-0.0303	0.0542	
Sprayers	0.0185	0.0098	*	-0.0071	0.0076		-0.0154	0.0153		-0.0307	0.0156	**	0.0347	0.0149	**
Plows	0.0000	0.0097		0.0125	0.0089		0.0052	0.0138		0.0059	0.0133		-0.0236	0.0137	*
Oxen	0.0098	0.0052	*	0.0040	0.0034		0.0038	0.0084		-0.0056	0.0078		-0.0121	0.0068	*
Motorcycles	0.0008	0.0103		0.0084	0.0088		-0.0218	0.0157		0.0178	0.0163		-0.0052	0.0156	
Bicycles	0.0018	0.0071		-0.0062	0.0045		0.0081	0.0096		0.0020	0.0103		-0.0057	0.0086	
Carts	0.0374	0.0229		-0.0011	0.0154		-0.0289	0.0289		-0.0156	0.0292		0.0082	0.0272	
%MenLiterate	0.0227	0.0362		-0.0046	0.0228		-0.0405	0.0498		-0.0366	0.0377		0.0590	0.0428	
%WomenLiterate	0.0477	0.0502		0.0432	0.0302		-0.0225	0.0781		-0.0010	0.0835		-0.0673	0.0542	
Bambara	0.0148	0.0291		-0.0471	0.0152	***	0.0372	0.0441		-0.0101	0.0411		0.0051	0.0506	
Senoufo	-0.0173	0.0365		0.0142	0.0264		0.0460	0.0530		-0.0236	0.0527		-0.0192	0.0512	
Peulh	-0.0069	0.0313		-0.0294	0.0204		0.0622	0.0534		0.0600	0.0859		-0.0859	0.0533	
OtherEthnic	-0.0408	0.0368		-0.0547	0.0179	***	0.0015	0.0723		0.1295	0.0967		-0.0355	0.0414	
Year_2010	0.0206	0.0425		-0.0316	0.0216		0.0121	0.0628		-0.0032	0.0725		0.0020	0.0541	
Village_2	-0.0228	0.0359		0.0925	0.0312	***	0.1438	0.0489	***	-0.1645	0.0356	***	-0.0490	0.0514	
Village_3_M	-0.0263	0.0445		0.0359	0.0248		-0.1289	0.0354	***	0.0323	0.0569		0.0870	0.0523	*
Village_4_M	-0.1164	0.0159	***	-0.0245	0.0200		-0.0068	0.0425		-0.0026	0.0481		0.1503	0.0580	***
Village_5_M	-0.0606	0.0286	**	0.0330	0.0263		-0.0528	0.0461		0.0507	0.0469		0.0296	0.0618	
Village_6	0.0040	0.0414		0.0201	0.0203		-0.0810	0.0422	*	-0.0927	0.0502	*	0.1496	0.0530	***

Legend: *** = P < .01 ** = P < .05 * = P < .10

Discussion by Crop Type

For cotton, the model estimated that the key determinants of increased land allocation were the number of infants per cultivated hectare and the number of pesticide sprayers and draft animals owned by the household. While infants cannot assist with cotton production nor consume it directly, this link may be because families with young children expect many out-of-pocket expenses—from medicine to future school fees—that require cash, which can be earned through cotton production. Also, since cotton growing has high start-up costs, the ownership of capital needed for successful cotton production, such as oxen and sprayers, helps farmers decide to grow more of it. Village location was also influential, suggesting that those households living in a village closer to Koutiala planted smaller shares of cotton. This may be because better access to Koutiala's markets reduced farmer's transaction costs for trading and earning income with other crops and purchasing inputs for coarse grains (which are delivered by CMDT if purchased on cotton credit).

For maize, the key determinants of land allocation were ethnic identity and village location. Households belonging to the Minianka and Senoufo ethnic devoted significantly more land to maize relative to Bambara, Peulh, or Other households, perhaps reflecting a cultural or historical preference for maize. If the difference in maize shares between ethnic groups is due to differences in household tastes or preferences, this illustrates the influence of household consumption needs into land allocation decisions—a main principle of the agricultural household model. However, maize is the only crop in the model where there is a statistical difference between two ethnic groups and the majority Minianka ethnic group. Also, one village allocated a significantly higher share of their land to maize, though interestingly, this village was designated as having difficult market access.

While sorghum and millet have similar subsectors, the results imply that the determinants for how much land is devoted to each differ. Increased sorghum share is associated with higher percentages of inactive women in the household. Further, a significantly higher share of sorghum is grown in a village with difficult market access, and a significantly smaller share of sorghum is grown in a village with easy market access. While I would have expected millet to fulfill a similar role, the results show that smaller shares of millet are associated with villages with difficult market access and ownership of pesticide sprayer equipment, which suggests that millet may be a marketable and hardy coarse grain. Moreover, every additional woman per hectare is correlated with a 13 percent larger share of millet—likely because women help with the critical weeding periods—while an additional youth per hectare is correlated with a 12 percent smaller share of millet. So while their subsectors may be similar, the determinants of land allocation to sorghum and millet have distinct differences.

For secondary crops, the results reveal several statistically significant determinants. First, village location seems important. This may be due to increased access to markets—two villages near Koutiala had significantly higher shares of secondary crops—or due to geographic differences: the village of Gantiesso, which also had high shares of secondary crops, may have been located near a river that made rice production feasible. In terms of equipment, ownership of pesticide sprayers was associated with higher shares of secondary crops, while ownership of plows and oxen were associated with slightly smaller shares of secondary crops. Finally, smaller shares of secondary crops were linked to increased number of women per hectare. Since it was thought that women are mostly responsible for secondary crop production, the negative effect of adult women per hectare on secondary crop shares was unexpected.

Interestingly, there were many variables that did not have any statistically significant effect on any crop share, including men per hectare, percentage of men who were inactive, or transportation variables (i.e., ownership of motorcycles, bicycles, and carts). Literacy also did not have a statistically significant effect on any crop share, failing to demonstrate the effect of literacy or what it represented on land allocation. Furthermore, the Senoufo and Peulh ethnic groups are not statistically different from the Minianki ethnic group, and the dummy variable for year has no significant effect either.

Discussion by Household Type

It is a useful exercise to apply the average marginal effects from Table 4 to a couple of hypothetical households, now that the results have been discussed by crop type. This will demonstrate the usefulness of the average marginal effects when trying to predict total crop shares for specific cases. In particular, I want to examine differences between a wealthier, larger household and a small, disadvantaged household. For simplicity, both cases will be from the same year, village, and ethnic group, dismissing the need to consider these coefficients in the calculation. The predicted crop shares for both households, and the differences between them, highlight many of the results discussed above.

The first household has six adult males, eight adult females, eight young boys and girls over the age five, and four children under five. The household farms on 16 hectares and owns two sprayers, three plows, six draft animals, two carts, two motorcycles and three bikes. Furthermore, three adult men are literate along with two of the adult women, and only the elderly grandmother is considered inactive. These figures are slightly better than the average 2008/09 Minianka household of this size. The second household is comprised of two adult males, two adult females, three boys and girls over five, and three children under five. The household farms on six hectares with one plow, one ox, one bicycle, but does not own a sprayer, motorcycle, or a cart. One adult male is literate and none are inactive. Again, these numbers are slightly lower or more disadvantaged than the average 2008/09 Minianka household of this size.

Assuming that both households belong to the same year, village, and ethnic group, the disadvantaged household is estimated to allocate 14% less land to cotton and 4% less land to maize, relative the larger household. In exchange, the disadvantaged household will allocate 7% more land to sorghum, 4% more land to millet, and 7% more land to secondary crops, relative the larger household. By definition, differences in crop shares between the two households sum to zero. These differences between the aggregate average marginal effects for two households emphasize the role of cotton and, to some extent, maize as a cash crop or preferred crop for households with the proper farming equipment and labor supply.

For the sake of another example, consider if the disadvantaged household now has one inactive male—an unfortunate but possible scenario. This one change impacts the differences in crop shares between the two households. Now, the disadvantaged household is estimated to allocate

19% less to cotton share and 14% more to sorghum share, relative to the larger household. The strengthening of the divide between cotton and sorghum shares highlights sorghum's role as a crop for vulnerable households with less land, labor, or other inputs.

The primary purpose of this exercise was to demonstrate how the average marginal effects on crop shares add up, though some clearly have more effect than others. It presented two scenarios of realistic, yet different, households in the Koutiala production zone and predicted how their crop shares may differ relative to each other. In the second example, the model was able to predict land allocation differences for almost a quarter of the household's total cultivated land, even though these families could have been neighbors. Using the coefficients, it is also possible to predict the expected crop shares for different years and ethnic groups represented in the data.

VI. CONCLUSION

Understanding land allocation of field crops in Mali is important for improving household food security and preparing for challenges facing Mali's cotton industry. Policymakers need to know how certain market and household characteristics affect planting decisions of cotton, coarse grains, and secondary crops. Whereas many studies examine this issue one crop at a time—perhaps building on a supply response model or estimating the probability of crop adoption with a logit or probit model—I employed a fractional multinomial logit to data from two survey rounds of 153 households in Mali's Koutiala Cercle. The fractional multinomial logit technique builds on a standard logit by allowing for categorical, non-binary, dependent variables whose values are fractions which sum to one. For my purposes, the share of land allocated to cotton, maize, sorghum, millet and other "secondary" crops served as the five dependent variables, all of which, when combined, equaled the total number of hectares cultivated by a household. The fractional multinomial logit results were estimated through quasi-maximum likelihood and transformed into the average marginal effects of the explanatory variables on each crop share category.

The study does suffer from limitations, including human error from survey respondents or enumerators and a low volume of observations that is unfortunately common for many studies of agricultural households in the developing world. There are also three primary limitations with the fractional multinomial logit model. First, the model's focus on land allocation disregards the effect of fertilizer and careful maintenance on production. Second, the use of the fractional multinomial logit prevents the inclusion of crop-specific variables. Third, the use of crop shares as dependent variables makes land allocation appear as a zero-sum game—that is, a situation in which all gains are some other's losses—though in reality, households could respond to new incentives by planting additional hectares of crops or increasing adoption of fertilizer to boost yields. However, these limitations were discussed and kept in mind throughout the interpretation.

The results found that the most influential sets of variables in determining land allocation of cotton, maize, sorghum, millet, and secondary field crops were those representing village location, the latter of which may be due to proximity to regional markets. Villages closer to Koutiala were closely associated with much higher shares of millet and lower shares of maize and cotton. Also, ownership of equipment was correlated with crop shares; particularly, ownership of each additional pesticide sprayer was associated with increase in cotton and secondary crop share and a decrease in millet share. Meanwhile, variables sets representing family and farm size, literacy, ethnicity, and modes of transportation had significant results for some variables on particular outcomes, but were not as revealing overall.

These results provide a few insights for agricultural policymakers hoping to improve food security in Mali's Koutiala Cercle. First, the results suggest that cotton shares are highest when Mali's cotton company CMDT is able to help farmers procure equipment and fertilizer on cheap cotton-backed credit and reduce farmers' transaction costs by transporting inputs and cotton output, particularly in economically isolated villages. Therefore, for those looking to reform the cotton industry, it is important to consider if these extra incentives can be sustainable or persevered through privatization. Second, while maize production has grown significant throughout Mali in the last decade (Staatz et. al. 2011), it remains cultivated on the least amount of hectares in the Koutiala Cercle, especially by ethnic groups not native to the area. This may because of varying taste and cooking preferences between ethnic groups or because CMDT extension, which introduced and promoted maize years ago, has dwindled and may not encourage maize production as much. Third, the results suggest that sorghum is the "safe" crop, grown more by vulnerable households, and millet is a marketable coarse grain. These differences in the roles that sorghum and millet play should be studied more to determine if these differences exist only in the dataset or truly represent trends across the production zone or country, especially because much of the policy discussion now is about the marketability of maize as opposed to millet. Finally, prevalence of secondary crops varies much by geographic location, potentially due to market access or environmental conditions (e.g., a river to grow rice and other water-intensive crops). Overall, the results suggest that promoting coarse grains and secondary crops and further developing their markets will benefit food production in the Koutiala Cercle.

The primary motive of this article was to develop a new method of modeling household land allocation to three or more crops in developing countries. To do this, I adapted the agricultural household model for use in Mali's Koutiala production zone and applied the relatively new fractional multinomial logit framework. Even with its limitations, the overall model had definite explanatory power and was useful in discussing determinants of crop choice and planting at the household level. What is needed now is additional work applying this model to different circumstances. To start, additional data including alternative sources of household income, last year's food stock at time of planting, seasonal labor supply estimates, and village-level variables (e.g., distance to paved road, nearest weekly market, or access to a microfinance institution), could contribute to the model. Within Mali, it can be applied to other regions and their alternative crop portfolios to better understand food security, especially after the country's political instability following the coup d'état in late March of 2012. Another idea is to open up the dependent variable representing shares of secondary crops to see how the explanatory factors affect peanuts, sweet potatoes, sesame, and vegetables differently. In fact, use of the model on a generous dataset in any developing country can help to provide evidence for theoretical discussions surrounding the agricultural household model, such as the extent of household preferences, literacy, or transaction costs on land allocation decisions. As it can compare all crops simultaneously, the fractional multinomial logit can serve as an additional tool to study the farming system on the household farm, which continues to be the most fundamental economic unit in the majority of the developing world.

REFERENCES

Ahn, C. Y., Singh, I., and Squire, L. (1981). A Model of an Agricultural Household in a Multi-Crop Economy: The Case of Korea. *The Review of Economics and Statistics*, *63(4)*, 520-525.

Askari, H., and Cummings, J. T. (1977). Estimating Agricultural Supply Response with the Nerlove Model: A Survey. *International Economic Review, 18(2)*, 257-292.

Barrett, C. B. (2008). Smallholder market participation: Concepts and evidence from eastern and southern Africa. *Food Policy, 33(4)*, 299-317.

Benjamin, D. (1992). Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models. *Econometrica, 60(2)*, 287-322.

Boer, A. J. D., and Chandra, S. (1978). A Model of Crop Selection in Semi-Subsistence Agriculture and an Application to Mixed Agriculture in Fiji. *American Journal of Agricultural Economics, 60(3)*, 436-444.

Boughton, D., and Dembele, N. N. (2010). *Rapid Reconnaissance of Coarse Grain Production and Marketing in the CMDT Zone of Southern Mali: field work report of the IER-CSA-PROMISAM team December 13 – 19, 2009* (Working paper 2010-1) East Lansing, MI: Food Security III Cooperative Agreement, Michigan State University.

Buis, M. L. (2008). FRACTIONAL MULTINOMIAL LOGIT: Stata module fitting a fractional multinomial logit model by quasi maximum likelihood [computer software]. Boston College Department of Economics. Retrieved from http://ideas.repec.org/c/boc/bocode/s456976.html

Buis, M. L. (2012). *fractional multinomial logit: module fitting a fractional multinomial logit model by quasi-maximum likelihood.* Retrieved September 5, 2012, from http://maartenbuis.nl/software/fractional multinomial logit.html

Dioné, J. (1989). Informing food security policy in Mali: interactions between technology, institutions and market reforms. (Doctoral dissertation, Michigan State University, 1989).

Dorward, A. (2011). *Conceptualising seasonal financial market failures and credit rationing in applied rural household models*. Retrieved from University of London, School of Oriental and African Studies, Centre for Development, Environment and Policy at: http://eprints.soas.ac.uk/10806/

Fafchamps, M. (1992). Cash Crop Production, Food Price Volatility, and Rural Market Integration in the Third World. *American Journal of Agricultural Economics, 74(1)*, 90-99.

Goetz, S. J. (1992). A Selectivity Model of Household Food Marketing Behavior in Sub-Saharan Africa. *American Journal of Agricultural Economics, 74(2)*, 444-452.

Hazell, P. B. R. (1982). Application of Risk Preference Estimates in Firm-Household and Agricultural Sector Models. *American Journal of Agricultural Economics, 64(2)*, 384-390.

Janvry, A. d., Fafchamps, M., and Sadoulet, E. (1991). Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained. *The Economic Journal, 101(409)*, 1400-1417.

Kelly, V., Murekezi, A., Me-nsope, N., Perakis, S., and Mather, D. (2012). *Cereal Market Dynamics: The Malian Experience from the 1990s to Present* (International Development Working Paper 2012-128). East Lansing, MI: Agricultural, Food and Resource Economics. Michigan State University

Koch, S. F. (2010). *Fractional Multinomial Response Models with an Application to Expenditure Shares* (Working paper 2010-21) Pretoria, South Africa: Economics, University of Pretoria.

Mullahy, J. (2010). Multivariate Fractional Regression Estimation of Econometric Share Models (NBER Working Papers 16354). Cambridge, MA: National Bureau of Economic Research. Retrieved September 9, 2012, from http://www.nber.org/papers/w16354.

Murteira, J. M. R., and Ramalho, J. J. S. (2012). Regression Analysis of Multivariate Fractional Data (Working Paper). Retrieved on September 9, 2012, from http://www.eea-esem.com/files/papers/eea-esem/2012/2277/multivfracdata.pdf.

Omamo, S. W. (1998). Farm-to-Market Transaction Costs and Specialisation in Small-Scale Agriculture: Explorations with a Non-separable Household Model. *Journal of Development Studies, 35(2)*, 152-163.

Papke, L. E., and Wooldridge, J. M. (1996). Econometric Methods for Fractional Response Variables With an Application to 401 (K) Plan Participation Rates. *Journal of Applied Econometrics, 11(6)*, 619-632.

Samake, A., et al. (2008). *Changements structurels des économies rurales dans la mondialisation : Programme RuralStruc Mali - Phase II* (World Bank Report). Bamako, Mali: Michigan State University, IInstitut d'Economie Rurale du Mali, Centre de Coopération Internationale en Recherche Agronomique pour le Développement.

Singh, I., Squire, L., and Strauss, J. (1986). A Survey of Agricultural Household Models: Recent Findings and Policy Implications. *The World Bank Economic Review, 1(1)*, 149-179.

Sivakumar, A., and C.R. Bhat. (2002). "A Fractional Split Distribution Model for Statewide Commodity Flow Analysis." *Transportation Research Record*. 1790: 80-88.

Staatz, J., et al. (2011). Mali Agricultural Sector Assessment (Report prepared for United States Agency for International Development). East Lansing, MI: Food Security III Cooperative Agreement, Michigan State University.

USAID: United States Agency for International Development. (2011). *Country Overview: Mali*. Bamako, Mali.

Van Dusen, M. E., and Taylor, J. E. (2003). Missing Markets and Crop Diversity: Evidence from Mexico. *Environment and Development Economics, 10*, 513-531.

Ye, X., and Pendyala, R. M. (2005). *A Model of Daily Time Use Allocation Using Fractional Logit Methodology*. Paper presented at the 16th International Symposium on Transportation and Traffic Theory, College Park Maryland, United States.