A MIXTURE OF EXPERTS MODEL TO EXPLAIN HOUSEHOLDS' CHOICE PATTERNS FOR TERMITE CONTROL OPTIONS IN LOUISIANA

Krishna Paudel, Mahesh Pandit, Michael A Dunn
Louisiana State University and LSU Agricultural Center
Baton Rouge, Louisiana

Vinaya Amatya
Department of Computer Science, Louisiana State University

Contact Information

Krishna P. Paudel
Associate Professor
Department of Agricultural Economics and Agribusiness
Louisiana State University and LSU Agricultural Center
Baton Rouge, LA 70803
Phone: (225) 578-7363
Fax: (225) 578-2716
Email: kpaudel@agcenter.lsu.edu

Poster prepared for presentation at the Agricultural & Applied Economics Association's 2011 AAEA & NAREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July 24-26, 2011

A Mixture of Experts Model to Explain Households' CHOICE Patterns for Termite Control Options in Louisiana

Krishna Paudel, Mahesh Pandit, Mike Dunn, and Vinay Chandra Amatya
Louisiana State University and LSU Agricultural Center, Baton Rouge, Louisiana, USA

Introduction

- **Invasive species**
  - Environmental damage (biodiversity and habitat loss)
  - Economic damage - approximate loss of $120 billion to the U.S. economy (Pimentel et al., 2005)
- **Formosan Subterranean termites (FST) (*Coptotermes formosanus* Shiraki),**
  - Native to China, introduced in the U.S. by returning war ships
  - Introduced after the second world war in the U.S.
  - Single colony may consist of 1-10 million termites
  - Attack wooden structures from the ground as well as aerially
  - Damage in the U.S. in the range of one billion dollars a year, damage in Louisiana $500 million a year
- **FST control subsidies**
  - Researchers and government agencies are attempting to identify the preferred treatment option by homeowners
  - Large scale subsidy payment may instituted in the future to control FST in Louisiana
  - Small scale subsidy payment is in effect since 2000 in the French Quarter area of New Orleans
- **Preference ranking**
  - Economists generally ask respondents to rank alternative choices but do not identify the most preferred option
  - The ranked nature of the preference data is modeled using the Benter model for ranked data (Benter, 1994) and a mixture model of these distribution gives a model based approach to clustering homeowners into preference blocs (Gormely and Murphy, 2008)

Objectives

- Use a mixture of experts model to analyze complete ranked preference data
- Identify how different preference block of rankings are impacted by the demographic and cognitive risk/benefit variables.

Data

- Collected from Louisiana homeowners in 2002 using mail survey following the Dillman's (2000) tailored design method
- A total of 5,641 single family homeowners were contacted through the use of our mail survey:
  - 1,490 from Monroe, 1,305 from Alexandria, 1,395 from Baton Rouge, and1,451 from the New Orleans Metropolitan areas
  - Pre-survey and focus group discussions were conducted prior to mailing the survey
  - A survey response rate of 25% was obtained, although not all respondents ranked the treatment options
- Four FST treatment options were provided for each individual homeowner to rank from the most preferred choice to the least preferred choice.

Treatment Control Option 1: No control option: cost $0/square foot,
Treatment Control Option 2: Liquid treatment option: cost $0.13/square foot,
Treatment Control Option 3: Bait treatment option: $0.43/ square foot,
Treatment Control Option 4: Liquid + Bait treatment option: $0.56/square foot

- Individual homeowners ranked these options as their first, second, third and fourth most preferred option to control FST.
- Complete rankings were provided by 716 respondents which are used for analysis.

Methods

- We believe heterogeneous population exists in Louisiana homeowners, so we need to cluster heterogeneous population into homogenous subpopulation. A preference bloc is defined to be a group of homeowners who have similar opinion on treatment control options for FST. A mixture of experts model (MoE) [Jacobs et al. (1991), Jordan and Jacobs(1994)] which combine the idea of mixture models [McLachlan and Peel (2000) and generalized linear models (Gormley and Murphy, 2008) works well. We used the mixture model developed by Gormley and Murphy (2008) to examine the influence of different factors on this clustering and to examine the characteristics preference of the preference bloc.

ABSTRACT

A mixture of experts model in which the model parameters are functions of covariates are used in rank data of preferred four control option of Formosan Subterranean Termites (FST) control by Louisiana homeowners to examine the influence of different factors on this clustering and to examine the characteristics preference of the bloc. The existence of preference bloc in the homeowners is established and it determined that survey location, household pre-tax income, and knowledge of FST are important factors in influencing choice in this survey.

Table 1. Percentage of respondents ranking preference i in position j

	Rank			
Method	1	2	3	4
No Control	18.99	7.96	2.79	70.25
Liquid	54.05	25.98	18.16	1.82
Bait	12.29	56.28	30.45	0.98
Liquid +Bait	14.66	9.78	48.60	26.96

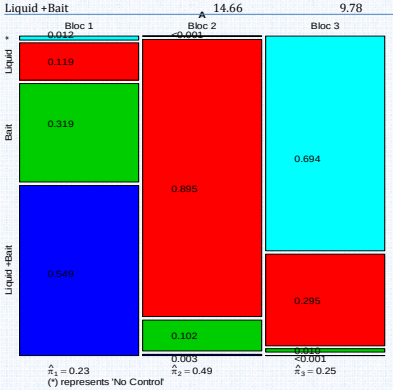

Figure 1. A graphical representation of the maximum likelihood estimates of the Benter support parameter for homeowner preference survey choice options

Figure 2. FST damage at the present

Figure 3. FST infestation likelihood

Table 2. Gating network parameter estimate ($\hat{\beta}$)

Covariates	Bloc 2		Bloc 3	
	Log odds	Odd ratio	Log odds	Odd ratio
Intercept	1.1261	3.0837	1.6506	5.2102
Survey location New Orleans (1=yes)	-0.5056	0.6031	-1.1249	0.3247
Home market value $300K or more (1=yes)	-0.5004	0.6063	-0.6800	0.5066
Concrete slab home foundation (1=yes)	0.0361	1.0368	-0.2631	0.7687
Termites an existing problem in neighborhood (1=yes)	-0.6206	0.5376	-1.3229	0.2664
Heard of FST (1=yes)	0.2759	1.3178	-0.3391	0.7124
Gender female (1=yes)	-0.2013	0.8176	-0.3236	0.7235

Major Reference

Gormley, I.C., and Murphy, T.B. (2008). A Mixture of Experts Model for Rank Data with Applications in Election Studies." *The Annals of Applied Statistics*2:1452-1477.
Other references are available upon request

Methods cont.

- The conditional probability of control options i ($i = 1, \ldots, 4$) rank y_i given their associated covariates w_i is $P(y_i|w_i) = \sum_{k=1}^{K} \pi_{ik} P(y_i|\theta_k)$, where K denotes the number of components in the mixture model.
- The gating network coefficients in MoE model $\pi_{ik} = \pi_k(w_i)$ is the probability of control option being a complete member of expert network k and θ_k represents the parameters of the probability model of the k^{th} expert network and are function of covariates. Specifically the gating network coefficients are assumed to be multinomial logistic functions of homeowners.
- Each network is characterized by a different parameterized Benter model which consists two types of parameter:
  - I. Support Parameter: Within expert network k, the support parameter vector is denoted by $\beta_k = (\beta_{k1}, \ldots, \beta_{kR})$, Where $0 \le \beta_{kr} \le 1, \sum_{r=1}^{R} \beta_{kr} = 1$ and n denotes the number of control options available for choice. The support parameter β_{kr} is interpreted as the probability of control option j being given a first preference by a complete member of choice option k.
  - II. Dampening parameter: The dampening parameter vector is denoted by $\alpha = (\alpha_1, \ldots, \alpha_R)$, where $\alpha_r \in [0,1]$ for $r = 1, \ldots, R$. To avoid over parameterization of the model, the constraints $\alpha_1 = 1$ and $\alpha_R = 0$ are imposed.
- The likelihood function to estimate MoE model for N rank observation is

$$L_c(\beta, \theta, \alpha | y, z, w) = P(y, z | w, \beta, \theta, \alpha) = \prod_{i=1}^{N} \prod_{k=1}^{K} [\pi_k(w_i) P(y_i | \theta_k, \alpha)]^{z_{ik}}$$

  When $i = 1, \ldots, N$, $z_i = (z_{i1}, \ldots, z_{iK})$ are latent variable which takes value 1 if homeowner i chose expert network k and, 0 otherwise. And the model parameters are estimated via the 'EMM' algorithm a hybrid algorithm of EM and MM [Gormley and Murphy, 2008]

Results

- Using Bayesian Information Criterion [BIC] [Schwartz, 1978], we found that the optimal number of expert networks is 3.
- Benter support parameter estimates
  - Figure 1 is a mosaic plot illustrating the Benter support parameters estimates within each of the three preference blocs in the optimal model.
  - Preference bloc 1 appears to favor of the expensive treatment control options as 72% of homeowners support for Liquid and Bait control options in this preference bloc with the largest support, and 32% support for Bait option.
  - Liquid has the largest support in second preference bloc with almost 90% support. Preference bloc 3 represents homeowners who do not want control termites as 69% supports for no-control treatment options.
- Benter dampening parameter estimates
  - Under the optimal model the Benter dampening parameter estimates are $\hat{\alpha} = (1, 1, 1, 1)$. The estimates suggest that the certainty with which homeowners rank their preference remains constant with respect to choice level.
- Gating network parameter estimates:
  - The gating network parameters associated with the "high expensive" preference bloc (i.e. preference bloc 1 are used as the reference parameters.
  - The odd ratio less than 1 in preference bloc 1 indicates that who know the termites an existing problem in neighborhood, has annual pretax household income $125 or more, have market value $300 or more west and stay in New Orleans chose expensive treatment option.
  - As our expectation, within voting bloc3, odd ratio for all covariates are negative which implies that the odds of all covariates for being best described by preference bloc3 are less than the odds for being described by the expensive voting bloc. These results suggest that the homeowners who belong to bloc3 do not prefer to spend money to control termite.

Conclusions

- If subsidy would continue to be provided by the government, it should be supported if households apply liquid treatment options. Heavily infested area such as New Orleans preferred the most expensive control option. It could be due to (i) Subsidy effect , (ii) Information effect
- Effective subsidy program may be tiered system: heavily infested area getting expensive control options and others getting lesser expensive control – Rather than supporting in cash or in-kind subsidy, education could be alternative option.
- Information should be targeted to different groups in Louisiana according to where they live, their prior experience with termites, and other demographic categories that relate to termite control option preferences and risk tolerances.