OMICS Group

OMICS Group International through its Open Access Initiative is committed to make genuine and reliable contributions to the scientific community. OMICS Group hosts over **400** leading-edge peer reviewed Open Access Journals and organizes over **300** International Conferences annually all over the world. OMICS Publishing Group journals have over **3 million** readers and the fame and success of the same can be attributed to the strong editorial board which contains over **30000** eminent personalities that ensure a rapid, quality and quick review process. OMICS Group signed an agreement with more than **1000** International Societies to make healthcare information Open Access.

OMICS Journals are welcoming Submissions

OMICS Group welcomes submissions that are original and technically so as to serve both the developing world and developed countries in the best possible way.

OMICS Journals are poised in excellence by publishing high quality research. OMICS Group follows an Editorial Manager® System peer review process and boasts of a strong and active editorial board.

Editors and reviewers are experts in their field and provide anonymous, unbiased and detailed reviews of all submissions.

The journal gives the options of multiple language translations for all the articles and all archived articles are available in HTML, XML, PDF and audio formats. Also, all the published articles are archived in repositories and indexing services like DOAJ, CAS, Google Scholar, Scientific Commons, Index Copernicus, EBSCO, HINARI and GALE.

For more details please visit our website: http://omicsonline.org/Submitmanuscript.php

Generalized Additive Models in Business and Economics

Sunil K Sapra
Department of Economics and Statistics
California State University
Los Angeles, CA 90032

SUMMARY

- The paper presents applications of a class of semi-parametric models called generalized additive models (GAMs) to several business and economic datasets. Applications include analysis of wage-education relationship, brand choice, and number of trips to a doctor’s office. The dependent variable may be continuous, categorical or count. These semi-parametric models are flexible and robust extensions of Logit, Poisson, Negative Binomial and other generalized linear models. The GAMs are represented using penalized regression splines and are estimated by penalized regression methods.

SUMMARY (Continued)

- The degree of smoothness for the unknown functions in the linear predictor part of the GAM is estimated using cross validation. The GAMs allow us to build a regression surface as a sum of lower-dimensional nonparametric terms circumventing the curse of dimensionality: the slow convergence of an estimator to the true value in high dimensions. For each application studied in the paper, several GAMs are compared and the best model is selected using AIC, UBRE score, deviances, and R-sq (adjusted). The econometric techniques utilized in the paper are widely applicable to the analysis of count, binary response and duration types of data encountered in business and economics.

Introduction and Overview

- Normal, Logit, Probit, and Poisson Regression Models are used widely in business and economics.
- These models belong to the family of Generalized Linear Models (GLMs).
- A GLM has a link function $\eta = g(\mu)$, which maps the mean into the set of real numbers.

Limitations of GLMs and Motivation for GAMs

- The GLMs are too restrictive in that these models have a linear predictor $\eta=\beta' x$. As a result, nonlinearities in the link function can be missed completely leading to unreliable inference.
- Generalized Additive Models (GAMs) are semi-parametric models, which provide an attractive alternative to GLMs in situations in which predictors may appear nonlinearly in the link function.

Generalized Additive Models (GAMs)

- GAMs replace the linear predictor $\eta = \beta' x$ with $\eta = g(\mu) = \alpha + \sum s_j(x_j)$, where $s_j(x_j)$ are non-parametric smooth functions.
- Generalized additive models are very flexible, and can provide an excellent fit in the presence of nonlinear relationships. The GAM approach gives us more flexibility in model form.
- Methods for estimating generalized additive Poisson models are discussed in Hastie and Tibshirani (1986, 1990) and Wood (2004) among others.

Additive Models

Consider the model

- $y_i = \alpha + \sum f_j(x_{ji}) + \varepsilon_{i,}, \ \varepsilon_i \sim N(0,\sigma^2)$

- f_j is an unknown smooth function of covariate x_j, which may be vector-valued.

- The f_j are confounded via the intercept, so that the model is only estimable under identifiability constraints on the fj .

- The most common constraints are

$$\sum f_j(x_{ji}) = 0$$

Backfitting Algorithm for Additive Models

$Step\,1$: Initialize $\hat{\alpha} = 1/N\sum_{i=1}^{N} y_i,\ f_j^0 = 0,\ j = 1,2,\ldots,p.$

$Step\ 2$: Given initial estimates $f_k^0(x_k)$ of $f_k(x_k)$, smooth the partial residuals

$y - \hat{\alpha} - \sum_{k \neq j} f_k^0(x_k)$ on $x_j,\ j = 1,2,\ldots,p$ to obtain an improved estimate $f_j^1(x_j)$.

$Step\ 3$: Use $f_j^1(x_j)$ to obtain improved estimates of $f_k^1(x_k)$ of $f_k(x_k), k \neq j,$
$k = 1,2,\ldots,p$

by smoothing the partial residual of $f_k(x_k)$ on x_k for each x_k.

$Step\ 4$: Continue steps 2 and 3 until the functions f_j do not change.

Fitting GAM Models

The generalized additive models are fit to count data or binary response data by maximizing a penalized log likelihood or a penalized log partial-likelihood. To maximize it, the backfitting procedure is used in conjunction with a maximum likelihood or maximum partial likelihood algorithm (Hastie and Tibshirani (1990) and Wood (2004)). The Newton-Raphson method for maximizing log-likelihoods in these models can be presented in a IRLS (iteratively reweighted least squares) form. It involves a repeated weighted linear regression of a constructed response variable on the covariates: each regression yields a new value of the parameter estimates which give a new constructed variable, and the process is iterated. In the generalized additive model, the weighted linear regression is simply replaced by a weighted backfitting algorithm (Hastie and Tibshirani (1986)).

Backfitting Algorithm for GAMs

Step 1: Initialize $\hat{\alpha} = g(1/N\sum_{i=1}^{N} y_i),\ f_j^0 = 0,\ j = 1,2,\ldots,p.$

Step 2: Construct an adjusted dependent variable z_{ij} as

$z_i = \eta_i^0 - (y_i - \mu_i^0)(\partial\eta_i / \partial\mu_i)_0$

$\eta_i^0 = g(\mu_i^0) = \alpha^0 + \sum_{j=1}^{p} f_j^0(x_{ij})$ and $\mu_i^0 = g^{-1}(\eta_i^0).$

Step 3: Compute weights

$w_i = (\partial\mu_i / \partial\eta_i)_0 (V_i^0)^{-1},$

where V_i^0 is the variance of y at μ_i^0,

Step 4: Fit a weighted additive model to z_i and compute f_j^1, η^1, and μ_i^1,

the second stage estimates of f_j, η, and μ_i.

Step 5: Repeat steps 2 and 3 replacing η^0 by η^1 until the difference between

two successive values of η is less than a small prespecified number and

convergence is obtained.

Estimation of Generalized Additive Models using Penalized Regression Method

- **Algorithm for Penalized Iteratively Re-weighted Least Squares (PIRLS) (Wood (2006)**
- The GAMs are fit to data by maximizing a penalized log likelihood or a penalized log partial-likelihood. The following algorithm is used to implement these methods. The R-package mgcv (Wood (2012)) was used for computations.

PIRLS Algorithm

$Step\,1$: Initialize $\hat{\alpha} = g(1/N\sum_{i=1}^{N} y_i), s_j^0 = 0, j = 1,2,\ldots,p.$

$Step\,2$: Construct an adjusted dependent variable z_{ij} as

$z_i = \eta_i^0 - (y_i - \mu_i^0)(\partial\eta_i / \partial\mu_i)_0$

$\eta_i^0 = g(\mu_i^0) = \alpha^0 + \sum_{j=1}^{p} s_j^0(x_{ij})$ and $\mu_i^0 = g^{-1}(\eta_i^0).$

$Step\,3$: Compute weights

$w_i = (\partial\mu_i / \partial\eta_i)_0 (V_i^0)^{-1},$

where V_i^0 is the variance of y at μ_i^0,

PIRLS Algorithm (Continued)

Step 4: *Penalized Spline Regression*

Minimize $\left\|\sqrt{W}(z - X\beta)\right\|^2 + \lambda\beta' S\beta$ with respect to β, where X is the matrix of data on basis functions used to represent the regression function, W is a diagonal matrix with i-th diagonal element w_i, S is a matrix of known coefficients in the penalty function $\beta' S\beta$ and λ is a smoothing parameter. Compute s_j^1, η^1, and μ_i^1, the second stage estimates of s_j, η, and μ_i.

Step 5: Repeat steps 2 - 4 replacing η^0 by η^1 until the difference between two successive values of η is less than a small prespecified number and convergence is obtained.

The generalized additive Gaussian model

- The generalized additive Gaussian model assumes that
- The adjusted dependent variable *z* and the weights *w* used in the algorithm above are
- The functions *s1, s2,…, sp* are estimated by an algorithm like the one described earlier.

An Empirical Application of GAM Gaussian Model to Wages

Variable Definitions and Data Description

- The dependent variable is *WAGE*.
- *WAGE* = Earnings per hour
- *EDUC* = Years of education
- *EXPER* = Post education years experience
- *HRSWK* = Usual hours worked per week
- *MARRIED* = 1 if married, 0 otherwise.

Table 1: Summary Statistics

Variable	Obs	Mean	Std Dev	Min	Max
WAGE	1000	20.20122	12.1038	2.03	72.13
EDUC	1000	10.689	2.44013	1	16
EXPER	1000	26.501	12.99041	3	64
HRSWK	1000	39.24	11.44611	0	99

Source: Data source: Dr. Kang Sun Lee, Louisiana Department of Health and Human Services reproduced in Principles of Econometrics by Hill, Griffith and Lu (2012).

Nonparametric Exploration of Nonlinearity

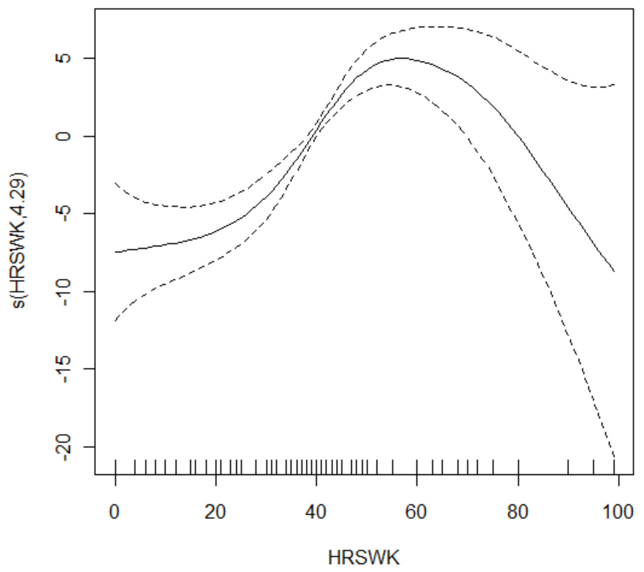

Figure 1: Partial residuals plot of HRSWK

The partial residual plot indicates that the link function is highly nonlinear in the variable HRSWK. Hence nonparametric smooth functions of HRSWK are included in our GAM models.

GAM Gaussian Models

$Model\,1: GLM\ \ Normal\ \ with\ Identity\ Link$

$\eta = g(\mu) = \beta_1 + \beta_2 EDUC + \beta_3 EXPER + \beta_4 HRSWK$

$Model\ 2: GLM\ \ Normal\ \ with\ a\ Quadratic\ Term\ \ and\ \ Identity\ Link$

$\eta = g(\mu) = \beta_1 + \beta_2 EDUC + \beta_3 EXPER + \beta_4 HRSWK$

$+ \beta_4 HRSWKSQ$

$Model\ 3: Generalized\ \ Additive\,\mathrm{Re}\,gression\ \ Models\ with\ Identity\ Link$
$for\, Wages$

$\eta = g(\mu) = \beta_1 + \beta_2 EDUC + \beta_3 EXPER + \beta_4 HRSWK + s(HRSWK)$

$Model\ 4: GLM\ \ Normal\ \ with\ Identity\ Link\ \ for\,\ln(Wages)$

$\eta = g(\mu) = \beta_1 + \beta_2 EDUC + \beta_3 EXPER + \beta_4 HRSWK$

$Model\ 5: GAM\ \ Normal\ \ with\ Identity\ Link\ \ for\,\ln(Wages)$

$\eta = g(\mu) = \beta_1 + \beta_2 EDUC + \beta_3 EXPER + s(HRSWK)$

$Model\ 6: GAM\ \ Normal\ \ with\ Identity\ Link\ \ for\,\ln(Wages)\ and\ \ Interaction$

$\eta = g(\mu) = \beta_1 + \beta_2 EDUC + \beta_3 EXPER + \beta_4 EXPER * MARRIED + s(HRSWK)$

$Model\ 7: GAM\ \ Normal\ \ with\ Identity\ Link\ \ for\,\ln(Wages)\ and\ \ Nonparametric$
$Interaction$

$\eta = g(\mu) = \beta_1 + \beta_2 EDUC + \beta_3 EXPER + s(EXPER * MARRIED) + s(HRSWK)$

Comparing the GAM Gaussian Models

Table 9: Models and the AICs

MODEL	AIC
1	7531.5
2	7527
3	7508.195
4	7524.032
5	7487.733
6	7479.896
7	7467.548

Results for GAM Gaussian Models

The estimation and results are presented in tables 2 through 8.

- A comparison of models using the AIC presented in Table 9 suggests that models 6 and 7, which employ *ln(WAGES)* as the response variable and allow for interaction between *EXPER* and *MARRIED* have the lowest AICs among the models considered and are therefore the best models.
- Model 7, a generalized additive model for *ln(WAGES)*, which includes a nonparametric interaction term between *EXPER* and *MARRIED* has the lowest AIC and UBRE score among the seven models studied .
- Model 6, which allows parametric interaction term between *EXPER* and *MARRIED* has the second lowest AIC and UBRE score.
- At the other extreme, Model 1, a generalized linear Gaussian model for *WAGES*, has the highest AIC suggesting that it is the poorest model among all models considered.

The Generalized Additive Logit Model

The generalized additive Logit model assumes that

$$g(\mu_i) = \mathrm{logit}(\mu_i) = \ln\left(\frac{\mu_i}{1-\mu_i}\right) = \alpha + \sum_{j=1}^{p} s_j(x_{ij}),$$

where $\mu_i = p_i = E(y_i = 1 | x_i) = \exp\{\alpha + \sum_{j=1}^{p} s_j(x_{ij})\} / \exp\{1 + \alpha + \sum_{j=1}^{p} s_j(x_{ij})\}.$

The adjusted dependent variable z and the weights w used in the algorithm abo

$$z_i = \eta_i + (y_i - p_i) / p_i(1 - p_i),$$

$$w_i = p_i(1 - p_i)$$

where $p_i = g^{-1}(\eta_i), \eta_i = \alpha + \sum_{j=1}^{p} s_j(x_{ij})$

The functions $s_1, s_2, \ldots, s_p$ are estimated by an algorithm like the one described earlier.

An Empirical Application of GAM Logit Model to Brand Choice

- **GAM Logit applied to Cracker Data**
- The dataset is from Jain et al. (1994) and Paap and Franses (2000). We consider an optical scanner panel data set on purchases of saltine crackers in the Rome (Georgia) market, collected by Information Resources Incorporated. The data set contains information on all purchases of crackers (3292) of 136 households over a period of two years, including brand choice, actual price of the purchased brand and shelf price of other brands, and whether there was a display and/or newspaper feature of the considered brands at the time of purchase. The data file contains 17 variables, arranged in five rows for each observation as follows.
- *ID*: individuals identifiers
- *CHOICE*: one of sunshine, kleebler, nabisco, private
- *DISP.z*: is there a display for brand z ?
- *FEAT*.z: is there a newspaper feature advertisement for brand z ?
- *PRICE*.z: price of brand z

Data Characteristics of Cracker Data

- Table 10 shows some data characteristics. There are three major national brands in our database, that is, Sunshine, Keebler and Nabisco with market shares of 7%, 7% and 54% respectively. The local brands are collected under `Private label', which has a market share of 32%

Table 10. Some data characteristics of Cracker data

BRAND	Sunshine	Kleebler	Nabisco	Private Label
MKT. SHARE	0.07	0.07	0.54	0.32
DISPLAY a	0.13	0.11	0.34	0.10
FEATURE b	0.04	0.04	0.09	0.05
AVERAGE PRICE	0.96	1.13	1.08	0.68
ESTIMATED S	0.39	0.07	0.35	0.19
BRAND K	0.09	0.50	0.30	0.11
SWITCHING N	0.04	0.04	0.84	0.08
PROBABILITIES c p	0.04	0.03	0.12	0.81

Variable Definitions

- The dependent variable *CHOICE* is defined as follows.
- *NABISCO* = 1 if Sunshine is chosen,
- = 0 if any other brand is chosen
- *PRICE.SUNSHINE* = Price of a box of Sunshine
- *PRICE.KEEBLER* = Price of a box of Sunshine
- *PRICE.NABISCO* = Price of a box of Sunshine
- *PRICE.PRIVATE* = Price of a box of Sunshine
- *DISP.SUNSHINE* = 1 if Sunshine is displayed at time of purchase, otherwise = 0
- *DISP.KEEBLER* = 1 if Keebler is displayed at time of purchase, otherwise = 0
- *DISP.NABISCO* = 1 if Nabisco is displayed at time of purchase, otherwise = 0
- *DISP.PRIVATE* = 1 if Private Label is displayed at time of purchase, otherwise = 0

Nonparametric Exploration of Nonlinearity

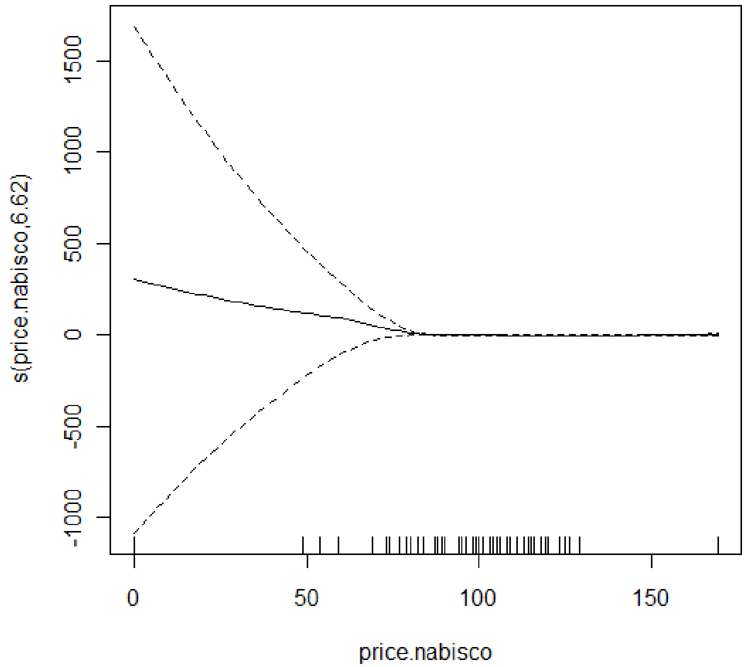

Figure 2: Partial residuals plot of price.nabisco

The partial residual plot of price.nabisco suggests high nonlinearity of the link function in the variable price.nabisco, Accordingly, the GAM Logit models include nonparametric smooth terms in price.nabisco.

GAM Logit Models

Model 1: Logit Regression Model 1

$$\eta = g(\mu) = \beta_1 + \beta_2 PRICE.NABISCO + \beta_3 DISP.NABISCO$$

$$+ \beta_4 FEAT.NABISCO$$

Model 2: Logit Regression Model 2

$$\eta = g(\mu) = \beta_1 + \beta_2 PRICE.KEEBLER + \beta_3 DISP.KEEBLER$$

$$+ \beta_4 FEAT.KEEBLER$$

Model 3: Generalized Additive Logit Regression Models with a Nonparametric Interaction Term

$$\eta = g(\mu) = \beta_1 + s(PRICE.NABISCO) + \beta_3 DISP.NABISCO$$

$$+ \beta_4 FEAT.NABISCO$$

Model 4: Generalized Additive Logit Regression Models with a Nonparametric Interaction Term and a dummy variable for whether a particular brand is featured at the time of purchase

$$\eta = g(\mu) = \beta_1 + \beta_2 FEAT.NABISCO + \beta_3 FEAT.SUNSHINE$$

$$+ \beta_4 FEAT.KEEBLER + s(PRICE.NABISCO) + s(PRICE.SUNSHINE)$$

$$+ s(PRICE.KEEBLER)$$

Model 5: Logit Regression Model with a Nonparametric Smooth for the Price of each Brand

$$\eta = g(\mu) = \beta_1 + s(PRICE.NABISCO) + s(PRICE.SUNSHINE)$$

$$+ s(PRICE.KEEBLER) + s(PRICE.PRIVATE)$$

Comparing the GAM Logit Models

Table 16: Models and the AICs

MODEL	AIC
1	4393.78
2	4538.65
3	4351.397
4	4210.49
5	3758.557

RESULTS FOR GAM LOGIT MODELS

Model 5, a GAM, which includes a nonparametric smooth term for the price of each brand, has the lowest AIC and UBRE score among the five models studied and is therefore the best model.

The Logit regression model (Model 2), which uses the characteristics of the competing brand, Keebler has the highest AIC. This is not surprising since the Logit model misses the nonlinearity in the price of each brand in the link function.

Model 5 also has the lowest deviance of 3945.672 on 3275 degrees of freedom, while Model 2 has the highest deviance of 4530.7 on 3288 degrees of freedom.

Results for GAM Logit Models

- The statistical significance of brand prices differs markedly between the Logit regression model and the various GAM Logit models employed here.
- The signs of the coefficient estimates are all expected in all of the models but Model 2. For instance, the sign of *DISP.KEEBLER* is positive in model 2 indicating that the odds of choosing *NABISCO* over *KEEBLER* increase if *KEEBLER* is displayed at the time of purchase and decrease if *NABISCO* is displayed at the time of purchase.
- The analysis of deviance in Tables 13, 14, and 15 indicates significant nonlinear contribution from the variables *PRICE.NABISCO* as well as other brand price variables.
- The high degree of nonlinearity in *PRICE.NABISCO* is also seen in the partial residual smoothing plot of PRICE.NABISCO in Figure 1. The dotted curves around the solid curve represent +-2 standard errors around the solid curve. The only surprising result is the negative sign of the variable *FEAT.KEEBLER* in Table 12.

Generalized Additive Poisson and Negative Binomial Models

- In the Poisson regression model the outcome, yi is a count variable, such as number of visits to a doctor's office as in Gurmu (1997). We wish to model p(yi|xi1, xi2,…, xip) the probability of an event given factors xi1, xi2,…, xip. The Poisson regression model assumes that the link function is linear:

$$\ln \mu_i = \beta_0 + x_{i1}\beta_1 + \ldots + x_{ip}\beta p$$

- The generalized additive Poisson model assumes instead that

$$\ln \mu_i = \beta_0 + f_1(x_{i1}) + \ldots + f_1(x_{ip})$$

- The functions *f1, f2,…, fp* are estimated by the PIRL algorithm described earlier.
- The generalized additive models are fit to count data or binary response data by maximizing a penalized log likelihood or a penalized log partial-likelihood. To maximize it, the backfitting procedure is used in conjunction with a maximum likelihood or maximum partial likelihood algorithm (Hastie and Tibshirani (1990) and Wood (2004)). The Newton-Raphson method for maximizing log-likelihoods in these models can be presented in a IRLS (iteratively reweighted least squares) form. It involves a repeated weighted linear regression of a constructed response variable on the covariates: each regression yields a new value of the parameter estimates which give a new constructed variable, and the process is iterated. In the generalized additive model, the weighted linear regression is simply replaced by a weighted backfitting algorithm (Hastie and Tibshirani (1986)).

An Empirical Application f GAM Poisson Model to Doctor Visits data

DATA AND VARIABLE DEFINITIONS

The data are a 1986 cross section sample from the US consisting of 485 observations and are drawn from Gurmu (1997). These data came from the 1986 Medicaid Consumer Survey sponsored by the Health Care Financing Administration. The following variables were used in econometric analysis.

DOCTOR = the number of doctor visits
CHILDREN = the number of children in the household
ACCESS = is a measure of access to health care
HEALTH = a measure of health status (larger positive numbers are associated with poorer health

Table 17: Summary Statistics for the Doctor data

Variable	Obs.	Mean	Std. Dev.	Min	Max
DOCTOR	485	1.610	3.346809	0	48
CHILDREN	485	2.264	1.319136	1	9
ACCESS	485	0.3812	0.186105	0	0.92
HEALTH	485	-0.00004124	1.433520	1	1

Nonparametric Exploration of Nonlinearity

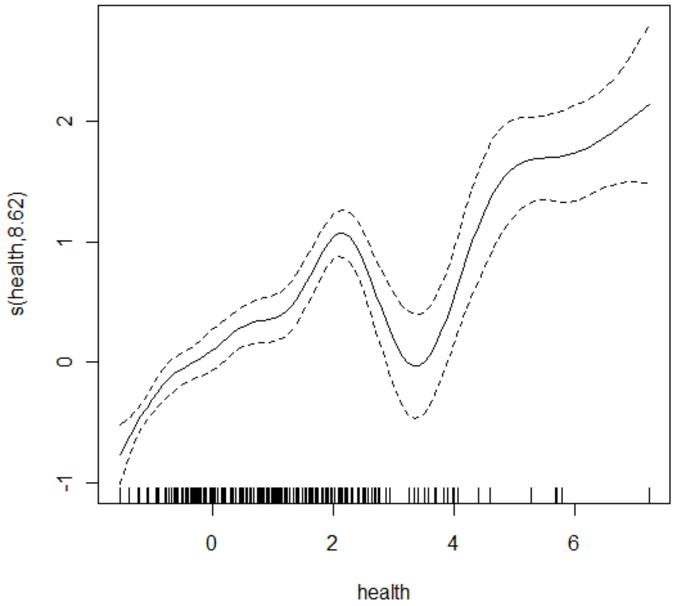

Fig. 3: Partial Residual plot of HEALTH

The partial residual plot of HEALTH suggests that the link function is highly nonlinear in the variable HEALTH. Accordingly, nonparametric smooth terms in the variable HEALTH are included in both GAM Poisson and GAM Negative Binomial regression models.

GAM Poisson and Negative Binomial Models

Model 1: Poisson Regression Model with Identity Link

$$\eta = g(\mu) = \beta_1 + \beta_2 CHILDREN + \beta_3 ACCESS + \beta_4 HEALTH$$

Model 2: Negative Binomial Regression with Identity Link

$$\eta = g(\mu) = \beta_1 + \beta_2 CHILDREN + \beta_3 ACCESS + \beta_4 HEALTH$$

Model 3: Generalized Additive Poisson Regression Model

$$\eta = g(\mu) = \beta_1 + \beta_2 CHILDREN + \beta_3 ACCESS + s(HEALTH)$$

Model 4: Generalized Additive Negative Binomial Regression Model

$$\eta = g(\mu) = \beta_1 + \beta_2 CHILDREN + \beta_3 ACCESS + s(HEALTH)$$

Model 5: Generalized Additive Poisson Regression Model with smooth nonparametric terms for access and health

$$\eta = g(\mu) = \beta_1 + \beta_2 CHILDREN + s(ACCESS) + s(HEALTH)$$

Model 6: Generalized Additive Negative Binomial Regression Model with smooth nonparametric terms for access and health

$$\eta = g(\mu) = \beta_1 + \beta_2 CHILDREN + s(ACCESS) + s(HEALTH)$$

Model 7: Generalized Additive Poisson Regression Model with smooth nonparametric terms for access and health and nonparametric interaction between ACCESS and HEALTH

$$\eta = g(\mu) = \beta_1 + \beta_2 CHILDREN + s(ACCESS) + s(HEALTH) + s(ACCESS * HEALTH)$$

Model 8: Generalized Additive Negative Binomial Regression Model with smooth nonparametric terms for access and health and nonparametric interaction between ACCESS and HEALTH

$$\eta = g(\mu) = \beta_1 + \beta_2 CHILDREN + s(ACCESS) + s(HEALTH) + s(ACCESS * HEALTH)$$

Comparing the Poisson and Negative Binomial Nodels

Table 28: GLM and GAM Poisson Models and the AICs

MODEL	AIC
1P	2179.487
2P	2140,685
3P	2067.047
4P	2024.952
5P	2060.085

Table 29: GLM and GAM Negative Binomial Models and the AICs

MODEL	AIC
1NB	1692.614
2NB	1686.335
3NB	1653.477
4NB	1641.151
5NB	1653.908

Results for GAM Poisson and Negative Binomial Models

- The estimation and results are presented in tables 11 through 15. A comparison of models using the AIC is presented in Tables 28 and 29. Table 28 compares GLM and GAM Poisson models and Table 29 compares GLM and GAM Negative Binomial models.
- Model 4P, a GAM Poisson model, which includes nonparametric smooth terms for *ACCESS*, *HEALTH* and the interaction *ACCESS*HEALTH* has the lowest AIC and UBRE score among the five models studied and is therefore the best model.
- The GLM Poisson regression model (Model 1P), which uses the predictors *CHILDREN*, *ACCESS* and *HEALTH* has the highest AIC. This is not surprising since the Poisson model misses the nonlinearity in the predictors *CHILDREN*, *ACCESS*, *HEALTH*, and interaction *ACCESS*HEALTH*.
- Similarly, Model 4NB, a GAM Negative Binomial model, which includes nonparametric smooth terms for *ACCESS*, *HEALTH* and the interaction *ACCESS*HEALTH* has the lowest AIC and UBRE score among the five models studied and is therefore the best model among the Negative Binomial GLM and GAM models.
- Finally, among the Poisson and Negative Binomial GLM and GAM models studied here, the clear winner is Model 4NB, which has the lowest AIC and UBRE scores. Model 4P also has the lowest deviance of 1380.621 on 471.0001 degrees of freedom among the Poisson GLM and GAM models, while Model 1P has the highest deviance of 1508.8 on 481 degrees of freedom.
- The statistical significance of the predictors *CHILDREN*, *ACCESS* and *HEALTH* are generally similar among all models since all of the variables are found highly significant in all of the models. The signs of the coefficient estimates are all expected in all of the models but Model 2. For instance, the sign of *ACCESS* is positive as is the sign of *HEALTH*, but the sign of *CHILDREN* is negative indicating that as access to health services increases, individuals make more trips to a doctor's office even for routine checkups. Individuals with poor health with higher numbers on *HEALTH* tend to make more trips to a doctor's office as reflected in the positive sign of *HEALTH*. The negative sign of the variable *CHILDREN* may be surprising, but may indicate that households with fewer children make more trips to a doctor's office perhaps because these households in the sample are more health-conscious than households with too many children in this sample. The analysis of deviance in Tables 21 through 27 indicates significant nonlinear contribution from the variables *ACCESS, HEALTH,* and interaction *ACCESS*HEALTH*. The high degree of nonlinearity in *HEALTH* is also seen in the partial residual smoothing plot of *HEALTH* in Figure 3. The dotted curves around the solid curve represent +-2 standard errors around the solid curve.

Conclusions

- The paper has studied four generalized additive Gaussian, Logit, Poisson, and Negative Binomial regression models as alternatives to the generalized linear Normal, Logit, Poisson, and Negative Binomial regression models respectively. The Gaussian generalized additive models (GAMs) were applied to data on wages, education, and experience. The logit GAMS were applied to data on brand choice. The Poisson and Negative Binomial GAMs were applied to data on doctor visits.
- In all of the empirical applications, each of the GAMs provides a much better fit than the corresponding generalized linear model (GLM) as reflected in lower AICs and lower deviances.
- The econometric techniques used in the paper are widely applicable to the analysis of count, binary response and duration types of data occurring frequently in economics and business. GAMs extend nonparametric regression to more than one regressor helping to circumvent the curse of dimensionality. Another advantage of the GAM approach used in the paper is that we take account of nonlinearities and interactions among explanatory variables non-parametrically. Nevertheless, the GAMs are not without drawbacks. The computational algorithms are complex and interpretations can be difficult. These models are useful mainly when simple models for the linear predictor provide an inadequate fit for the data.

References

- Baltas, George 1997. "Determinants of store brand choice: a behavioral analysis", Journal of Product & Brand Management, Vol. 6 (5), 315 – 324.
- Cameron, A. C. and P. Trivedi, 1998. Microeconometrics, Cambridge University Press.
- Greene, W. H. (2008). Econometric Analysis, Pearson/Prentice Hall, New York.
- Guadagni, Peter M. and Little, John D C 1983. A Logit Model of Brand Choice Calibrated on Scanner Data Marketing Science, Summer Vol. 2(3), 203-238.
- Gurmu, Shiferaw (1997) “Semiparametric estimation of hurdle regression models with an application to medicaid utilization”, *Journal of Applied Econometrics*, 12(3), 225-242.
- Hastie, T. & Tibshirani, R. (1986), Generalized Additive Models, Statistical Science 1, 297-318
- Hastie, T. & Tibshirani, R., 1990. Generalized Additive Models, Chapman and Hall, London.
- Hill, R. C., W. E. Griffiths, and G. C. Lim (2012). Principles of Econometrics, Wiley, New Jersey
- .
- Jain, Dipak C., Naufel J. Vilcassim and Pradeep K. Chintagunta (1994) “A random–coefficients logit brand–choice model applied to panel data”, *Journal of Business and Economics Statistics*, **12(3)**, 317-328
- McCullagh, P. and J. Nelder, 1989. Generalized Linear Models, Chapman and Hall, London.
- Manski, Charles and Daniel McFadden 1981. Structural Analysis of Discrete Data and Econometric Applications, MIT Press, Cambridge.

References

- Paap, R. and Philip Hans Frances (2000) “A dynamic multinomial probit model for brand choices with different short–run effects of marketing mix variables”, *Journal of Applied Econometrics*, **15(6)**, 717–744.

- Wood, S. N. 2012, R-package mgcv

- Wood, S.N. 2006. Generalized Additive Models: an introduction with R, CRC, Boca Raton.

- Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semi-parametric generalized linear models. Journal of the Royal Statistical Society (B) 73(1):3-36

- Wood, S.N. (2004) Stable and efficient multiple smoothing parameter estimation for generalized additive models. *J. Amer. Statist. Asociation.* 99:673-686.

Business and Economics Related Journals

- Arabian Journal of Business and Management Review
- Business and Economics Journal
- International Journal of Economics and Management Sciences
- Journal of Public Affairs

Business and Economics Related Conferences

- ✓ International conference on Business, Economics and Management
- ✓ International Conference on Advertising and Marketing Expo

OMICS Group **Open Access Membership**

OMICS publishing Group Open Access Membership enables academic and research institutions, funders and corporations to actively encourage open access in scholarly communication and the dissemination of research published by their authors.

For more details and benefits, click on the link below:

http://omicsonline.org/membership.php