Appraisal Using Generalized Additive Models

*R. Kelley Pace**

Abstract. **Many of the results from real estate empirical studies depend upon using a correct functional form for their validity. Unfortunately, common parametric statistical tools cannot easily control for the possibility of misspecification. Recently, semiparametric estimators such as generalized additive models (GAMs) have arisen which can automatically control for additive (in price) or multiplicative (in ln(price)) nonlinear relations among the independent and dependent variables. As the paper shows, GAMs can empirically outperform naive parametric and polynomial models in exsample predictive behavior. Moreover, GAMs have well-developed statistical properties and can suggest useful transformations in parametric settings.**

Introduction

Many functional forms of the variables and parameters lead to pricing functions that agree with the information amassed by the substantial theoretical and empirical work in hedonic pricing and mass assessment.[1] Consequently, the exact specification to adopt remains one of the central uncertainties of empirical work, especially since the ‘‘wrong’’ functional form leads to all sorts of disastrous consequences for traditional estimators. In response to this problem, many nonparametric estimators have been proposed which adapt to the data and do not require an *a priori* functional specification. However, nonparametric estimator performance typically declines as the dimensionality of the problem increases.[2] As a compromise, various semiparametric estimators have arisen that possess the adaptive traits of nonparametric regression while retaining the estimation efficiency of parametric estimators. Projection pursuit (Friedman and Stuetzle, 1981), neural nets, alternating conditional expectations (Brieman and Friedman, 1985), additivity and variance stabilization (Tibshirani, 1988), regression trees (Brieman, Friedman, Olshen and Stone, 1984), Π (Brieman, 1991), multivariate adaptive regression splines (Friedman, 1991) and sliced inverse regression (Duan and Li, 1991) estimators represent some of the efforts in this direction.[3]

In real estate, various applications of nonparametric and semiparametric regression have appeared from time to time. For example, analysis of pairs of houses matched on all but two or fewer characteristics via graphical methods would qualify as nonparametric estimation. Isakson (1986) used a form of nearest neighbor

*Department of Finance, Louisiana State University, Baton Rouge, LA 70803 or kelleypace@compuserve.com.

nonparametric estimation.[4] Sunderman, Birch, Cannaday and Hamilton (1990) explored the relation between assessed value and market price using a bivariate spline regression estimator. Meese and Wallace (1991) and Pace (1993b) applied nonparametric multivariate regression estimators to real estate data. Meese and Wallace employed locally weighted regression (see Cleveland and Devlin, 1988) to form hedonic price indices. They conducted diagnostics on the sample fit to document its generally good performance. Using two hedonic pricing examples, Pace (1993) demonstrated the kernel nonparametric regression estimator could out-perform ordinary least squares (OLS) in ex-sample prediction. Pace (1995) applies the semiparametric index estimator ($y=g(X\beta)+\varepsilon$) of Hardle and Stoker (1989) and Powell, Stock and Stoker (1989) to real estate data and showed it could compete with OLS and the kernel regression estimator.

Coulson (1992) used a model incorporating some parametric components and a bivariate spline estimator for the nonparametric component. Anglin and Gencay (1993) used a model incorporating parametric components and a multivariate kernel estimator for the nonparametric component to investigate the hedonic pricing functional form. Their semiparametric estimates clearly outperformed their best parametric ones.

Both the Coulson and the Anglin and Gencay estimators fall in the category of Generalized Additive Models (GAMs). GAMs constitute perhaps the simplest class of semiparametric estimators in terms of computation and visualization. Essentially, GAMs in Equation (1) estimate the dependent variable as a sum of functions of the independent variables.

$$y=f_1(X_1)+f_2(X_2)+\ldots+f_k(X_k)+\varepsilon. \tag{1}$$

Naturally, GAMs include linear models.

$$y=X_1\beta_1+X_2\beta_2+\ldots+X_k\beta_k+\varepsilon.$$

GAMs can extend their range in the same way linear models extend theirs through transformations and functions of the individual regressors.

$$y=\beta_1(X_1)+f_2(X_2)+f_3(X_1X_2)+f_4(g(X_3))+f_5(X_2X_3)+\ldots+f_k(X_k)+\varepsilon.$$

Effectively, Coulson used a model involving the first two terms while Anglin and Gencay used a model involving the first and the fifth terms. Specifically, Anglin and Gencay used a kernel estimator involving six dimensions or characteristics. Naturally, the dependent variable, y, could represent a transformation of some other variable (*e.g.*, $y=\ln(z)$ or $y=z^{1/2}$), which means GAM could also include multiplicative modeling of z.

Graphs of the estimated transformation $\tilde{f}_i(X_i)$ versus X_i constitute one of the main products of the GAM estimator. These may have interest in their own right or can serve as guides to transforming variables in the ordinary linear model. Alternatively, these estimated transformations allow one to check on the linearity of and y in a posited linear model.

This article examines the computation of the GAM estimator in section two and applies the estimator in section three. Specifically, beginning with a typical semilogarithmic specification using 442 observations from the Memphis Multiple Listing Service (MLS), the GAM estimator suggests transformations that lead to a linear double logarithmic model. For comparison, section three includes the semilogarithmic and double logarithmic GAM and polynomial regression models. Section three also includes a cross-validation prediction experiment that shows the superiority of the GAM and the retransformed linear model to the original semiparametric and polynomial regression models. Section four summarizes the key results.

Computation of Generalized Additive Models

As mentioned previously, the GAM estimator is one of the simplest semiparametric estimators to compute and visualize. One minimizes some loss function, typically squared error, through the choice of functions as opposed to individual parameters.

$$\min \tilde{e}'\tilde{e} \text{ where } \tilde{e}=(y-\tilde{f}_1(X_1)+\tilde{f}_2(X_2)+...+\tilde{f}_k(X_k)). \tag{2}$$

Hastie and Tibshirani (1990) extensively discussed the use of the backfitting algorithm that iteratively minimizes Equation (2) an estimated function at a time. Let i (i=1,2,...k) represent the individual estimated functions ($\tilde{f}_i(\cdot)$) and j (j=1,2,...m) represent the iteration. For each iteration j one minimizes Equation (2) with respect to each of the estimated functions $\tilde{f}_i(\cdot)$. One continues the iterations until convergence. Hastie and Tibshirani prove this algorithm will converge to an unique solution independent of the starting values for symmetric smoothing functions such as smoothing or regression splines. Interestingly, if for all i $\tilde{f}_i(\cdot)=X_i\beta_i$, the backfitting algorithm yields, albeit slowly, the least-squares solution for a squared-error loss function.

One can employ a variety of methods to estimate the functions $\tilde{f}_i(\cdot)$. For example, one can employ the kernel method, locally weighted smoothing, smoothing splines, regression splines, nearest neighbor and polynomials.[5]

The advantage to GAM, as opposed to purely nonparametric methods, lies in the reduction of the problem of estimating nonparametric surfaces to a sequence of bivariate smoothing problems. These allow (1) visual inspection of the smooth; and (2) the estimates converge as rapidly as parametric estimators.

In the following estimates, I used smoothing splines as the bivariate or scatterplot smoother. Smoothing splines minimize Equation (3):

$$\min [(y-\tilde{f}_i(X_i))'(y-\tilde{f}_i(X_i))+\lambda \int (\tilde{f}_i''(t))^2dt] \tag{3}$$

where λ represents a roughness penalty. If $\tilde{f}_i(X_i)$ had a linear form, the second derivative of a linear function, $\tilde{f}_i''(X_i)$, would be 0. Alternatively, if $\tilde{f}_i(X_i)$ rapidly changed with X_i, the second derivative, $\tilde{f}_i''(X_i)$, would have a large magnitude. If λ equals 0, the smoothing spline would cause $\tilde{f}_i(X_i)$ to match every point in y, resulting in no error. If λ equals infinity, the heavy penalty on roughness would cause $\tilde{f}_i(X_i)$ to return a linear fit, resulting in the least squares regression line.

Naturally, the parameter λ greatly affects the smoothing splines behavior. A small value of λ means $\tilde{f}(X_i)$ is very flexible in the same way a high order polynomial is flexible. As most individuals do not have much prior information concerning λ, Buja, Hastie and Tibshirani (1989) provided a way of measuring the equivalent degrees-of-freedom sacrificed by making $\tilde{f}_i(X_i)$ very flexible. This greatly reduces the difficulty of smoothing parameter selection. The equivalence between λ and degrees-of-freedom makes it possible to perform approximate inference for GAM.

As a final note concerning λ, by appropriate selection of $l_i = g(X_i)$ one could maximize the linearity of $\tilde{f}_i(l_i)$. This could greatly reduce the value of $\tilde{f}_i''(l_i)$ which reduces the sensitivity of the overall solution to an inappropriate choice of λ. We shall use this technique latter in the actual estimation.

Polynomials constitute the traditional way of modeling functions of X_i ($\tilde{f}_i(X_i) = a + bX_i + cX^2)_i + \ldots + pX_i^{p-1}$) in linear models. A series of polynomials leads to a model linear-in-the-parameters which least squares can fit directly. The difference between nonparametric smoothers such as smoothing splines or the kernel method and polynomial regression lies in the local nature of the nonparametric estimator fits versus the global nature of the polynomial regression estimator fit. If the (y, X_i) plot linearly over part of X_i but have a curved portion over another part of X_i, nonparametric estimators can follow this (even with two degrees-of-freedom). A second degree polynomial will have some curvature over all of X_i. Hence, polynomials of limited degree do react to nonlinearities. Their global fit means any nonlinearity polynomials detect in $\tilde{f}_i(X_i)$ for some values of X_i will be spread over all X_i.

Finally, GAM are a generalization of generalized linear models (GLM) of McCullagh and Nelder (1989). GLM parametrically fits models $y = g(X\beta) + \varepsilon$ for different distributions (with different variance specifications). Consequently, one can easily apply GAM using other distributions such as Poisson, gamma and multinomial. Hence, one can estimate count or duration data, survival data and probabilities with the same flexibility in functional form.

Estimation Results

This section provides an empirical illustration of the advantages of GAM using real estate data. Specifically, the first subsection provides the models and variables used in the latter subsections, the second subsection discusses the data, the third subsection focuses upon the graphs of the estimated functions versus their arguments, the fourth subsection presents the global sample estimates and the final subsection contains a predictive cross-validation trial of the various estimators and models.

Models and Variables

AREAID, an dichotomous variable, refers to one of twenty-four districts within Memphis. *CARPORTS*, *GARAGE*, *CENAC*, *NONAC*, *FIRE*, *POOL*, *BRICK* and *ALUM* (aluminum siding) are also dichotomous variables with one representing the presence of the characteristic. *KITSF* (kitchen area) and *NONKITSF* (non-kitchen area) added together equal total area. *LOTSF* denotes lot area in square feet. *BATHS* denotes number of bathrooms. ln(*AGE*) actually equals ln(*AGE*+*e*).[6] In the results I will make reference to the following models.

Common Model

$$\ln(\mathit{Price}) = \beta_1 \text{intercept} + \beta_{2\text{-}24}\mathit{AREAID} + \beta_{25}\mathit{CARPORTS} + \beta_{26}\mathit{GARAGE} + \beta_{27}\mathit{CENAC} \\ + \beta_{28}\mathit{NONAC} + \beta_{29}\mathit{FIRE} + \beta_{30}\mathit{POOL} + \beta_{31}\mathit{BRICK} + \beta_{32}\mathit{ALUM}.$$

Models 1-6 Common Model +. . .

1. $\beta_{33}\mathit{BATHS} + \beta_{34}\mathit{NONKITSF} + \beta_{35}\mathit{KITSF} + \beta_{36}\mathit{LOTSF} + \beta_{37}\mathit{AGE}$.
2. $\beta_{33}\mathrm{BATHS} + \beta_{34}\mathit{BATHS}^2 + \beta_{35}\mathit{NONKITSF} + \beta_{36}\mathit{NONKITSF}^2 + \beta_{37}\mathit{KITSF}$ $+ \beta_{38}\mathit{KITSF}^2 + \beta_{39}\mathit{LOTSF} + \beta_{40}\mathit{LOTSF}^2 + \beta_{41}\mathit{AGE} + \beta_{42}\mathit{AGE}^2$.
3. $\beta_{33}\mathit{BATHS} + \beta_{34}\mathit{BATHS}^2 + \beta_{35}\ln(\mathit{NONKITSF}) + \beta_{36}\ln(\mathit{NONKITSF})^2$ $+ \beta_{37}\ln(\mathit{KITSF}) + \beta_{38}\ln(\mathit{KITSF})^2 + \beta_{39}\ln(\mathit{LOTSF}) + \beta_{40}\ln(\mathit{LOTSF})^2 + \beta_{41}\ln(\mathit{AGE})$ $+ \beta_{42}\ln(\mathit{AGE})^2$.
4. $\beta_{33}\mathit{BATHS} + \beta_{34}\ln(\mathit{NONKITSF}) + \beta_{35}\ln(\mathit{KITSF}) + \beta_{36}\ln(\mathit{LOTSF}) + \beta_{37}\ln(\mathit{AGE})$.
5. $\beta_{33}\mathit{BATHS} + \beta_{34\text{-}35}s(\mathit{NONKITSF}) + \beta_{36\text{-}37}s(\mathit{KITSF}) + \beta_{38\text{-}39}s(\mathit{LOTSF}) + \beta_{40\text{-}41}s(\mathit{AGE})$.
6. $\beta_{33}\mathit{BATHS} + \beta_{34\text{-}35}s(\ln(\mathit{NONKITSF})) + \beta_{36\text{-}37}s(\ln(\mathit{KITSF})) + \beta_{38\text{-}39}s(\ln(\mathit{LOTSF})) +$ $\beta_{40\text{-}41}s(\ln(\mathit{AGE}))$.

Data

The sample data came from the Memphis MLS's *Multiple Listing Book* (Memphis Board of Realtors, January 1987). The actual transactions price came from the cumulative index of sold properties. Characteristics data on each of the selected properties came either from this index or from the original listing description. The sample contains observations on 442 single-family dwellings sold within the previous six-month period with complete information on each variable. Stratified random sampling, whereby the proportion of properties in the sample from the twenty-four different city areas matched the population proportion in these areas, was used to insure a truly representative sample of the population of sold properties. As a result, the sample means of both the dependent and independent variables closely match their population counterparts.

Depictions of Nonlinearities

This subsection examines the GAM and polynomial regression model estimated functions $\tilde{f}_i(X_i)$ for various values of X_i. For the linear model $\tilde{f}_i(X_i)=X_i\beta_i$, hence any departures from this provide evidence of nonlinearities.

As the semilogarithmic specifications seems the most common in real estate, I began with it. Essentially, Model 1, the simple semilogarithmic model, represents this type. Using this model but allowing the nondichotomous variables to act as an arguments to a nonparametrically estimated functions gave rise to Model 5. The GAM estimated transformations (and their confidence regions) of the selected independent variables *BATHS*, *LOTSF*, *AGE* and *KITSF* for this model appear in Exhibit 1.[7] I allocated two degrees-of-freedom for each of these variables as I did for the polynomial regressions. This set of graphs reveals the apparent need for some type of transformations for the *BATHS*, *LOTSF* and *AGE* variables. Note, the estimated transformation for *LOTSF* actually turns down after a 50,000 square feet. However, the confidence region for the graph still admits of a monotonic transformation.

I used a logarithmic transformation of the *AGE*, *LOTSF* and *KITSF* variables as well as the *NONKITSF* variable (not shown due to lack of space). Recall, the coefficients in a double logarithmic specification estimate elasticities. Hence, GAM allows estimation of variable elasticities. Subsequent estimation of a GAM model involving the transformed variables (Model 6) produced the estimated transformations in Exhibit 2. The estimated transformations have become much more linear than those in Exhibit 2.[8] The new transformed independent variables gave rise to Model 4, the simple double logarithmic model.

As a check upon the GAM, I estimated the equivalent polynomial models (Model 2, the polynomial semilogarithmic specification and Model 3, the polynomial double logarithmic specification) using quadratic polynomials (two degrees-of-freedom). The polynomial semilogarithmic specification (Model 2) estimated transformations of the original variables and associated confidence intervals appear in Exhibit 3. The polynomial semilogarithmic specification provides estimated nonmonotonic transformations of *BATHS*, *LOTSF* and *AGE*. This model also estimates nonmonotonic confidence regions for the latter two variables.

Using ln(*AGE*), ln(*LOTSF*), ln(*KITSF*) and ln(*NONKITSF*) and their squares yielded Model 3. Subsequent estimation of the polynomial double log specification produced the estimated transformations in Exhibit 4. The estimated transformations have become more linear than those in Exhibit 3. However, the polynomial double logarithmic specification still estimates a non-monotonic transformation of *BATHS* and ln(*AGE*).[9] However, the confidence regions for both admit of monotonic transformations.

Generally, polynomial models approximate functions well within the factor space. However, attempts to extrapolate outside of this, especially for the untransformed

Exhibit 1
Spline GAM Fits for Semi-Log Model

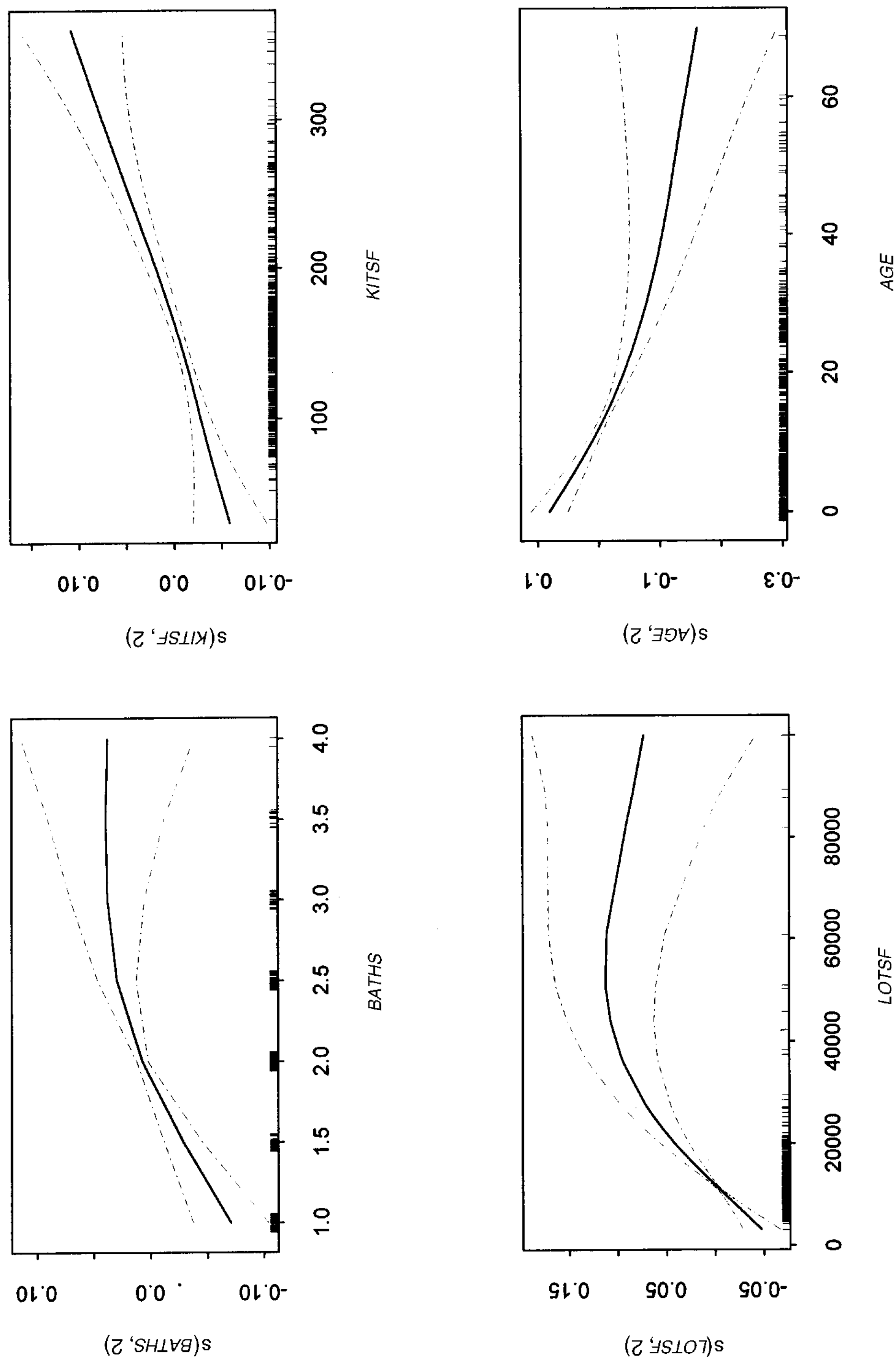


Exhibit 2
Spline GAM Fits for Double Log Model

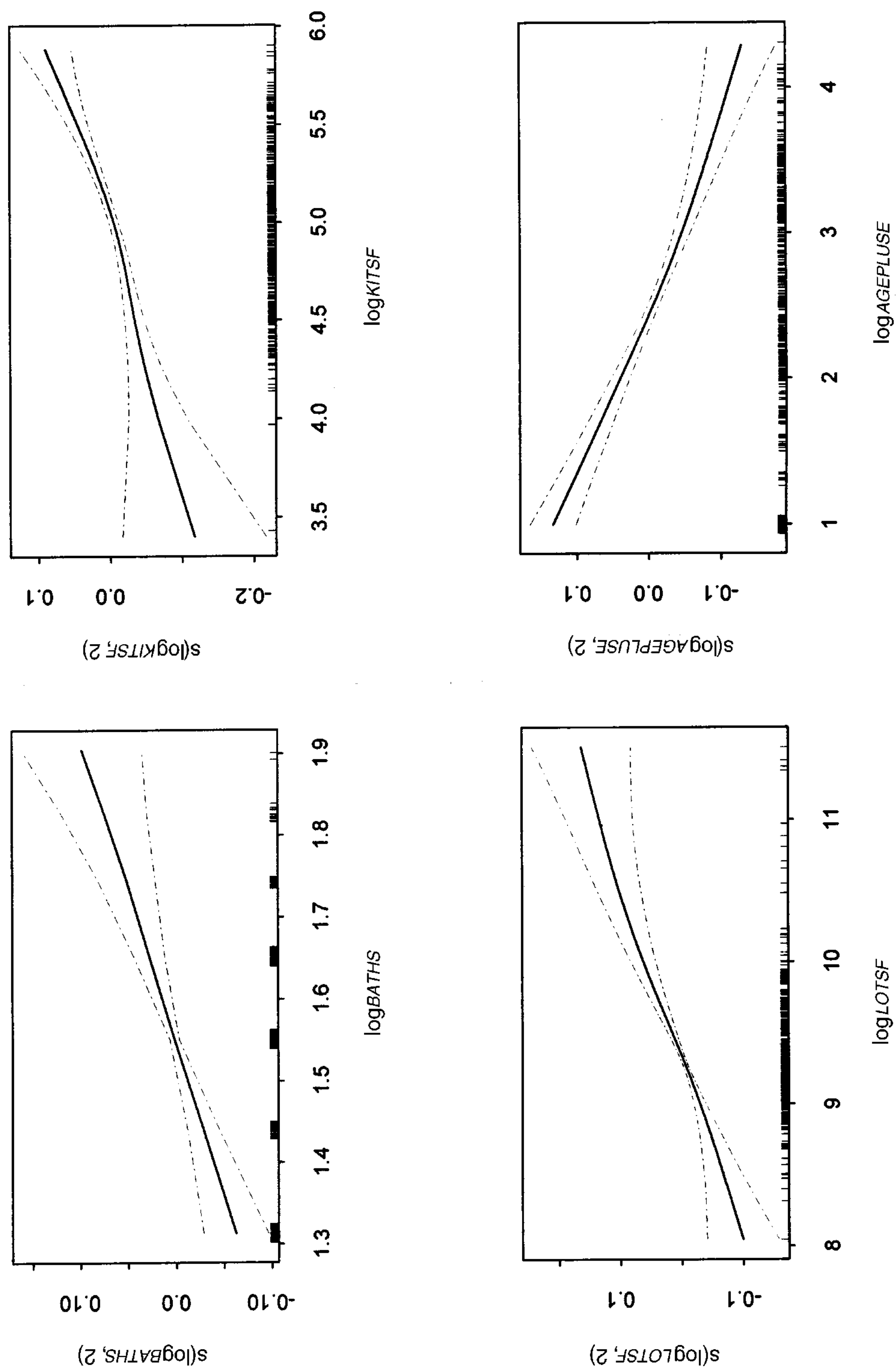


Exhibit 3
Polynomial GAM Fits for Semi-Log Model

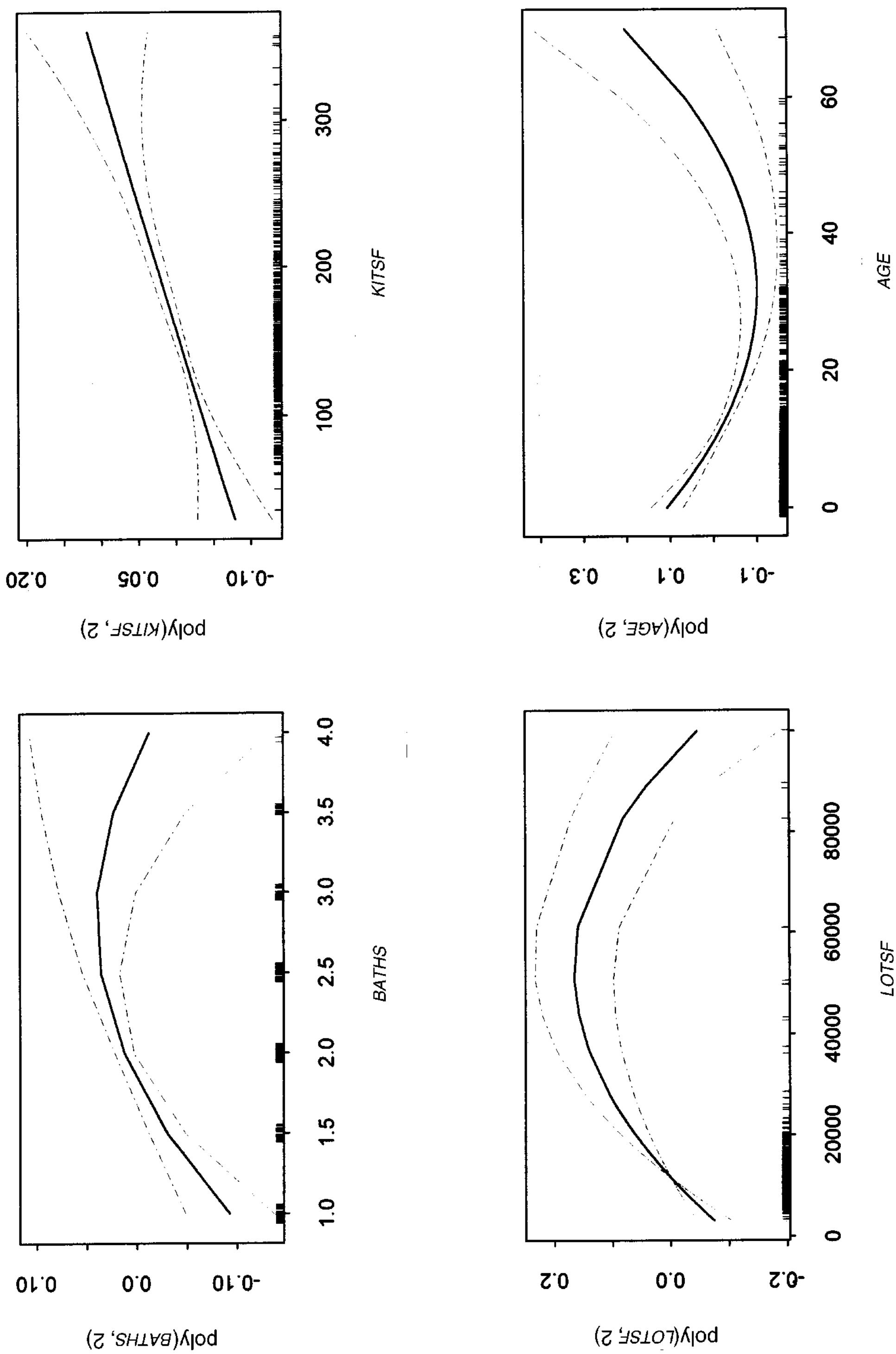
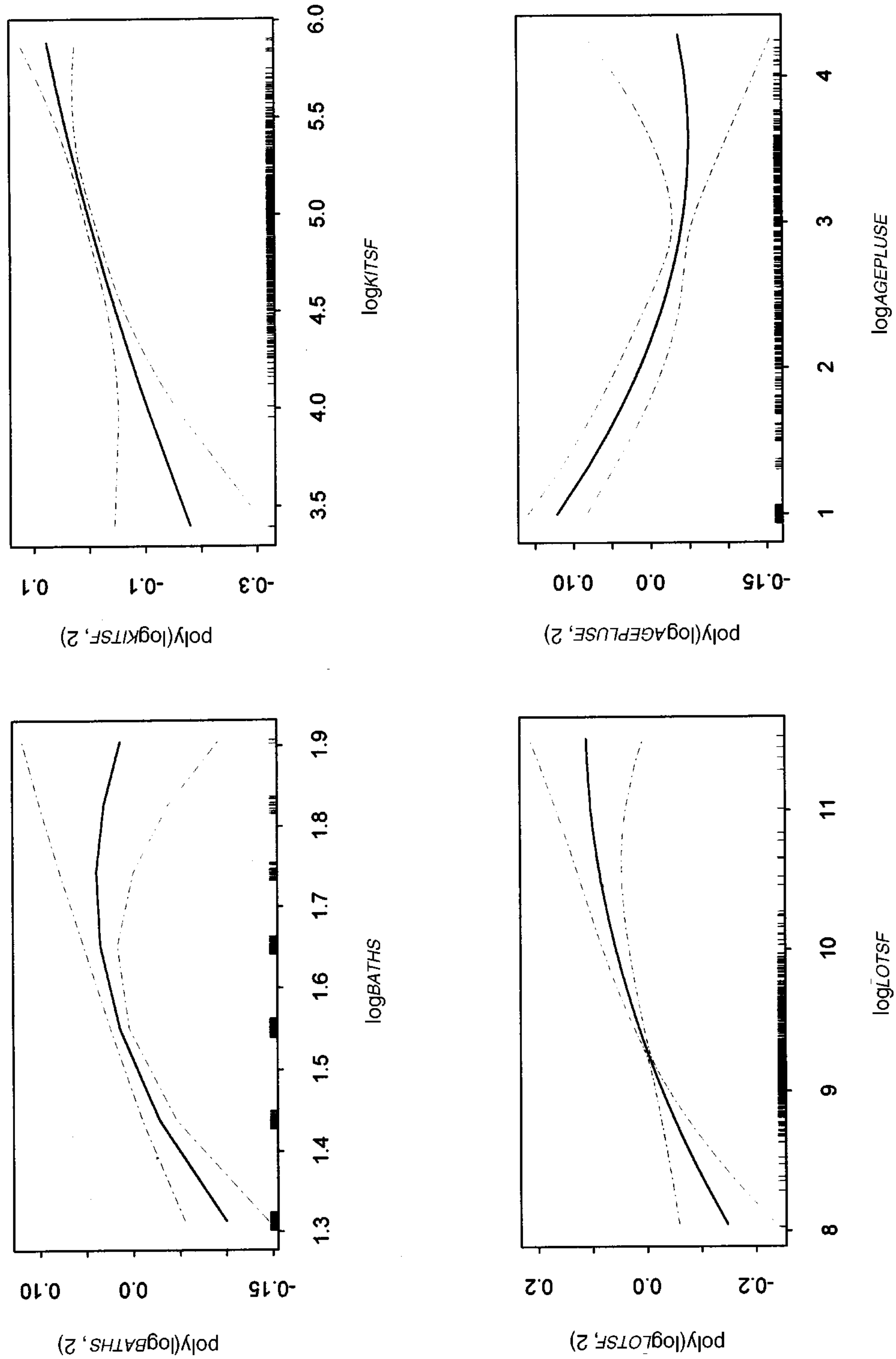


Exhibit 4
Polynomial GAM Fits for Double Log Model



variables, would likely produce poor results. For example, in Model 2 having over seventy years of *AGE* actually adds value to the house. In Model 3 having six or more *BATHS* would reduce the value of the house. On the other hand, the logarithmic transformation of *AGE* did help in Model 3. It would take over 1000 years of *AGE* before this would add to the value of the house.

Using GAM, as opposed to the traditional polynomial specification, resulted in transformations more in accord with prior information. Specifically, one would expect positive monotonic transformations of characteristics representing goods and negative monotonic transformations of characteristics such as *AGE*. The GAM transformations satisfied these priors while the polynomial specification often suggested nonmonotonic transformations.[10]

Finally, the GAM results agree with those of Anglin and Gencay (1993) who found the hedonic pricing function concave in bedrooms and in lot size using a combination of linear (seven variables) and nonparametric kernel estimation (six variables).

Global Sample Estimates

Exhibits 5-10 contain the estimates for the respective models using the global sample of 442 observations. Each of the models produced estimates with the expected signs. Only two of the nonarea variables, *ALUM* (for aluminum siding) and *POOL*, had estimates not significant at the 5% level or better.

Exhibit 11 presents the estimates for common characteristics across the six models. The last two columns provide the range of estimates across models and the highest estimated standard error across the six regressions. As an informal way of identifying model differences, I have shaded cells associated with variables where the range of estimates exceeded the maximum standard error by a factor of two or more. Eight of the thirty-one common nonintercept variables changed by this magnitude or more.[11]

Model 1, the original semilogarithmic model, and Model 4, the simple double logarithmic model, yielded the most extreme estimates of the six models. For example, Model 1 yielded the maximum estimate in five and the minimum estimate in one of the eight nonintercept variables where the estimate ranges exceeded twice the maximum standard error. Model 4 yielded the minimum estimate in five of the eight nonintercept variables where the estimate ranges exceeded twice the maximum standard error.

Can we state anything concerning the realism of these estimates given our prior information? Model 1 produced an estimate of fireplace value of $5,439 when transformed into price space.[12] Also, Model 1 produced an estimate of central air-conditioning value of $10,134 over that of window unit air-conditioning.[13] This seems somewhat unrealistic. In contrast, Model 4, the simple double logarithmic model yields a fireplace value of $2462 and a central air-conditioning value of $7198 over that of window unit air-conditioning. Hence, Model 1 exceeds the cost-bound priors reported by Pace and Gilley (1993) for both fireplace values and central air-

Exhibit 5
Original Semi-Logarithmic Specification

Variable	Estimate	Std. Err.	*t*-ratio	Pr(>|t|)
Intercept	10.336	0.043	239.2	0.000
AreaID 1	0.113	0.030	3.8	0.000
AreaID 2	0.120	0.016	7.7	0.000
AreaID 3	−0.051	0.012	−4.2	0.000
AreaID 4	0.035	0.008	4.3	0.000
AreaID 5	−0.020	0.006	−3.4	0.001
AreaID 6	0.006	0.005	1.4	0.175
AreaID 7	0.001	0.006	0.2	0.835
AreaID 8	0.013	0.005	2.5	0.014
AreaID 9	0.005	0.004	1.2	0.249
AreaID 10	0.007	0.004	1.6	0.112
AreaID 11	−0.004	0.003	−1.4	0.157
AreaID 12	−0.003	0.002	−1.1	0.257
AreaID 13	−0.009	0.003	−3.4	0.001
AreaID 14	−0.003	0.002	−1.6	0.104
AreaID 15	−0.005	0.002	−2.5	0.012
AreaID 16	−0.007	0.002	−3.2	0.001
AreaID 17	−0.010	0.002	−6.1	0.000
AreaID 18	−0.015	0.002	−7.4	0.000
AreaID 19	−0.014	0.002	−7.8	0.000
AreaID 20	−0.002	0.001	−1.9	0.059
AreaID 21	0.003	0.001	2.1	0.040
AreaID 22	0.000	0.001	0.3	0.750
AreaID 23	−0.006	0.002	−3.3	0.001
CarPorts	0.023	0.009	2.4	0.015
Garage	0.055	0.009	5.9	0.000
CenAC	0.129	0.023	5.7	0.000
NonAC	−0.205	0.050	−4.1	0.000
Fire	0.067	0.017	4.0	0.000
Pool	0.019	0.028	0.7	0.504
Brick	0.058	0.015	3.9	0.000
Alum	0.022	0.040	0.6	0.582
Baths	0.070	0.019	3.6	0.000
NonKitSF	0.000	0.000	15.9	0.000
KitSF	0.001	0.000	5.7	0.000
LotSF	0.000	0.000	2.4	0.016
Age	−0.003	0.001	−2.8	0.006

Note: There were 442 observations; df=405; R^2=.928; and F=145.6.

conditioning ($3700 and $6582, respectively). Model 4 exceeds the cost-bound prior for central air-conditioning, but by a much lesser margin. The other models yield estimates for these characteristics between those of Models 1 and 4.

Examining Exhibits 9 and 11 directly, note the close agreement in coefficients between Model 4, the simple double logarithmic model, and Model 6, the GAM double logarithmic model.[14] Finally, the standard errors on the common coefficients differed little across models.

Exhibit 6
Polynomial Semi-Logarithmic Specification

Variable	Estimate	Std. Err.	*t*-ratio	Pr(>|t|)
Intercept	11.072	0.025	439.7	0.000
AreaID 1	0.025	0.034	0.7	0.460
AreaID 2	0.120	0.015	8.2	0.000
AreaID 3	−0.032	0.012	−2.7	0.007
AreaID 4	0.046	0.008	5.6	0.000
AreaID 5	−0.011	0.006	−1.9	0.064
AreaID 6	0.013	0.005	2.8	0.005
AreaID 7	0.001	0.005	0.1	0.893
AreaID 8	0.012	0.005	2.5	0.012
AreaID 9	0.001	0.004	0.3	0.784
AreaID 10	0.003	0.004	0.8	0.419
AreaID 11	−0.006	0.003	−2.2	0.026
AreaID 12	−0.004	0.002	−1.6	0.101
AreaID 13	−0.005	0.003	−1.9	0.060
AreaID 14	0.000	0.002	−0.1	0.957
AreaID 15	−0.005	0.002	−2.5	0.012
AreaID 16	−0.006	0.002	−2.7	0.008
AreaID 17	−0.008	0.002	−5.1	0.000
AreaID 18	−0.013	0.002	−6.4	0.000
AreaID 19	−0.012	0.002	−7.1	0.000
AreaID 20	−0.001	0.001	−1.0	0.340
AreaID 21	0.002	0.001	1.9	0.057
AreaID 22	0.001	0.001	0.4	0.708
AreaID 23	−0.011	0.002	−5.6	0.000
CarPorts	0.026	0.009	3.0	0.003
Garage	0.045	0.009	5.2	0.000
CenAC	0.098	0.022	4.3	0.000
NonAC	−0.180	0.047	−3.9	0.000
Fire	0.043	0.016	−2.6	0.010
Pool	0.028	0.026	1.1	0.277
Brick	0.052	0.014	3.7	0.000
Alum	0.023	0.037	0.6	0.532
poly(*Baths*,2)1	0.712	0.217	3.3	0.001
poly(*Baths*,2)2	−0.456	0.170	−2.7	0.008
poly(*NonKitSF*,2)1	4.213	0.251	16.8	0.000
poly(*NonKitSF*,2)2	−0.248	0.155	−1.6	0.111
poly(*KitSF*,2)1	0.744	0.124	6.0	0.000
poly(*KitSF*,2)2	−0.010	0.120	−0.1	0.935
poly(*LotSF*,2)1	0.500	0.132	3.8	0.000
poly(*LotSF*,2)2	−0.589	0.138	−4.3	0.000
poly(*Age*,2)1	−1.065	0.337	−3.2	0.002
poly(*Age*,2)2	1.071	0.198	5.4	0.000

Note: There were 442 observations: df=400; R^2=.939; and F=150.

Cross-Validation of Prediction Errors

To gain insight into the six models, I conducted a predictive cross-validation experiment. Specifically, for each of the 500 iterations I divided the sample into 221 insample and 221 exsample observations. I estimated the six models on the 221

Exhibit 7
Polynomial Double Logarithmic Specification

Variable	Estimate	Std. Err.	*t*-ratio	Pr(>|t|)
Intercept	11.070	0.026	432.7	0.000
AreaID 1	0.072	0.028	2.6	0.011
AreaID 2	0.115	0.015	7.9	0.000
AreaID 3	−0.038	0.011	−3.4	0.001
AreaID 4	0.039	0.008	5.1	0.000
AreaID 5	−0.018	0.006	−3.3	0.001
AreaID 6	0.010	0.004	2.3	0.021
AreaID 7	0.006	0.006	1.0	0.296
AreaID 8	0.016	0.005	3.3	0.001
AreaID 9	0.004	0.004	0.9	0.354
AreaID 10	0.004	0.004	1.0	0.315
AreaID 11	−0.004	0.003	−1.3	0.186
AreaID 12	−0.002	0.002	−0.9	0.349
AreaID 13	−0.006	0.003	−2.3	0.023
AreaID 14	−0.001	0.002	−0.5	0.585
AreaID 15	−0.002	0.002	−1.2	0.230
AreaID 16	−0.007	0.002	−3.1	0.002
AreaID 17	−0.009	0.002	−5.8	0.000
AreaID 18	−0.013	0.002	−6.7	0.000
AreaID 19	−0.011	0.002	−6.3	0.000
AreaID 20	−0.001	0.001	−0.6	0.538
AreaID 21	0.003	0.001	2.8	0.005
AreaID 22	0.002	0.001	1.4	0.176
AreaID 23	−0.009	0.002	−4.9	0.000
CarPorts	0.018	0.009	2.1	0.036
Garage	0.041	0.009	4.8	0.000
CenAC	0.093	0.022	4.2	0.000
NonAC	−0.169	0.048	−3.5	0.000
Fire	0.045	0.017	2.7	0.008
Pool	0.025	0.026	1.0	0.328
Brick	0.049	0.014	3.5	0.001
Alum	0.007	0.038	0.2	0.855
poly(*Baths*,2)1	0.764	0.217	3.5	0.000
poly(*Baths*,2)2	−0.487	0.165	−3.0	0.003
poly(ln(*NonKitSF*),2)1	4.113	0.251	16.4	0.000
poly(ln(*NonKitSF*),2)2	0.963	0.174	5.5	0.000
poly(ln(*KitSF*),2)1	0.757	0.122	6.2	0.000
poly(ln(*KitSF*),2)2	−0.056	0.122	−0.5	0.644
poly(ln(*LotSF*),2)1	0.700	0.151	4.7	0.000
poly(ln(*LotSF*),2)2	−0.167	0.124	−1.3	0.180
poly(ln(*Age*),2)1	−1.127	0.308	−3.7	0.000
poly(ln(*Age*),2)2	0.425	0.211	2.0	0.045

Note: There were 442 observations; df=400; R^2=.940; and F=151.5.

insample observations and recorded their prediction errors on the 221 exsample observations.[15] I examined both root mean squared error (RMSE) and the median absolute error.[16] Exhibit 12 provides the mean and standard deviation of the errors (both RMSE and median absolute errors) as well as relative mean and standard

Exhibit 8
Double Logarithmic Specification

Variable	Estimate	Std. Err.	*t*-ratio	Pr(>|t|)
Intercept	6.062	0.251	24.1	0.000
AreaID 1	0.075	0.026	2.9	0.004
AreaID 2	0.119	0.015	7.9	0.000
AreaID 3	−0.044	0.011	−4.0	0.000
AreaID 4	0.030	0.007	4.1	0.000
AreaID 5	−0.016	0.006	−2.8	0.006
AreaID 6	0.007	0.004	1.5	0.126
AreaID 7	−0.001	0.004	−0.2	0.830
AreaID 8	0.014	0.004	3.4	0.001
AreaID 9	−0.003	0.003	−0.8	0.442
AreaID 10	−0.001	0.004	−0.3	0.797
AreaID 11	−0.007	0.003	−2.9	0.004
AreaID 12	−0.006	0.002	−2.9	0.004
AreaID 13	−0.007	0.003	−2.5	0.014
AreaID 14	−0.002	0.002	−1.1	0.255
AreaID 15	−0.005	0.002	−3.2	0.002
AreaID 16	−0.007	0.002	−3.1	0.002
AreaID 17	−0.009	0.002	−5.9	0.000
AreaID 18	−0.012	0.002	−6.0	0.000
AreaID 19	−0.011	0.002	−6.6	0.000
AreaID 20	−0.002	0.001	−1.8	0.080
AreaID 21	0.002	0.001	1.6	0.108
AreaID 22	0.001	0.001	0.5	0.629
AreaID 23	−0.010	0.002	−5.6	0.000
CarPorts	0.020	0.009	2.3	0.022
Garage	0.044	0.009	4.9	0.000
CenAC	0.091	0.022	4.1	0.000
NonAC	−0.136	0.048	−2.8	0.005
Fire	0.030	0.017	1.8	0.076
Pool	0.031	0.027	1.2	0.246
Brick	0.045	0.014	3.1	0.002
Alum	0.017	0.038	0.4	0.656
Baths	0.074	0.018	4.0	0.000
ln(*NonKitSF*)	0.529	0.034	15.8	0.000
ln(*KitSF*)	0.098	0.016	6.2	0.000
ln(*LotSF*)	0.078	0.017	4.5	0.000
ln(*Age*)	−0.086	0.013	−6.6	0.000

Note: There were 442 observations; df=405; R^2=.933; and F=157.8.

deviation (scaled by the mean and standard deviation of Model 6). In addition, Exhibit 12 provides the proportion of iterations where each model improved over the others for both types of error.

For both types of error, Model 1 performs the worst and Model 6 performs the best. I found the latter result rather surprising as the difference between the estimates on Model 4 and Model 6 were rather small given the extra degrees of freedom used by Model 6. In terms of median absolute error, Model 1 performs 15.1% worse in a

Exhibit 9
GAM Semi-Logarithmic Specification

Variable	Estimate	Std. Err.	*t*-ratio	Pr(>|t|)
Intercept	10.359	0.041	255.1	0.000
AreaID 1	0.085	0.028	3.0	0.003
AreaID 2	0.117	0.015	8.0	0.000
AreaID 3	−0.044	0.011	−3.9	0.000
AreaID 4	0.037	0.008	4.9	0.000
AreaID 5	−0.017	0.006	−3.1	0.002
AreaID 6	0.009	0.004	2.0	0.047
AreaID 7	0.000	0.005	0.1	0.956
AreaID 8	0.013	0.005	2.6	0.008
AreaID 9	0.003	0.004	0.7	0.454
AreaID 10	0.005	0.004	1.3	0.203
AreaID 11	−0.006	0.003	−2.0	0.051
AreaID 12	−0.003	0.002	−1.3	0.194
AreaID 13	−0.007	0.003	−2.8	0.005
AreaID 14	−0.002	0.002	−1.0	0.343
AreaID 15	−0.005	0.002	−2.5	0.013
AreaID 16	−0.007	0.002	−3.3	0.001
AreaID 17	−0.009	0.002	−6.0	0.000
AreaID 18	−0.014	0.002	−7.2	0.000
AreaID 19	−0.013	0.002	−7.6	0.000
AreaID 20	−0.002	0.001	−1.5	0.139
AreaID 21	0.002	0.001	2.1	0.039
AreaID 22	0.001	0.001	0.5	0.611
AreaID 23	−0.009	0.002	−5.0	0.000
CarPorts	0.023	0.009	2.6	0.009
Garage	0.049	0.009	5.7	0.000
CenAC	0.107	0.021	5.0	0.000
NonAC	−0.191	0.047	−4.1	0.000
Fire	0.052	0.016	3.2	0.001
Pool	0.021	0.026	0.8	0.419
Brick	0.056	0.014	4.0	0.000
Alum	0.022	0.038	0.6	0.567
s(*Baths*,2)	0.065	0.018	3.6	0.000
s(*NonKitSF*,2)	0.000	0.000	17.5	0.000
s(*KitSF*,2)	0.001	0.000	5.9	0.000
s(*LotSF*,2)	0.000	0.000	3.4	0.001
s(*Age*,2)	−0.004	0.001	−3.1	0.002

Note: There were 442 observations; df=400; R^2=.937; and F=166.

relative sense and 1.0% in an absolute sense. This means Model 1 has about $800 extra error than Model 6 when converting this into price space. As the proportions show, Model 6 yields less error than Model 1 in 98.0% to 96.4% of the trials. As expected, Model 6 outperforms the polynomial models.

Model 4, the simple double logarithmic model, outperformed every model except Model 6 in terms of median absolute error. Model 4 outperformed Model 2 in terms of mean RMSE and Model 3 in terms of improvement proportions. Hence, the results

Exhibit 10
GAM Double Logarithmic Specification

Variable	Estimate	Std Err.	*t*-ratio	Pr(>|t|)
Intercept	6.030	0.242	24.9	0.000
AreaID 1	0.074	0.025	3.0	0.003
AreaID 2	0.117	0.014	8.1	0.000
AreaID 3	−0.042	0.011	−4.0	0.000
AreaID 4	0.034	0.007	4.8	0.000
AreaID 5	−0.017	0.005	−3.0	0.003
AreaID 6	0.008	0.004	1.9	0.052
AreaID 7	0.002	0.004	0.4	0.670
AreaID 8	0.015	0.004	3.7	0.000
AreaID 9	0.000	0.003	<0.1	0.996
AreaID 10	0.001	0.004	0.3	0.742
AreaID 11	−0.006	0.002	−2.4	0.016
AreaID 12	−0.004	0.002	−2.1	0.036
AreaID 13	−0.006	0.003	−2.3	0.020
AreaID 14	−0.001	0.002	−0.7	0.508
AreaID 15	−0.004	0.002	−2.5	0.014
AreaID 16	−0.007	0.002	−3.2	0.002
AreaID 17	−0.009	0.002	−5.9	0.000
AreaID 18	−0.012	0.002	−6.2	0.000
AreaID 19	−0.011	0.002	−6.8	0.000
AreaID 20	−0.001	0.001	−1.2	0.240
AreaID 21	0.003	0.001	2.3	0.020
AreaID 22	0.001	0.001	1.0	0.321
AreaID 23	−0.010	0.002	−5.6	0.000
CarPorts	0.021	0.009	2.4	0.017
Garage	0.044	0.009	5.0	0.000
CenAC	0.095	0.021	4.5	0.000
NonAC	−0.143	0.046	−3.1	0.002
Fire	0.037	0.016	2.3	0.022
Pool	0.026	0.026	1.0	0.324
Brick	0.047	0.014	3.4	0.001
Alum	0.017	0.037	0.5	0.652
s(*Baths*,2)	0.070	0.018	4.0	0.000
s(ln(*NonKitSF*),2)	0.528	0.032	16.3	0.000
s(ln(*KitSF*),2)	0.098	0.015	6.4	0.000
s(ln(*LotSF*),2)	0.079	0.017	4.8	0.000
s(ln(*Age*),2)	−0.076	0.013	−6.0	0.000

Note: There were 442 observations; df=400; R^2=.939; and F=170.2.

for Model 4 were mixed using RMSE. However, due to the measured lack of normality of the residuals, I believe the median absolute error gives a clearer picture.

Conclusion

GAM allows for the estimation of the dependent variable as a series of general functions of the independent variables. Insofar as the dependent variable can be in logarithms or powers, GAM can model multiplicative specifications as well. GAM

Exhibit 11
Estimates of Common Variables Across Models

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Range	Max SE
Intercept	10.336	11.072	11.070	6.062	10.359	6.030	5.042	0.251
1	0.113	0.025	0.072	0.075	0.085	0.074	0.088	0.034
2	0.120	0.120	0.115	0.119	0.117	0.117	0.005	0.016
3	−0.051	−0.032	−0.038	−0.044	−0.044	−0.042	0.019	0.012
4	0.035	0.046	0.039	0.030	0.037	0.034	0.016	0.008
5	−0.020	−0.011	−0.018	−0.016	−0.017	−0.017	0.009	0.006
6	0.006	0.013	0.010	0.007	0.009	0.008	0.007	0.005
7	0.001	0.001	0.006	−0.001	0.000	0.002	0.007	0.006
8	0.013	0.012	0.016	0.014	0.013	0.015	0.004	0.005
9	0.005	0.001	0.004	−0.003	0.003	0.000	0.008	0.004
10	0.007	0.003	0.004	−0.001	0.005	0.001	0.008	0.004
11	−0.004	−0.006	−0.004	−0.007	−0.006	−0.006	0.003	0.003
12	−0.003	−0.004	−0.002	−0.006	−0.003	−0.004	0.004	0.002
13	−0.009	−0.005	−0.006	−0.007	−0.007	−0.006	0.004	0.003
14	−0.003	0.000	−0.001	−0.002	−0.002	−0.001	0.003	0.002
15	−0.005	−0.005	−0.002	−0.005	−0.005	−0.004	0.003	0.002
16	−0.007	−0.006	−0.007	−0.007	−0.007	−0.007	0.001	0.002
17	−0.010	−0.008	−0.009	−0.009	−0.009	−0.009	0.002	0.002
18	−0.015	−0.013	−0.013	−0.012	−0.014	−0.012	0.003	0.002
19	−0.014	−0.012	−0.011	−0.011	−0.013	−0.011	0.003	0.002
20	−0.002	−0.001	−0.001	−0.002	−0.002	−0.001	0.001	0.001
21	0.003	0.002	0.003	0.002	0.002	0.003	0.001	0.001
22	0.000	0.001	0.002	0.001	0.001	0.001	0.002	0.001
23	−0.006	−0.011	−0.009	−0.010	−0.009	−0.010	0.005	0.002
CarPorts	0.023	0.026	0.018	0.020	0.023	0.021	0.008	0.009
Garages	0.055	0.045	0.041	0.044	0.049	0.044	0.014	0.009
CenAC	0.129	0.098	0.093	0.091	0.107	0.095	0.038	0.023
NonAC	−0.205	−0.180	−0.169	−0.136	−0.191	−0.143	0.069	0.050
Fire	0.067	0.043	0.045	0.030	0.052	0.037	0.037	0.017
Pool	0.019	0.028	0.025	0.031	0.021	0.026	0.009	0.028
Brick	0.058	0.052	0.049	0.045	0.056	0.047	0.013	0.015
Alum	0.022	0.023	0.007	0.017	0.022	0.017	0.016	0.040

estimate each function of the independent variables nonparametrically through smoothing splines, locally weighted regression or the kernel method. Hence, GAM attempt to combine the interpretability of additive modeling with the flexibility of nonparametric estimation. Since GAM generally rely upon bivariate nonparametric estimation, these estimates converge with $\sqrt{n}$, the same rate as parametric estimators and above the multivariate nonparametric estimator rate, which declines with dimensionality. Moreover, the bivariate fits lend themselves to graphical presentation. Hence, the user can see the estimated transformation versus its argument. This knowledge can lead to variable transformations as an input to further estimation or prove useful in its own right. Finally, GAM are a generalization of GLMs. Consequently, one can easily apply GAM using other distributions such as Poisson, gamma and multinomial using different variance specifications. Hence, one can estimate count or duration data, survival data and probabilities with the same flexibility in functional form.

Exhibit 12
Cross-Validation of Prediction Errors[a]

RMSE	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Mean	0.135	0.132	0.129	0.130	0.128	0.127
σ	0.011	0.012	0.012	0.010	0.011	0.011
Relative Mean	1.062	1.037	1.013	1.021	1.011	1.000
Relative σ	1.028	1.165	1.084	0.965	1.014	1.000
Panel A: Proportions of Trials where Model i RMSE<Model j RMSE						
2<1	0.756					
3<1,2	0.906	0.644				
4<1,2,3	0.896	0.538	0.358			
5<1,2,3,4	0.992	0.650	0.504	0.668		
6<1,2,3,4,5	0.980	0.732	0.610	0.934	0.656	
Panel B: Median |e|						
Mean	0.074	0.066	0.068	0.066	0.069	0.064
σ	0.006	0.005	0.005	0.005	0.005	0.005
Relative Mean	1.151	1.026	1.048	1.016	1.071	1.000
Relative σ	1.109	1.005	1.032	1.012	1.032	1.000
Panel C: Proportions of Trials Where Model i Median|e|<Model j Median|e|						
2<1	0.9480					
3<1,2	0.9000	0.3720				
4<1,2,3	0.9700	0.5600	0.6660			
5<1,2,3,4	0.9320	0.2200	0.3580	0.2060		
6<1,2,3,4,5	0.9640	0.6580	0.7820	0.6480	0.8900	

[a]Based upon 500 iterations of resampling 221 insample and 221 exsample observations.

This article applied GAM to a sample of 442 houses with transactions data from the Memphis area. The GAM estimator suggested logarithmic transformations to the nondichotomous variables. In prediction cross-validation experiments the resulting double logarithmic model outperformed the original semiparametric model by a large margin. Specifically, the double logarithmic model improved on the semilogarithmic model anywhere from 89.6% to 97.0% of the trials. In turn, the GAM model, despite the five extra degrees-of-freedom used, improved over the simple double logarithmic model (and all others). Specifically, the GAM improved on the semilogarithmic model anywhere from 96.4% to 98.0% of the trials and improved over the double logarithmic model in 64.8% to 93.4% of the trials. The GAM displayed about $800 less error on average than the original semilogarithmic model.

In his discussion of the things statisticians can learn from neural net experiences, Tibshirani (1994) stated, ‘‘Models with very large numbers of parameters can be useful for prediction, especially for large data sets and problems exhibiting high signal-to-noise ratios.’’ GAM, along with many other semiparametric and nonparametric estimators, fall into this category. Moreover, we often have large data sets in real estate with high signal-to-noise (goodness-of-fit). Based upon this study and others, semiparametric and nonparametric estimators would appear to have great potential in real estate empirical research.

Notes

[1]Debate continues concerning the guidance theory provides concerning the directions or signs of first and second derivatives of prices with respect to the bundled characteristics. Traditional hedonic pricing theory (Rosen, 1974) does not offer much guidance. Coulson (1989) presented conditions which could lead to a linear form. Pace and Gilley (1990) provided some arbitrage arguments based upon home remodeling costs which bounded the derivative of price with respect to the characteristics. Colwell (1991, 1993) presented in detail the theory behind such arbitrage, which could result in nonmonotonic hedonic price functions in special cases. In addition to the theory, numerous hedonic pricing and computer aided mass assessment applications most often support hedonic pricing functions linear or concave in characteristics representing goods.

[2]The so-called "curse of dimensionality." Bivariate nonparametric estimators can often match the rate on parametric estimators ($\sqrt{n}$), but multivariate nonparametric estimators converge at much slower rates.

[3]Projection pursuit predates and subsumes neural nets. However, neural nets often use so many parameters that they often resemble nonparametric estimators more than semiparametric ones. See Cheng and Titterington (1994) for more on the statistical interpretation of neural nets (many of the eight discussants' comments are quite interesting as well). See Do and Grudnitski (1993) for a real estate application.

[4]He used parametric estimation to form the weights used in the nonparametric smoothing. His results clearly show the advantages of nonparametric estimation. Unfortunately, he did not link his method to the literature on nonparametric estimation.

[5]The S-PLUS program gives a choice of locally weighted regression smoothing, smoothing splines, two forms of regression splines and polynomials for estimating GAM. Of these, polynomials and regressions splines have parametric forms which methods such as least squares can fit directly. For nonparametric estimation, S-PLUS also implements the supersmoother and kernel methods. Similarly, GAUSSX supports kernel regression. Also, by combining the kernel regression and maximum likelihood routines in GAUSSX, one could estimate GAM. However, the kernel used in GAUSSX is not guaranteed to converge in GAM estimation, unlike smoothing splines or locally weighted regression.

[6]Since new houses appeared in the sample, *AGE* equals 0 for some observations. This necessitated the addition of a constant. Adding a constant changes the curvature of the function with respect to the variable. Neural nets call this "bias." I chose *e* as a constant based upon a graph of ln(*Price*) versus *AGE*.

[7]One can identify the spline functions up to a constant (the polynomial functions are completely identified). Also, due to scaling of the functions, the actual numbers on the *y* axis are difficult to interpret. However, this does not affect the pattern of linearity or nonlinearity of the function.

[8]I also estimated a GAM using ln(*BATHS*). Both the *BATHS* and ln(*BATHS*) variables produced approximately linear graphs. I retained the *BATHS* variable to act as a nondichotomous variable common to all models.

[9]Using a neural network model, Do and Grudnitski (1993) found a nonmonotonic relation between sales price and age. In estimating GAM using sales price, I noted a similar phenomenon (as also present with the polynomial regressions). However, the introduction of an interaction between total area and lot area eliminated this and led to greater linearity among all the variables. In fact, the estimated functions were more linear when using *Price* than ln(*Price*).

Incidentally, the neural net estimator predictions did not improve on those of OLS for a variant of this data.

[10]Restricting the coefficients in the polynomial regression to produce monotonic transformations over the factor space might cure some of its problems. One could accomplish this through inequality constrained least squares. See Pace and Gilley (1993) for an application of this to semilogarithmic models. Note, new software exists to do this. Both GAUSS and Matlab have quadratic programming routines while S-PLUS includes a non-negative least-squares routine. SHAZAM performs Bayesian inequality estimation. This would also prove quite simple to implement in GAUSS or GAUSSX with the RNDTN function.

[11]The intercept variable possessed the greatest divergence of estimates relative to their standard errors. However, the change in the definition of variables and the addition of other terms would naturally cause this to vary greatly across regressions.

[12]One can estimate the value of a coefficient in price space given some model fitted in some function of price space. See Pace and Gilley (1993) for an application and references. Note, the adjustment varies depending upon whether X_i is continuous or discrete.

[13]The intercept would capture the effect of window unit air-conditioning since two dichotomous variables measure no air-conditioning and central air-conditioning.

[14]It requires adjustment to compare the effects of the other variables across models. Spot calculations show the different models tend to produce similar estimates at the mean of X. Note, the polynomial regression routine in S-PLUS uses orthogonal polynomials and rescales X_i to unit column length to prevent numerical problems. Hence, the scaling makes the numbers appear large.

[15]The 500 iterations took 11.5 hours using S-PLUS on a 486DX250 computer.

[16]While all of these estimators minimize sum of squared error or equivalently RMSE, the median absolute error may measure more precisely, in some sense, the performance of these estimators. Squared error can have an asymmetric, long-tailed distribution. While sums of such random variates approach normality, in this case, the resulting sums had not yet reached normality. The mean does not possess much efficiency for such distributions. Hence, the mean may not measure central tendencies as well as the median. Note, the median of squared error is the same as the squared median absolute error (for series with odd-numbered lengths). Hence, the median absolute error may provide information relevant to squared errors.

References

Anglin, P. M. and R. Gencay, *Semiparametric Estimation of a Hedonic Price Function*, Windsor, CA: University of Windsor, 1993.

Breiman, L., The β Method for Estimating Multivariate Functions from Noisy Data, *Technometrics*, 1991, 33, 125–60.

Breiman, L. and J. Friedman, Estimating Optimal Transformations for Multiple Regression and Correlation, *Journal of the American Statistical Association*, 1985, 80, 580–19.

Breiman, L., J. Friedman, R. Ohlsen and C. J. Stone, *Classification and Regression Trees*, Belmont, CA: Wadsworth Press, 1984.

Buja, A., T. Hastie and R. Tibshirani, Linear Smoothers and Additive Models, *Annals of Statistics*, 1989, 17, 453–555

Cheng, B. and D. M. Titterington, Neural Networks: A Review From a Statistical Perspective (with discussion), *Statistical Science*, 1994, 9, 2–54.

Cleveland, W. S. and S. J. Devlin, Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting, *Journal of the American Statistical Association*, 1988, 83, 596–10.

Colwell, P. F., Functional Obsolescence and an Extension of Hedonic Theory, *Journal of Real Estate Finance and Economics*, 1991, 4, 49–58.

——, Semiparametric Estimates of the Marginal Price of Floorspace: Comment, *Journal of Real Estate Finance and Economics*, 1993, 7, 73–74.

Coulson, N. E., The Empirical Content of the Linearity-As-Repackaging Hypothesis, *Journal of Urban Economics*, 1989, 25, 295–09.

——, Semiparametric Estimates of the Marginal Price of Floorspace, *Journal of Real Estate Finance and Economics*, 1992, 5, 73–84.

Do, A. Q. and G. Grudnitski, A Neural Network Analysis of the Effect of AGE on Housing Values, *Journal of Real Estate Research*, 1993, 8, 253–64.

Duan, N. and K. C. Li, Slicing Regression: A Link-free Regression Method, *Annals of Statistics*, 1991, 19, 505–30.

Friedman, J., Multivariate Adaptive Regression Splines, *Annals of Statistics*, 1991, 19, 1–141.

Friedman, J. and W. Stuetzle, Projection Pursuit Regression, *Journal of the American Statistical Association*, 1981, 76, 817–23.

Hardle, W. and T. M. Stoker, Investigating Smooth Multiple Regression by the Method of Average Derivatives, *Journal of the American Statistical Association*, 1989, 84, 986–95.

Harrison, A., GAIM 1.1: A Review, *Journal of Applied Econometrics*, 1993, 8, 101–07.

Hastie, T. J. and R. Tibshirani, *Generalized Additive Models*, London: Chapman and Hall, 1990.

Hilbe, J. M., Generalized Additive Models Software, *The American Statistician*, 1993, 47, 59–64.

Isakson, H. R, The Nearest Neighbors Appraisal Technique: An Alternative to the Adjustment Grid Methods, *Journal of the American Real Estate and Urban Economics Association*, 1986, 14, 274–86.

McCullagh, P. and J. A. Nelder, *Generalized Linear Models*, New York: Chapman and Hall, 1989.

Meese, R. and N. Wallace, Nonparametric Estimation of Dynamic Hedonic Price Models and the Construction of Residential Housing Price Indices, *Journal of the American Real Estate and Urban Economics Association*, 1991, 19, 308–22.

Memphis Board of Realtors, *Multiple Listing Book 2*, Memphis, 1987.

Pace, R. K., Nonparametric Methods with Applications to Hedonic Models, *Journal of Real Estate Finance and Economics*, 1993, 7, 185–04.

——, Parametric, Semiparametric, and Nonparametric Estimation of Characteristic Values Within Mass Assessment and Hedonic Pricing Models, *Journal of Real Estate Finance and Economics*, 1995, 11, 195–17.

Pace, R. K. and O. W. Gilley, Appraisal Across Jurisdictions Using Bayesian Estimation With Bootstrapped Priors for Secondary Mortgage Market Applications, *Property Tax Journal*, 1989, 8, 27–42.

——, Estimation Employing A Priori Information With Mass Appraisal and Hedonic Pricing Models, *Journal of Real Estate Finance and Economics*, 1990, 3, 55–72.

——, Improving Prediction and Assessing Specification Quality in Non-Linear Statistical Valuation Models, *Journal of Business and Economics Statistics*, 1993, 11, 301–10.

Powell, J. L., J. H. Stock and T. M. Stoker, Semiparametric Estimation of Index Coefficients, *Econometrica*, 1989, 57, 1403–30.

Rosen, S., Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, *Journal of Political Economy*, 1974, 82, 34–55.

Sunderman, M., J. Birch, R. Cannaday and T. Hamilton, Testing for Vertical Inequity in Property Tax Systems, *Journal of Real Estate Research*, 1990, 5, 319–45.

Tibshirani, R., Estimating Optimal Transformations for Regression via Additivity and Variance Stabilization, *Journal of the American Statistical Association*, 1988, 83, 394–05.

——, Neural Networks: A Review from a Statistical Perspective: Comment, *Statistical Science*, 1994, 9, 48–49.

My thanks to Mustafa Chowdhury who discussed an earlier version of this manuscript at the 1995 ASSA. I also gratefully acknowledge the generous research support provided from the University of Alaska.