

**Input Separability in the Canadian Cable  
Television Industry: An Application of  
Generalized Additive Models**

by

Morteza Haghiri, Stephen M. Law,  
and James F. Nolan

Working Paper Series 2004-03



**DEPARTMENT OF ECONOMICS  
THE UNIVERSITY OF NEW BRUNSWICK  
FREDERICTON, CANADA**

# Input Separability in the Canadian Cable Television Industry: An Application of Generalized Additive Models

MORTEZA HAGHIRI, STEPHEN M. LAW <sup>†</sup>, and JAMES F. NOLAN

---

Morteza Haghiri, Assistant Professor, Department of Economics  
Mount Allison University, 144 Main Street, Sackville, New Brunswick, Canada, E4L 1A7  
Phone: +001 (506) 364-2334, Fax: +001 (506) 364-2625, E-mail: [mhaghiri@mta.ca](mailto:mhaghiri@mta.ca)

Stephen M. Law, Associate Professor, Department of Economics  
Mount Allison University, 144 Main Street, Sackville, New Brunswick, Canada, E4L 1A7  
Phone: +001 (506) 364-2355, Fax: +001 (506) 364-2625, E-mail: [slaw@mta.ca](mailto:slaw@mta.ca)  
and Adjunct Professor, Department of Economics  
University of New Brunswick, P.O. Box 4400, Fredericton, New Brunswick, Canada E3B 5A3  
Phone (Department): +001 (506) 453-4828, Fax: +001 (506) 453-4514, E-mail: [slawl@unb.ca](mailto:slawl@unb.ca)

James F. Nolan, Associate Professor, Department of Agricultural Economics  
University of Saskatchewan, 51 Campus Drive, Saskatoon, Saskatchewan, Canada, S7N 5A8  
Phone: +001 (306) 966-8412, Fax: +001 (306) 966-8413, E-mail: [james.nolan@usask.ca](mailto:james.nolan@usask.ca)

---

<sup>†</sup> Corresponding author.

**Keywords:** generalized additive models, cable-television industry, input separability.

**JEL Codes:** C14, L80.

## ABSTRACT

*We develop an example estimating a non-parametric cost function with Generalized Additive Models (GAMs). We demonstrate the feasibility of testing for input separability, providing a test for bias in GAMs. Our example focuses on cable television (CATV) provision, and we estimate a new non-parametric cost function for the Canadian industry using financial and operating data collected between 1990 and 1996. This period is of particular importance from a policy perspective because CATV in Canada was a rate- and entry-regulated industry prior to 1997. The results show that the input separability assumption holds and the GAM system provides a good fit to the CATV data.*

# 1 INTRODUCTION

The use of aggregated data has been the subject of much debate among economists. Opponents believe that working with aggregated data yields biased results leading to poor policy decisions. Proponents claim that the use of data at the aggregated level is unavoidable. For instance, when measuring productivity through index numbers, data should be aggregated. Otherwise the basic calculation is almost impossible due to a lack of sufficient degrees of freedom. Typically, data is aggregated into smaller groups of input and output variables. For example, although labour might exist in various categories such as “own”, “family”, and “hired”, in much applied econometric practice all of these categories are aggregated into one group of input known as “labour.” Similarly considering outputs, different agricultural products, such as wheat, barley, and rice may also be aggregated into one group of output called “grains.”<sup>1</sup>

Empirical studies estimating production, cost, and/or supply and demand functions can be divided into two major categories depending on the purposes of the studies. The first category attempts to obtain generalized results, approaching the problem by weakening restrictions derived from basic axioms. In contrast, there exist many studies that impose strong restrictions on the functional form to obtain ad hoc results. The problems arising from working with aggregated data stem from the latter group of studies. Richmond (2000) stated that the structure of applied microeconomic studies needs disaggregated data, which implies that practitioners must impose severe restrictions on their models to make them consistent with theoretical prescriptions. As a consequence, most applied studies are *data specific* since any changes in the data will often yield different results.

Separability of predictors is the most usual types of restriction used in applied microeconomic studies. Sono (1945) and Leontief (1947a; 1947b) were the first to discuss the issue of functional separability to deal with aggregation problems in consumer and producer theory, respectively. Simply put, a separable technology is a technology that can be divided into several stages. For example, in production theory we assume that the underlying technology is inherently weakly separable when we estimate production functions. In consumer demand theory the role of separability is more important because it allows practitioners to group various commodities into main budget categories. Since the results of such studies are dependent on

these crucial assumptions, it would be a wise decision if practitioners chose to examine whether the input separability assumption holds. Otherwise the overall validity of the study is questionable.

This paper aims to test the separability of inputs in a cost function that is estimated non-parametrically using the theory of *generalized additive models* (GAMs). This work is an extension of the study conducted by Law & Nolan (2002) who measured the impact of regulation on the Canadian basic cable television (CATV) with both parametric and non-parametric methods (data envelopment analysis). From a policy perspective, knowledge about input separability will better assist policy makers in understanding the nature of costs in this industry. For example, if capital and labour are not additively separable, then the marginal rate of technical substitution (MRTS) between these two inputs depends on the amount of each input chosen by the firm and this, in turn, increases the potential regulatory pricing constraints to distort input ratios from the levels that would have been chosen in the absence of the regulation. Law (1997, 2002) identified features of the regulatory structure established by the Canadian Radio-television and Telecommunications Commission (CRTC) that have some scope to generate these effects and Henderson et al. (1992) suggested that one of the provisions in the CRTC pricing structure results in significantly higher capital expenditures.<sup>2</sup>

This paper has five main sections. The next section reviews the studies examining input separability using non-parametric regressions. Section 3 provides an explanation of the methodology of GAMs. This is followed by a discussion of the non-parametric technique, known as *spline smoothing*. This technique is used to estimate the CATV cost function constructed in GAMs. Due to the specialised structure of GAMs, section 3 presents a statistical test that examines the separability of predictors in the model. Section 4 presents and discusses the findings of the model including tests for input separability. Finally, section 5 concludes the study and presents areas for further research.

## **2 LITERATURE REVIEW**

There are a few studies that have statistically tested the additive structure of GAMs, although their estimation process of the non-parametric regressions is different than the one that

will be used here. For example, Linton & Nielsen (1995) proposed a method to discriminate between the additive and multiplicative specifications of a mean response function. Since non-parametric regressions are less satisfactory with respect to the rate of convergence (to true parameters) whenever the number of predictors (independent variables) are more than two and simple plots are not available to assist practitioners in model selection, Linton & Nielsen (1995) used a simple kernel estimation procedure constructed based on the theory of marginal integration method (MIM) to estimate the model. The researchers used the information related to a sample of 534 individuals drawn randomly from the 1985 Current Population Survey collected by the U.S. Department of Commerce. Using the MIM, Linton & Nielsen (1995) could estimate a univariate predictor in both additive and multiplicative non-parametric regressions and investigated the additive structure of the model.

Chen et al. (1996) proposed a method to test the additive separability assumption in a GAM featuring a Cobb-Douglas production function using five inputs. Their model was superior to the Linton-Nielsen model from two perspectives: first, more than two predictors were used in the model, and second, they proposed an estimator for additive models with an explicit hat matrix, which did not use an iteration process. The researchers applied their model to a sample containing 250 observations drawn from a Wisconsin livestock farm database collected in 1987. Chen et al. (1996) concluded that all predictors were additively separable from each other except for hired labour.

Gozalo & Linton (2001) developed several kernel-based statistical tests to examine the additivity hypothesis in non-parametric regressions. The researchers specified a semi-parametric generalized method of moments (GMM) model in which discrete covariates were allowed. Gozalo & Linton (2001) were interested in examining the additivity hypothesis for the predictors since this structure-based assumption provides a basis from which the model can be interpreted and its estimators converted to their true values with reasonable convergence rates. Within the process, Gozalo & Linton (2001) were able to find the asymptotic distribution of the parameter estimators. In addition, they derived a locally asymptotic distribution for the additivity test statistics that they used. Since the Gozalo-Linton model yielded different results, the researchers specified a local asymptotic power system to rank the degree of validity of the results obtained

from conducting the semi-parametric structure of the model. Gozalo & Linton (2001) also conducted a simulation experiment with 2000 replications to study a binary response model of migration between East and West Germany as a function of age, income, rent, and a qualitative predictor, which measures migrants' degrees of satisfaction. Gozalo & Linton (2001) concluded that the predictors in the model were additively separable. Despite obtaining such results, the input separability statistical test proposed in the Gozalo-Linton model had some drawbacks. Specifically, the model was sensitive to the choice of bandwidths and the degree of the asymptotic approximations being used.

Finally, Haghiri et al. (2004) examined and compared technical efficiency measures of Ontario and New York dairy producers for the period 1992 to 1998 using a GAM, which was estimated by a non-parametric technique known as the locally weighted scatterplot smoothing (LOWESS). The researchers specified a production function to evaluate variation in the annual total milk production measured in hundredweight as response (dependent variable). To examine the additive separability of inputs, these authors conducted a residual deviance analysis method and concluded that labour and total feed costs were indeed additively separable.

### 3 METHODOLOGY

#### 3.1 Generalized Additive Models (GAMs)

The method of generalized additive models (GAMs), as proposed by Hastie & Tibshirani (1990), are extensions of generalized linear models (GLMs), which in turn, are extensions of classical linear models. A GAM maintains the additivity assumption of predictors and relaxes the linearity postulate. Suppose there are  $n$  observations on a random response  $Y$ , whose mean variation is measured by predetermined and/or random variable covariates  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)'$ . In this case, any multiple regression model can be written as

$$(1) \quad \begin{aligned} \mathbf{Y}_i &= \alpha + x_{i1}\beta_1 + x_{i2}\beta_2 + \dots + x_{ik}\beta_k + \varepsilon_i && \text{or} \\ \mathbf{Y}_i &= m(\mathbf{x}_i) + \varepsilon_i, && \text{for } i = 1, \dots, n \end{aligned}$$

where  $m(\mathbf{x}_i) = \alpha + \mathbf{x}_i' \boldsymbol{\beta}$ ,  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_k)'$ ,  $E(\varepsilon_i) = 0$ , and  $\text{Var}(\varepsilon_i) = \sigma^2$ . The driving assumption in equation [1] is that the relationship between the expected value of  $\mathbf{Y}_i$  and each of  $k$ -covariate elements of  $\mathbf{x}_i$  is linear and additive.

One way to relax the linearity assumption in equation [1] is to use surface smoothers, which can be thought of as non-parametric estimates of the regression model. A group candidate of surface smoothers is *kernel functions* whose characteristics in finding a local neighbourhood in  $k$  dimensions have some major drawbacks. Perhaps the most important shortcoming of kernel functions is the *curse of dimensionality*, which precludes practitioners from including more than two variables in the model. This problem can be waived if one uses GAMs.

To see how a GAM is constructed, consider equation [2]

$$(2) \quad E[\mathbf{Y}_i | \mathbf{x}_i = (x_{i1}, \dots, x_{ik})] = G \left[ \alpha + \sum_{j=1}^k f_j(x_{ij}) \right] = G[f(\mathbf{x}_i)]$$

in which  $f(\mathbf{x}_i) = \alpha + \sum_{j=1}^k f_j(x_{ij})$ ,  $\alpha$  is a constant, the  $f_j$ s are arbitrary univariate smooth functions; one for each predictor, and  $G(\cdot)$  is a fixed link function. The distribution of  $\mathbf{Y}_i$  follows an exponential family similar to GLMs. To avoid having free constant in each of the functions  $f_j$ , it is assumed that  $E[f_j(x_{ij})] = 0$ . This requirement, which is set in the range of  $1 \leq i \leq n$  and  $1 \leq j \leq k$ , implies  $E[f(\mathbf{x}_i)] = \alpha$  and is necessary for the purpose of identification. Therefore, GAMs allow the conditional mean of a response to be dependent on a sum of individual univariate functions where each of them contains one predictor of the covariate matrix. By relaxing of the linearity assumption in a GAM, the predictor's effects in equation [2] might be non-linear, because the functions  $f_j$  are now arbitrary (Schimek & Turlach, 2000).

This paper uses a non-parametric technique, the so-called *spline smoothing* approach, (also known as a *cubic spline smoothing*) developed by Wahba (1990), to estimate the mean response function modelled in GAMs. Spline smoothing provides a flexible methodology for fitting data in a non-parametric regression, and has been applied in a wide variety of sciences

such as analysis of growth data, medicine, remote sensing experiment, and economics (e.g., Pagan & Ullah, 1999, pp.91-93).

### 3.2 Spline Smoothing

Suppose the main purpose in equation [1] is to estimate  $m$  from sample data. A conventional econometric approach for estimating  $m$  and the parameters of the regression function is to use the least squares (LS) method by minimising the residual sum of squares

$$RSS(m) = \sum_{i=1}^n [y_i - \hat{m}(x_i)]^2$$

over all observations in relation to the assumed functional form of

$m(x_i) = \alpha + x_i\beta$ . However, the problem with using the LS approach is that there may not be a linear relationship between the response and predictors. One way to find the source of such failure is to use a Taylor-expansion series such as

$$(3) \quad m(x) = m(x_0) + m'(x_0)(x - x_0) + o(|x - x_0|^2)$$

in which  $m$ , an unknown function, is at least twice differentiable and there is a point  $x$  close to some fixed point  $x_0$ . Equation [3] states that for  $x$  close to  $x_0$ ,  $m$  follows a linear model whose intercept and slope are,  $m(x_0) - m'(x_0)x_0$  and  $m'(x_0)$ , respectively. This may occur in two extreme cases. First,  $m$  is assumed to be linear, which implies the slope  $m'(x_0)$  remains invariant and the residual term  $o(|x - x_0|^2)$  is small; an unrealistic assumption. Although this case uses too little of the information available in the data, it provides a useful summary of the sample observations and presents a satisfactory description of the features in the data. Second,  $m$  is assumed to have varying slope, which means that at each point  $x$  the slopes of the lines that connect each pair of responses may be different. Unlike the first scenario, this case uses too much information, but it does not provide a useful summary of the data and fails to deliver a satisfactory description of basic trends in the sample observations. This failure may occur due to the regression function in equation [3] rather than to the random-noise component of the model.

To circumvent this problem, assume that the rate of change in the slope of a function  $m$  is given by  $m''$ . Since this slope varies from one point to another, taking the integral from the entire changes in slope of the fitted function generates

$$(4) \quad \Phi(m) = \int_{x_1}^x m''(x)^2 dx$$

which suggests a new approach that can take into account the possibility that as the predictor changes the slope may change rapidly. Hence, we consider

$$(5) \quad RSS(m) + \tau \Phi(m), \quad \tau \geq 0,$$

which can be minimised over all functions provided that they are capable of being twice differentiated. In equation [5],  $\tau$  is called the *smoothing parameter* (or *span degree*) indicating the level of importance placed on the structure of the function given that the slope of the fitted function, to some extent, is flexible. For example, as  $\tau$  approaches infinity we get linear regressions with fixed slopes and, as  $\tau$  approaches zero, linear regressions with flexible slopes are obtained.

Eubank (2000) showed that if  $n$  in equation [4] is greater than or equal to two, then there exists a unique and computable minimizer  $\mathcal{G}_\tau$  for equation [5], known as a *cubic spline smoothing*. Cubic spline smoothing estimators are linear in the sense that one can find constants such as  $g_i(x)$ ,  $i = 1, \dots, n$  for each estimation point  $x$  such that

$$(6) \quad \mathcal{G}_\tau(x) = \sum_{i=1}^n g_i(x) y_i$$

Cubic spline smoothing estimators can solve the problem of fitting regressions with variant slopes, but at a cost. This problem, which arises due to a lack of theory and an appropriate algorithm, concerns how an appropriate smoothing degree is determined, given sample observations. As a result, cubic spline smoothing estimators are inconsistent and sensitive to the choice of smoothing degree, which means the choice of span degree is data specific. To mitigate

such problem, this study uses the *cross validation* (CV) method of Stone (1974) that is explained in the next section.

### 3.3 Model Specification

Consider a multiple variable cost function, which seeks to find a relationship between a response and predictors. The model can be written as

$$(7) \quad Y_{it} = f_i(X_{it}) + \varepsilon_{it}, \quad i = 1, 2, \dots, n_t, \text{ and } t = 1, 2, \dots, T. \quad \text{or}$$

$$f_i(X_{it}) = \alpha + \sum_{j=1}^k f_{ji}(X_{jit}) + \varepsilon_{it}$$

where  $Y_{it}$  is the total cost,  $f_i$  is the unknown functional form of the regression,  $X_{it}$  is a multidimensional series of input prices and output level with real values, i.e.,  $X_{it} \in \mathfrak{R}^k$ . Furthermore,  $\varepsilon_{it}$  is the random error term assumed to be distributed identically and independently with zero mean and constant variance. In equation [7],  $f_i$  must be estimated throughout a series of common  $k$  regressors used by each decision making unit,  $i$ , in the sample observations drawn independently from a population and can be estimated through a non-parametric method, such as spline smoothing. Moreover, the identification condition described earlier still holds.

For brevity, the estimation process is described for two predictors and for simplicity, the subscripts  $i$  and  $t$  are dropped. Assume a simple GAM such as equation [2], which can be written as

$$(8) \quad E[Y | x_1, x_2] = f(x_1, x_2) = \alpha + f_1(x_1) + f_2(x_2)$$

Furthermore, we may notice that

$$(9) \quad f_1(x_1) = \int f(x_1, x_2) g(x_2) dx_2 = E[f(x_1, X_2)]$$

By considering the assumption  $E\left[f_2(x_2)\right] = 0$ , we can obtain the following result,

$$(10) \quad \int f_2(x_2)g(x_2)dx_2 = 0$$

Thus,  $f_1(x_1)$  is estimated by  $\hat{f}_1(x_1) = T^{-1} \sum_{t=1}^T \hat{f}(x_1, x_{t2})$  where  $\hat{f}(x_1, x_{t2})$  is some semi-parametric estimator  $f(x_1, x_2)$ . Given the assumption that  $f$  is a twice-differentiable smooth function, the reliable semi-parametric estimators can now be obtained by using the *backfitting algorithm* introduced by Friedman & Stuetzle (1981) and subsequently modified by Breiman & Friedman (1985) as the *iterative smoothing process*. This algorithm, first, estimates  $\hat{f}_1(x_1)$  in equation [8] and then, while fixing the fitted function  $\hat{f}_1(x_1)$ , tries to estimate the mean regression function on  $x_2$  by smoothing the residual  $Y - \hat{\alpha} - \hat{f}_1(x_1)$ , which leads to the estimation of  $\hat{f}_2(x_2)$ . Since the backfitting algorithm is an iterative process, the next step is to improve the estimation of  $\hat{f}_1(x_1)$  by smoothing the residual  $Y - \hat{\alpha} - \hat{f}_2(x_2)$  on  $x_1$ , which, in turn, leads to enhance the estimators that are used to smooth the residual  $Y - \hat{\alpha} - \hat{f}_1(x_1)$  on  $x_2$  in the second step. This procedure continues until reliable and efficient estimators are achieved. It is important to know that the algorithm needs the initial estimation of  $\hat{f}_1(x_1)$  and hence an estimate of  $\hat{f}(x_1, x_{t2})$ , which is provided by using the spline smoothing technique. Interested readers can find more about the backfitting algorithm in Schimek (2000).

As mentioned earlier, the smoothing process is associated with the inconsistency and sensitivity of the estimates to the choice of span degree. Using the CV method, which provides optimal smoothing degree and, as a result, generates consistent estimators, can solve this problem. In particular, given equation [9] and [10], the CV method minimizes

$$(11) \quad \sum_{i,-t} \left[ Y_{it} - \hat{\alpha} - \hat{f}_{i,-t}(x_{it}) \right]^2$$

in which the estimation process of  $\hat{f}_{i,-t}$  follows two steps. In the first step, for a fixed individual data generator  $i, i = 1, 2, \dots, n_t$  and in every sequence of time period  $t, t = 1, 2, \dots, T$ , one pair of sample data, i.e., the  $i$ -th and  $t$ -th observations, is put aside and the mean response function  $f_i$ , defined in equation [7], is re-estimated based on the  $n - 1$  remaining observations. In the second step, the algorithm is repeated and continued to estimate  $f_i$  until convergence. According to Kneip & Simar (1996, p.192), this method is known as *leaving out the observation*  $(y_{it}, x_{it})$ .

### 3.4 Statistical Tests

We use the approach described to investigate features of the Canadian CATV cost function. Like other parametric and non-parametric regressions, the estimation of a cost function using GAMs has some drawbacks. From the theoretical point of view, although GAMs relax the linearity assumption of GLMs, they maintain the additivity structure of these types of models. This implies that the use of GAMs assumes that the predictors used in the model must be additively separable. However, “the additive separability of the predictors could be a wrong approximation of the real function” (Kneip & Simar, 1996, p.209). The additive assumption is not so important as separability because the former is a feature of parametric estimation. Separability can be tested through examination of the MRTS. Conversely, the assumption of input separability needs to be addressed in those non-parametric regressions, which use GAMs. Therefore, to verify applicability of the results obtained from the non-parametric approaches, we need to construct a statistical test to examine whether the input separability assumption holds.

Another major pitfall using GAMs arises from how these classes of models are estimated. This problem also occurs for those non-parametric techniques, such as LOWESS and/or spline smoothing. We stress that finding the optimal smoothing degree is data specific, which means that the span degree is different from one model to another. This is the reason the CV method is used in this study. In sum, given that we make relatively innocuous assumptions with respect to technology in the Canadian CATV industry, we feel that the gains from employing a GAM over the alternatives argue for its use in this analysis.

To examine the input separability assumption of the model this paper uses an analysis of *residual deviance*. Residual deviances, obtained from estimating GAMs, are aggregated into an *analysis of deviance* table. The value of the deviance is the logarithm of the likelihood ratio (LR), which follows a chi-squared distribution (Hastie & Tibshirani, 1990). The analysis of deviance can be used for statistical inferences in those models whose structure is based on the theory of GAMs (Bowman & Azzalini, 1997). The value for the residual deviance would be identical to the value of the residual sum of squares if only one predictor had been used in the model (Schimek & Turlach, 2000).

The value of deviance can be obtained by computing the values of two LR statistic tests: restricted and unrestricted. Equation [12] provides the computed value of deviance,  $\hat{\eta}$

$$(12) \quad D(m; \hat{\eta}) = 2 \{ l(\eta_{\max}; m) - l(\eta; m) \}$$

in which  $l(\eta_{\max}; m)$  and  $l(\eta; m)$  indicate the values of the restricted and unrestricted LR, respectively. Using a simulation experiment design based on the theory of GAMs, Hastie & Tibshirani (1990, p.282) showed that  $D(m; \hat{\eta})$  has asymptotic degrees of freedom equal to the differences in the dimensions between the two restricted and unrestricted models. Thus, a chi-square distribution is still a useful asymptotic approximation for screening the applicability of GAMs (Schimek, 2000).

#### 4 EMPIRICAL ANALYSIS

In this section, results of the estimation procedure are presented. We estimate a cost function for Canadian CATV industry, using financial and operating data collected between 1990 and 1996. The model specification is based on the theory of GAMs, and is estimated using a non-parametric technique called spline smoothing. In estimating the cost function we indirectly assume that the duality assumptions hold, which theoretically implies that a production function can be retrieved through estimating the cost function. This section is divided into two parts: *data analysis* and *statistical inference*. The former part presents a brief explanation of the sample data

followed by a description of the variables. The latter provides the result of a statistical test conducted to examine the separability of predictors.

#### *4.1 Data Analysis*

This study uses an unbalanced panel drawn from the CRTC database, and represents cable television providers operating between 1990 and 1996. The cable television firms submit their data annually for each licensed service area (LSA), which is then compiled by the CRTC. In total, this study uses 2804 sample observations obtained from 607 unique undertaking identification (UID) codes. A UID code is allocated to each LSA and thus the number of these codes indicates the number of cable systems in the sample. A new UID may be issued for an LSA whenever the identity of the operating CATV firm changes. In this paper, we consider each of the LSAs as one data-generating unit of CATV production over the time period of the study. In addition, we do not pool any firm's information in any sample period from preceding years since previous work suggested that the data at each point in time (annual) is sufficient to reflect long run decisions in the industry (see Law & Nolan, 2002, p.234). Finally, the richness of the database allows us to specify a relationship between total cost with input prices and the level of output for the entire sample of observations. Interested readers can find comprehensive descriptions of the Canadian CATV industry, the CRTC and its regulations, and the structure of LSAs, in Law (1997) and Law & Nolan (2002).

We used response and predictors similar to those employed by Law & Nolan (2002). Table 1 (Appendix) provides statistical descriptions of the variables. Output is defined as the total number of direct and indirect subscribers. Total cost is obtained by adding operating expenses to the user cost of capital, for both basic and non-basic services. Operating expenses, including the cost of labour, are computed by taking the difference between total expenses and those expenses incurred for programming. The price of labour is computed through a simple ratio of total salaries over the total number of staff employed within each of the LSAs. Finally, the rental price of capital is found by adding depreciation and financing rates. The former rate is derived from the CRTC database. The latter rate is obtained by the summation of a risk free interest rate and a risk premium. The "risk free" interest rate is the average of monthly ten-year-plus bond rates for a financial year ending in August. The product of the market risk premium and a systematic risk factor yields the risk premium for capital. According to Law & Nolan

(2002, p.247), the values of market risk premium and the systematic risk factor in Canada were equal to 0.042 and 0.60, respectively. All the dependent and independent variables are transformed to logarithms in order to mitigate possible heteroskedasticity in the model.

#### 4.2 *Estimation Results and Statistical Inference*

Following the methodology proposed in section 3, we estimate the total cost function of the Canadian CATV industry, constructed as equation [7] through a non-parametric regression using the spline-smoothing approach. The estimation procedure is conducted by writing code in S-Plus version 3.4 in the Unix Operation System. Table 2 and 3 (Appendix) shows the estimates of the cost function for the restricted and unrestricted regression models, respectively.

Using the statistical test described in section 3.4 we examined whether the non-parametric cost function model satisfies the implicit assumption of additive separability of the predictors. To conduct the analysis, first we introduce a new predictor obtained from the product of labour and capital prices. The idea behind the introduction of the new predictor is as follows. If the labour and capital were not separable then their product would add more information to the model. Since GAMs are susceptible to the problem of curse of dimensionality, the introduction of a new variable is valid. Second, using the LR statistic test described in equation [12], we test the null hypothesis of additive separability in the predictors in the non-parametric regression (restricted model) as opposed to the alternative hypothesis (unrestricted model). The computed value of deviance is 1.08. This is less than the critical value of chi-squared (3.84) at the 0.05 level of significance with one degree of freedom, so there is not enough evidence to show the population parameter of the new variable is significantly different from zero. In other words, the new variable did not add any information to the non-parametric model. Thus, we can conclude that the additive separability assumption between capital and labour holds.

Chambers (1997, p.42) stressed that in studying separability a couple of issues should be taken into account: (i) not all technologies are separable, and (ii) the concept of separability is most easily described in the context of continuously differentiable technologies. These points have been considered in the model. In fact, we assumed that the mean response variable is smooth enough to be differentiable at least twice, which is also a crucial assumption to any non-

parametric regression analysis. Using the LR statistic test, we concluded that the additive separability assumption concerning the predictors holds at the 0.05 level of significance.

Input separability is derived by examining how the MRTS is affected by the changes in another input available in a third dimension. By failing to reject the null hypothesis we conclude that any changes in labour or capital in the sample observations might alter the span degree in the space of these inputs individually, but do not affect the span degree of labour and capital together. This implies that the MRTS between capital and labour does not respond to any changes imparted by the new variable, i.e., the multiplicative product of labour and capital.

In summary, we find that CATV is an industry where the additive separability assumption between capital and labour holds. This finding supports the potential applicability of the model to other studies. Moreover, we strongly recommend that researchers examine the input separability assumption if they use GAMs due to the specific structure of these types of models.

## **5 CONCLUSIONS**

The primary purpose of this paper is to provide a successful example of the use of GAMs and demonstrate that the choice of a non-parametric technique need not imply the loss of ability to test specific characteristics of the model. If one wishes to examine a cost function within the framework of GAMs, a researcher must proceed with the determination of additive separability of inputs. In this case, we have shown that labour and capital are additively separable in the production of cable television services in Canada, confirming previous work about the nature of the provision of CATV services.

This example using GAMs suggests avenues for future research. It would be useful to be able to test not only separability but also the substitutability of inputs. This ability to test whether inputs are used more as complements or as substitutes would also have policy relevance for this particular industry. Optimal rate of return regulation can take different forms if returns are earned on an input (capital) which is complementary rather than a substitute for other inputs.

Another area of future research would be to test separability across output categories. Over the sample period, the CRTC maintained a rate system that indicated the formula by which

CATV operators would be able to obtain rate increases on the basis of increases in the amount of capital they have installed for the provision of basic service. The CATV firms were permitted to designate assets or declare what fractions of capital expenditures were for basic service.<sup>3</sup> The use of the CRTC database assumes that there is no bias introduced by the incentive to overstate the fraction of capital that is dedicated to the production of basic service. It would be interesting to test whether basic and non-basic assets are indeed separable. The results would have implications for other regulated situations in which artificial reporting requirements create separate categories for production either for inputs or for outputs.

## ENDNOTES

- [1] See Bertin et al. (1996), which discusses one condition in which both groups may be satisfied, that is, when there is little scope for substitution and hence aggregation does not impose on the data any conditions other than those which the economic actors face.
- [2] “Over the period 1986 to 1990, a substantial increase occurred in the gross and net fixed assets per subscriber -- a reflection in part of the industry’s response to the CAPEX [capital expenditure] provision for rate increases. For example, if no increase would have occurred in the value of average gross fixed assets per subscriber in that period without the CAPEX provision, the net increase in capital expenditures that may be attributed to the CAPEX provision amounts to roughly \$100 per subscriber over the five-year period -- a 25 per cent increase in gross fixed assets per subscriber.” (Henderson et al., 1992, p.22).
- [3] See Law (1997), especially Table A5: Proportion of Historical Assets Allocated to Basic Service.

## REFERENCES

- Bertin, A.L., Bresnahan T.F., and D.M.G. Raff D. (1996) "Localised Competition and the Aggregation of Plant-Level Increasing Returns: Blast Furnaces, 1929-1935," *Journal of Political Economy*, Vol.104, pp. 241-266.
- Bowman, A. D. and A. Azzalini (1997), *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustration*, New York: Oxford University Press.
- Breiman, L. and J.H. Friedman (1985), "Estimating Optimal Transformations for Multiple Regression and Correlation," *Journal of American Statistics Association*, Vol.80, pp.580-619.
- Chambers, R.G. (1997), *Applied Production Analysis: The Dual Approach*, New York: Cambridge University Press.
- Chen, R., Hardle, W., Linton, O.B. and E. Sevarance-Lossin (1996), "Nonparametric Estimation of Additive Separable Regression Models," *Statistical Theory and Computational Aspects of Smoothing*, New York: Physica Verlag, pp.247-65.
- Eubank, R.L. (2000), "Spline Regression," in M.G. Schimek (ed.), *Smoothing and Regression: Approaches, Computation, and Application*, New York: John-Wiley and Sons, pp.1-18.
- Friedman, J.H. and W. Stuetzle (1981), "Projection Pursuit Regression," *Journal of American Statistics Association*, Vol.76, pp.817-23.
- Gozalo, P.L. and O.B. Linton (2001), "Testing Additivity in Generalized Nonparametric Regression Models with Estimated Parameters," *Journal of Econometrics*, Vol.104, pp.1-48.
- Haghiri, M., Nolan, J.F. and K.C. Tran (2004), "Assessing the Impact of Economic Liberalisation across Countries: A Comparison of Dairy Industry Efficiency in Canada and the United States," *Journal of Applied Economics* (forthcoming).
- Henderson, J., Elek, A. and C. Gibson (1992), *A Financial and Regulatory Analysis of the Canadian Cable Television Industry to the Year 2001*, Toronto: KPMG Peat, Marwick, Stevenson & Kellog.
- Hastie, T.J. and R.J. Tibshirani (1990), *Generalized Additive Models*, London: Chapman and Hall.
- Kneip, A. and L. Simar (1996), "A General Framework for Frontier Estimation with Panel Data," *Journal of Productivity Analysis*, Vol.7, pp.187- 212.
- Law, S.M. (1997), "Economic Policy Interactions: Intellectual Property Rights and Competition Policy; Exclusive Licensing and Rate Regulation," PhD Dissertation, Department of Economics, University of Toronto.

- Law, S.M. (2002), "The Problem of Market Size for Canadian Cable Television Regulation," *Journal of Applied Economics*, Vol.34 (1), pp.87-99.
- Law, S.M. and J.F. Nolan (2002), "Measuring the Impact of Regulation: A Study of Canadian Basic Cable Television," *Review of Industrial Organisation*, Vol.21, pp. 231-249.
- Law, S.M. and J.F. Nolan (2004), "Unintended and Persistent Consequences of Regulation: The Case of Cable Television Provision in Canada," *Journal of Network Industries*, Vol.4 (4), pp.389-408.
- Leontief, W.W. (1947a), "A Note on the Interrelation of Subsets of Independent Variables of a Continuous Function with Continuous first Derivatives," *Bulletin of the American Mathematical Society*, Vol.53, pp.343-350.
- Leontief, W.W. (1947b), "Introduction to a Theory of the Internal Structure of Functional Relationships," *Econometrica*, Vol.15, pp.361-373.
- Linton, O.B. and J.P. Nielsen (1995), "A Kernel Method of Estimating Structured Nonparametric Regression Based on Marginal Integration," *Biometrika*, Vol.82, pp. 93-100.
- Pagan, A. and A. Ullah (1999), *Nonparametric Econometrics*, Cambridge: Cambridge University Press.
- Richmond, J. (2000), "Separability and Specification Tests," Discussion papers, No.527, U.K.: Department of Economics, University of Essex.
- Schimek, M.G. (2000), *Smoothing and Regression: Approaches, Computation, and Application*, (Ed.), New York: John-Wiley and Sons.
- Schimek, M.G. and B.A. Turlach (2000), "Additive and Generalized Additive Models," in M.G. Schimek (ed.), *Smoothing and Regression: Approaches, Computation, and Application*, New York: John-Wiley and Sons, pp.277-327.
- Sono, M. (1945), "The Effect of Price Changes on the Demand and Supply of Separable Goods," (in Japanese), *Kokumin Keisai Zasshi*, Vol.74, pp.1-51.
- Stone, M. (1979), "Cross-Validatory Choice and Assessment of Statistical Predictions," *Journal of the Royal Statistical Society, Series B*, Vol.36, pp.111-147.
- Wahba, G. (1990), "Spline Functions for Observational Data," *CBMS-NSF Regional Conference Ser. SIAM*, Philadelphia.

## APPENDIX

**Table 1. Statistical Description of the Variables**

<b>Variable</b>	<b>Mean</b>	<b>S. D.</b>	<b>Min.</b>	<b>Max.</b>
<b>Total Costs (\$ 000)</b>	3,461.5	13,038.6	3.5	227,593.2
<b>Price of Labour (\$)</b>	34,348.1	17,369.7	2021.0	198,041.0
<b>Price of Capital</b>	0.29436	0.12234	0.12027	1.86899
<b>Total number of Subscribers</b>	13,584	46,578	21	714,338

Source: Sample data.

**Table 2. Estimation Results from GAM (Restricted Model)**

<b>Predictors</b>	<b>DF</b>	<b>Non-parametric DF</b>	<b>Non-parametric F-value</b>	<b>P-value</b>
Intercept	1			
Price of Labour	1	3	7.7438	0.00004
Price of Capital	1	3	4.3320	0.00469
Total number of Subscribers	1	3	69.8768	0.00000
Residual Deviance		220.9124	Null Deviance	8165.884
Residual Degrees of Freedom		2790.993	Null Degrees of Freedom	2803.000

Source: Sample data.

**Table 3. Estimation Results from GAM (Unrestricted Model)**

<b>Predictors</b>	<b>DF</b>	<b>Non-parametric DF</b>	<b>Non-parametric F-value</b>	<b>P-value</b>
Intercept	1			
Price of Labour	1	3	4.0462	0.00698
Price of Capital	1	3	4.8711	0.00221
Total number of Subscribers	1	3	70.7926	0.00000
Product (Labour & Capital)	1	3	4.2875	0.00500
Residual Deviance		219.8237	Null Deviance	8165.884
Residual Degrees of Freedom		2786.992	Null Degrees of Freedom	2803.000

Source: Sample data.

