# Non-parametric Estimates of Technology Using Generalized Additive Models: Input Separability and Regulation in Canadian Cable Television

Morteza Haghiri

Memorial University at Corner Brook

University Drive, Corner Brook, NL, A2H 6P9, Canada

Tel: 1-709-637-6200 ext 6145          E-mail: mhaghiri@mun.ca


Stephen M. Law (Corresponding author)

Department of Economics, Mount Allison University

144 Main Street, Sackville, NB, E4L 1A7, Canada

Tel: 1-506-364-2355          E-mail: slaw@mta.ca


James F. Nolan

Department of Bioresource Policy, Business and Economics

University of Saskatchewan

51 Campus Drive, Saskatoon, SK, S7N 5A8, Canada

Tel: 1-306-966-8412          E-mail: james.nolan@usask.ca

## Abstract

We develop a non-parametric cost function using generalized additive models and demonstrate how to test for input separability. Our empirical example focuses on Canadian cable television (CATV) provision. We estimate a new non-parametric cost function for this 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 generalized additive model yields reliable estimates of CATV technology during this time period.

**Keywords**: Generalized additive models, Cable television industry, Input separability

JEL classification codes: C14, L50, L97

## 1. Introduction

Empirical estimates of production, cost, and/or supply and demand functions can be divided into two major categories depending on the purpose of the study. The first category describes general results, whereby a problem is analysed by weakening certain estimation restrictions founded on economic axioms. In contrast, there are studies that impose strong restrictions on functional form, which often obtain ad hoc results. Richmond (2000) observed that the structure of most applied microeconomic studies calls for disaggregated data. This implies that practitioners must impose severe restrictions on their models to make them consistent with theoretical prescriptions. As a consequence, many applied studies are in fact data-specific, since any modifications to the data can yield different results.

In applied microeconomic studies of production, separability of inputs is a very commonly used restriction in estimation. Sono (1961) [first published in Japanese in *Kokumin Keizai Zasshi*, LXXIV (1945), 1-51] and Leontief (1947a; 1947b) were the first to discuss the issue of functional separability in producer (and consumer) theory. 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 primary budget categories. Since empirical results are often dependent on these crucial assumptions, it is almost always a good decision for practitioners to formally examine whether the assumption of input separability holds. Here, we show an example where this practice is crucial to develop a non-parametric econometric estimation of technology.

This particular application is also an extension of earlier research by Law & Nolan (2002), who sought to measure the impact of regulation by the Canadian Radio-television and Telecommunications Commission (CRTC) on the Canadian basic cable television industry using both parametric and non-parametric (data envelopment analysis) methods. We estimate a novel non-parametric cost function for the Canadian cable television industry (CATV) using a generalized additive model (GAM). The structure of Canadian CATV leads us to believe that this particular specification is appropriate for cost estimation. In addition, we explain the need to conduct a formal test for input separability with respect to this estimation process.

This paper has five sections. The next section reviews those studies examining input separability using non-parametric regressions. Section three provides an explanation of generalized additive models. This is followed by a discussion of our chosen non-parametric technique, known as spline smoothing. The technique is used to estimate a Canadian CATV cost function constructed using GAMs. This section also motivates a statistical test for the separability of inputs in the model. Section four presents and discusses the findings of the estimates, including the tests for input separability. Finally, section five concludes the study and discusses areas of future research.

## 2. Studies on Input Separability Using GAMs

A number of recent studies discuss aspects of the estimation of generalized additive models in a context similar to our research, but in most cases the estimation processes used are somewhat different than the methods to 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 have problems with rate of convergence (to the true parameters) whenever independent variables number more than two, the authors used a simple kernel estimation procedure to estimate the model. In turn, they estimated a univariate predictor in both additive and multiplicative non-parametric regressions to examine the additive structure of the model.

Chen et al. (1996) proposed a method to test additive separability in a generalized additive model based on a Cobb-Douglas production function with five inputs. This model is superior to the Linton-Nielsen model for two reasons: first, more than two independent variables can be used in the model; and second, they developed an estimator for additive models with an explicit "hat" matrix which did not rely on an iterative process.

Gozalo & Linton (2001) developed a semiparametric generalised method of moments model in which discrete covariates were permitted. The authors were interested in examining the additivity hypothesis for the predictors since this structural assumption provided a basis upon which the model could be interpreted, while its estimators converged with reasonable convergence rates. Gozalo & Linton were able to find the asymptotic distribution of the parameter estimators and they derived a locally asymptotic distribution for the additivity test statistics. They concluded that the predictors in the model were additively separable. However, the input separability statistical test proposed by Gozalo-Linton has some drawbacks. Specifically, it is sensitive to the choice of bandwidths and the degree of the asymptotic approximations being used.

Finally, Haghiri et al. (2004) examined the technical efficiency of dairy farming using a generalized additive model. They estimated their model using a non-parametric technique known as locally-weighted scatterplot smoothing (LOWESS), and specified a production function to evaluate variation in their dependent variable. To examine the additive separability of inputs, the authors formulated a residual deviance analysis method and were able to conclude that labour and feed costs were indeed additively separable.

In the spirit of the recent research on this topic, we will address the question of input separability in the Canadian CATV industry using non-parametric specifications for production. In this manner, we seek to advance applied econometric research on measurement of separability, as well as extend analysis of regulatory transition and efficiency in a major industry for which the network characteristic of production is important.

## 3. Methodology

### 3.1 Generalized Additive Models (GAMs)

The model developed for this research builds on earlier studies. The method of GAMs, as proposed by Hastie & Tibshirani (1990) is extension of generalized linear models, which in turn, are extensions of classical linear models. A GAM maintains the additivity assumption for the predictors and relaxes the linearity postulate (Hastie & Tibshirani, 1990). Suppose there are $n$ observations on a random response $Y$, whose mean variation is measured by

predetermined and/or random variable covariates $\mathbf{X} = (\mathbf{x}_1,..,\mathbf{x}_n)'$. In this case, multiple regression models can be written as

$$\mathbf{Y}_i = \alpha + x_{i1}\beta_1 + x_{i2}\beta_2 + .... + x_{ik}\beta_k + \varepsilon_i \qquad \text{or}$$

$$\mathbf{Y}_i = m(\mathbf{x}_i) + \varepsilon_i, \qquad \text{for} \qquad i = 1,\ldots, n \qquad (1)$$

where $m(\mathbf{x}_i) = \alpha + \mathbf{x}_i'\%$, $\beta = (\beta_1,...,\beta_k)'$, $E(\varepsilon_i) = 0$, and $\text{Var}(\varepsilon_i) = \sigma^2$. The key 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 (Hastie & Tibshirani, 1990).

From Hastie & Tibshirani (1986), 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 of candidates for surface smoothers are kernel functions, but with respect to finding a local neighbourhood in $k$ dimensions, these have 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 (Hastie & Tibshirani, 1986). This problem can be avoided if the method of generalized additive models is used.

Schimek & Turlach (2000) explain how a GAM is constructed in this context. Consider equation [2]

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

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 $G(.)$ is a fixed link function. The distribution of $\mathbf{Y}_i$ follows an exponential family, in a manner similar to generalized linear models. To avoid having free constants in each of the functions $f_j$, it is assumed that

$E\lfloor f_j(x_{ij})\rfloor = 0$. This requirement, which is set in the range of $1 \le i \le n$ and $1 \le j \le k$, implies $E[f(\mathbf{x}_i)] = \alpha$ and is necessary for the purpose of identification (Schimek & Turlach, 2000). Thus 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 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).

For this paper, we employ a non-parametric technique known as the spline smoothing approach. This was developed by Wahba (1990) and we use it to estimate the mean response function modelled in our generalized additive models. Spline smoothing provides a flexible methodology for fitting data in a nonparametric regression, and has been used in a wide variety of applications, including analysis of growth data, medical research, remote sensing experiments, and economics (e.g., Pagan & Ullah, 1999, pp. 91-93).

*3.2 Spline Smoothing*

Suppose the primary goal of equation [1] is to estimate $m$ from the sample data. A conventional econometric approach for estimating $m$ and the parameters of the regression function is to use the least squares method by minimizing 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 least squares approach is that there may not be a linear relationship between the response and predictors (Wahba, 1990). One way to avoid this problem is to employ a Taylor-series expansion, such as

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

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, if $m$ is assumed to be linear, which implies the slope $m'(x_0)$ remains invariant and the residual term $o\left|x - x_0\right|^2$ is small, the latter

being an unrealistic assumption (Wahba, 1990). Although this case uses too little of the information available in the data, it does provide a useful summary of the sample observations and a satisfactory description of the features in the data. Second, $m$ could be assumed to have varying slope, meaning that at each point $x$ the slopes of the lines that connect each pair of responses may be different. Unlike the first case, the latter demands too much information, and it does not provide a useful summary of the data while failing to deliver a satisfactory description of basic trends in the sample observations. Note that this failure can occur because of the regression function in equation [3] rather than the random-noise component of the model (Wahba, 1990).

To circumvent this problem, Wahba (1990) assumes 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, we can take the integral from the set of changes in the slope of the fitted function. This generates equation [4]

$$\Phi\left(m\right) = \int_{x_1}^{x_n} m''\left(x\right)^2 \, \mathrm{d}x \tag{4}$$

which suggests a new approach that can account for the possibility that as the predictor changes, the slope may change rapidly (Wahba, 1990). Hence, we consider the following function

$$RSS\left(m\right) + \tau\,\Phi\left(m\right), \quad \tau \geq 0, \tag{5}$$

that can be minimised over all admissible functions provided that they are twice differentiable. 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. By way of example, as $\tau$ approaches infinity we obtain linear regressions with fixed slopes and, as $\tau$ approaches zero, we get linear regressions with flexible slopes (Wahba, 1990).

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

$$\vartheta_\tau\left(x\right) = \sum_{i=1}^{n} g_i\left(x\right) y_i \tag{6}$$

Cubic spline smoothing estimators can solve the problem of fitting regressions with variant slopes, but at a cost. The problem concerns how an appropriate smoothing degree is determined, given a set of sample observations. Consequently, cubic spline smoothing estimators are inconsistent and sensitive to the choice of smoothing degree, meaning that the choice of span degree is data specific. In order to mitigate this issue we employ the cross validation (CV) method of Stone (1974), a technique explained in the next section.

*3.3 Model Specification*

Consider a multi-variable cost function, which is seeks a relationship between a response and predictors. Such a model can be written as

$$Y_{it} = f\left(X_{it}\right) + \varepsilon_{it}, \quad i = 1, 2, \ldots, n_t, \text{ and } t = 1, 2, \ldots, T. \qquad \text{or}$$

$$f\left(X_{it}\right) = \alpha + \sum_{j=1}^{k} f_{ji}\left(X_{jit}\right) + \varepsilon_{it} \tag{7}$$

where $Y_{it}$ is total cost, $f$ is an unknown functional form of the cost relationship, $X_{it}$ is a multidimensional series of input prices and outputs with real values, i.e., $X_{it} \in \Re^k$ and $\varepsilon_{it}$ is a random error term, assumed to be distributed identically and independently with zero mean and constant variance. In equation [7], $f$ must be estimated throughout a series of common $k$ regressors employed by each firm $i$ in the sample observations. These are drawn independently from a population and can be estimated via non-parametric methods, such as the aforementioned spline smoothing. Moreover, the identification condition described earlier still holds.

The estimation process is described below for two predictors. For simplicity of exposition, the subscripts $i$ and $t$ are dropped. Assume a simple generalized additive model such as that found in equation [2], that can be written as

$$E\left[Y \mid x_1, x_2\right] = f\left(x_1, x_2\right) = \alpha + f_1\left(x_1\right) + f_2\left(x_2\right) \tag{8}$$

Furthermore, notice that

$$f_1(x_1) = \int f(x_1, x_2) g(x) \, dx_2 = E[f(x_1, X_2)] \tag{9}$$

By considering the assumption $E[f_2(x_2)] = 0$, we obtain the following result (Stone, 1974),

$$\int f_2(x_2) g(x_2) \, dx_2 = 0 \tag{10}$$

Note that Chen et al. (1996) show that $f_1(x_1)$ can then be 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 semiparametric estimator $f(x_1, x_2)$. Given the assumption that $f$ is a twice-differentiable smooth function, reliable semi-parametric estimators can be obtained by using the backfitting algorithm of Friedman & Stuetzle (1981), a technique that was subsequently modified by Breiman & Friedman (1985), as the iterative smoothing process. The latter algorithm first estimates $\hat{f}_1(x_1)$ in equation [8] and then, while fixing the fitted function $\hat{f}_1(x_1)$, estimates the mean regression function on $x_2$ by smoothing the residual $Y - \hat{\alpha} - \hat{f}_1(x_1)$, leading 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, enhances 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 statistically reliable and efficient estimators are achieved. It is important to know that the algorithm needs an initial estimation of $\hat{f}_1(x_1)$ and hence an estimate of $\hat{f}(x_1, x_{t2})$, and this is provided using the spline smoothing technique. For those interested, additional information about the backfitting algorithm can be found in Schimek & Turlach (2000).

The smoothing process is associated with the inconsistency and sensitivity of the estimates to the choice of span degree. The CV method, which provides an optimal smoothing degree and thus generates consistent estimators, can solve this problem (Stone, 1974). In particular, given equation [9] and [10], the CV method minimises

$$\sum_{-(i,t)} \left[ Y - \hat{\alpha} - \hat{f} \right]^2 \tag{11}$$

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

*3.4 Statistical Inference*

We use this combined approach to investigate features of a Canadian cable television cost function. Like most regressions, the estimation of a cost function using GAMs has some drawbacks. From a theoretical point of view, although GAMs relax the linearity assumption of generalised linear models, they maintain the additivity structure of these types of models. This implies that the use of GAMs also assumes that the predictors used in the model must be additively separable. But as some have noted, "the additive separability of the predictors could be a wrong approximation of the real function" (Kneip & Simar, 1996, p. 209). In fact, the additivity assumption is not as important as separability because the former is a feature of parametric estimation. And separability in production can be tested through examination of the estimated marginal rates of technical substitution (MTRS). However, the assumption of input separability needs to be addressed in applicable non-parametric regressions that use GAMs. Therefore to verify the general applicability of the results obtained from these non-parametric approaches to the estimation of production, we must construct a statistical test to examine whether the input separability assumption holds.

Another major pitfall that occurs using GAMs is a consequence of how these models are estimated, noting that this problem also occurs with other non-parametric techniques such as LOWESS and spline smoothing. In fact, finding the optimal smoothing degree is data-specific, meaning that the span degree will necessarily be different from one model to another. The latter is the primary reason why the CV method is incorporated into this study. Ultimately, given that we need to make relatively innocuous assumptions with respect to the structure of technology in the industry under analysis, we offer that the gains in estimation and inference from employing a GAM over the alternatives argue for its use in this type of applied econometric analysis.

Input separability in Canadian CATV will be examined through an analysis of residual deviance. When performing estimates with GAMs, analysis of deviance can be used for statistical inference (Bowman & Azzalini, 1997). Residual deviances obtained from our generalized additive model estimates are aggregated into a deviance table. The value of the deviance is the logarithm of the likelihood ratio (LR) and follows a chi-squared distribution (Hastie & Tibshirani, 1990). Note as well that the value for residual deviance is identical to the value of the residual sum of squares if only a single predictor is used (Schimek & Turlach, 2000).

Residual deviance is obtained by computing two LR statistical tests (restricted and unrestricted). Equation [12] shows the computed value of deviance, $\hat{\eta}$

$$D\left(m\,;\hat{\eta}\right)=2\left\{l\left(\eta_{\max}\,;m\right)-l\left(\eta\,;m\right)\right\} \tag{12}$$

in which $l\left(\eta_{\max}\,;m\right)$ and $l\left(\eta\,;m\right)$ indicate the restricted and unrestricted LR values, respectively. Hastie & Tibshirani (1990, p. 282) showed that $D\left(m\,;\hat{\eta}\right)$ has asymptotic degrees of freedom equal to the differences in the dimensions between the restricted and unrestricted models. Thus, a chi-square distribution is a good asymptotic approximation for testing the applicability of generalized additive models (Schimek & Turlach, 2000).

## 4. Empirical Analysis

Given that GAMs are appropriate to the estimation of production, we estimate a new non-parametric cost function for the Canadian cable television industry covering the period between 1990 and 1996. In turn, the GAMs model is estimated using spline smoothing techniques. We specify the cost function as a function of total output, i.e., the total number of basic and non-basic cable subscribers, as well as input prices, including the prices of labor, capital, and materials. In performing this estimation, we are also indirectly assuming that standard duality assumptions hold, so that a production function can be retrieved via estimation of this cost function.

One reason for our interest in input separability in the context of Canadian CATV is the importance of input choice as a consequence of industry regulation. In particular, the regulatory rules in this industry at that time provided for allowable price increases based on the costs of installation of additional capital for the provision of basic CATV service. This aspect of the CATV regulatory issue in Canada is related to concerns about rate of return regulation because of the potential presence of the Averch-Johnson-Wellisz (AJW) effect (Averch & Johnson, 1962, and Wellisz, 1963). As is well known in the regulatory economics literature, the AJW effect distorts the capital-labour ratio for the output level chosen by the regulated firm. While it is possible that the regulated firm might have chosen the same capital-labour ratio while providing a different level of service, one interpretation of the AJW model is that regulated firms make a simultaneous choice of capital, labour and output that leads to some distortion or productive inefficiency when compared to the choices made by an identical unregulated firm. Although Law (1997, 2002) identified key features of the regulatory structure established by the CRTC for Canadian CATV that led to cost inefficiencies, and Henderson et al. (1992, p.22) suggested that the CRTC pricing structure for Canadian CATV resulted in significantly higher capital expenditures than would be expected, to date there are no studies that conclusively identify the AJW effect in this candidate industry.

Furthermore, theory tells us that if aggregated capital and labor inputs are not additively separable, then the marginal rate of technical substitution (MRTS) between these inputs depends on the amount of each input chosen by the firm. This situation increases the possibility that regulatory pricing constraints will distort input ratios from those levels that would have been chosen in the absence of regulation. However if input separability in production is confirmed, then the likelihood that this negative effect of regulation could have occurred is reduced (see also Law, 2008). We conclude that if the latter situation is identified empirically, although regulation as practiced in Canadian CATV at that time might have led to other unfortunate consequences for economic efficiency, this regulation would not likely have been responsible for any distortion of input choice.

Estimating this particular non-parametric cost function for the Canadian cable television industry allows for testing whether various capital and labour inputs in this industry were additively separable. Necessarily, we are studying the technology of the regulated firms in the sample. Färe and Logan (1983) noted that "to reconstruct the rate-of-return

regulated production function, it is necessary to have knowledge about the rate-of-return constraint as well as to know the rate-of-return regulated cost function." In fact, we do not attempt to construct an assumed regulatory constraint that captures exactly the interaction between the firm and the regulator, and if we were to make this attempt all conclusions would be conditional on the accuracy of the formulation of the regulatory constraint. Instead, following Nelson and Wohar (1983, 1987) among others, we present results and conclusions which pertain to the operations of the actual regulated firms rather than those of hypothetical unregulated (but otherwise equivalent) firms. Given the latter set of related studies in regulatory economics, we offer that our chosen approach still generates useful results about the impact of regulatory constraints.

This section is divided into three sub-sections. First, the sources and the general characteristics of data are discussed and we provide a summary of the descriptive analysis of the variables used in the model. The second sub-section presents the estimation results followed by statistical inferences drawn from examining the null hypotheses of separability amongst inputs in the model. Finally, the third sub-section discusses the findings from the estimated model and discusses implications of this research.

*4.1 Data and Variable Descriptions*

The non-parametric cost function we estimate for the Canadian cable television industry uses an unbalanced panel of data drawn from several Canadian Radio-television and Telecommunications Commission (CRTC) databases. The CRTC databases contain information on cable television providers who operated in Canada between 1990 and 1996. Cable television firms were assigned operating areas by the CTRC, and each of these is referred to as a licensed service area (LSA). Each CATV operator was obliged to submit extensive data on operations, which was compiled by the CRTC.

The unbalanced panel contains 1,041 observations related to 242 unique undertaking identification (UID) codes. A UID code is allocated to each LSA so that the number of the codes indicates the number of individual cable operations in the sample. During the sample period, a new UID code would also be issued for an LSA whenever the identity of the operating CATV firm changed.

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 individual operation's information in any sample period from preceding years since previous work suggests that the data at each point in time (annual) is sufficient to reflect long run decisions in the industry (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) or Law & Nolan (2002).

Here we use dependent and independent variables similar to the set used by Law & Nolan (2002), with the exception that the data for each independent variable are provided separately for basic and non-basic cable subscribers. Subscribers to basic service receive a package of signals corresponding to the set of "local" television broadcasts; subscribers to non-basic service also receive "pay" or "discretionary" services which are packages of signals assembled by the CATV operator and include specialty services such as sports channels, movie channels, or community programming. Note that a subscription to basic service is a condition of a subscription to non-basic service.

Table 1 presents statistical summaries of the variables. Operating expenses, including the cost of labor, are computed by taking the difference between total expenses and those expenses incurred for programming. Total cost is obtained by adding operating expenses to the user cost of capital, for both basic and non-basic services. Output is defined as the number of basic and non-basic subscribers. The price of labor is computed through a simple ratio of total salaries over the total number of staff employed to provide services for basic and non-basic subscribers within each of the LSAs. The rental price of capital is found by adding depreciation and financing rates. The former rate is derived from the CRTC database, while the latter rate is obtained by the summation of a risk free interest rate and a risk premium specific to CATV operations. The "risk free" interest rate is the average of monthly ten-year-plus bond rates for a financial year, and 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. The price of materials is the per unit expenditure on intermediate material inputs, meaning that cost not allocated to expenditures for labor or for capital is distributed on a per cable kilometre and per channel basis for basic and for non-basic service for all years. Finally, all variables are computed per unit of output and transformed to logarithms in order to mitigate possible heteroskedasticity in the data.

Insert Table 1 Here

*4.2 Estimation Results & Statistical Inference*

Using the methodologies appropriate to the research issues as described in the previous sections, we estimate a total cost function for the Canadian CATV industry using the 1990-1996 data. We construct the function as in equation [7], estimated with a non-parametric regression using the spline-smoothing approach. The regression was run using the *R* software package (see, Ihaka and Gentleman, 1996), which is a more recent implementation of *S-Plus* (Version 4.0). Table 2 reports these estimates of the primary cost function using GAMs – this is our restricted model. The adjusted *R*-squared for this model is 98.6 per cent. All the coefficients are statistically significant at the 95 per cent confidence level, with the exception of the price of labor shared in basic subscribers, which is statistically different from zero at the 0.069 level of significance (see Table 2).

Insert Table 2 Here

Next, using the statistical test described in equation [12], we check whether the non-parametric cost function estimated satisfies the implicit assumption of additive separability of the predictors. Specifically, we examine whether changes in one of the factor prices could affect the total cost through the output variable. From a policy perspective, regulators might be interested to know whether changes in the price of labor shared by non-basic services have any impact on the total cost of providing services for basic subscribers. Or, if the price of capital shared in basic services changes, does this have any effect on the total cost through the component of output that is basic subscribers? To answer these types of policy questions along with the econometric issue of input separability, we need to estimate an unrestricted version of the CATV cost function in a similar fashion to the restricted model.

To build our unrestricted version of the CATV cost function, a new explanatory variable must be produced using the product of the pair of inputs under analysis. By way of example, consider the steps required to test the null hypothesis of whether changes in the price of labor shared by non-basic subscribers could affect the total cost of providing services for the basic subscribers. First, the new explanatory variable is obtained from the product of basic subscribers and the price of labor shared in non-basic subscribers. The idea behind the introduction of the interactive predictor is that if the number of basic CATV subscribers influences the labor shared in non-basic subscribers, it would affect the total cost of operations. In other words, if the workloads required by the non-basic subscribers affect the amount of labor needed by basic subscribers, then this would have an effect on the total cost of operations. Hence the presence of this new variable would necessarily add more information to the model, indicating that the original pair of explanatory variables was not additively separable. Finally, using the LR statistic as described in Section 3.4, the null hypothesis of additive separability between predictors in the primal non-parametric regression (the restricted model) can be examined against the alternative hypothesis (the unrestricted model).

Table 3 presents the results of statistical inference obtained from examining the null hypotheses. In total, we developed and estimated six different unrestricted models of the cost function and tested each one of them against the restricted cost function. The unrestricted models have one factor in common which distinguishes them from the restricted model, and that is the presence of the interactive predictor generated to examine the specific null hypothesis being tested. For the sake of brevity, none of the coefficient estimates of the unrestricted models are shown here. Instead, the estimated values of the residual deviance used in the calculation of the deviance value, and, in turn, the LR statistic test are reported. The computed values of deviance for all six models range from zero to 2.018, which are all less than the critical value of the chi-squared distribution (3.84) at the 0.05 level of significance with the appropriate (one) degree of freedom. We conclude the tests show that the interactive explanatory variable introduced in each of the six unrestricted models is not significantly different from zero, meaning that the new predictor variables add no new information to the initial non-parametric (restricted) cost function. Hence, we conclude that the additive separability assumption between predictors for cost estimation in Canadian CATV is valid.

Insert Table 3 Here

By design, our findings have both theoretical and policy relevance. From a theoretical perspective, input separability can be examined by studying how the marginal rate of technical substitution is affected by the changes in another input available in a third dimension. By failing to reject the null hypothesis on the inclusion of appropriate interaction variables, we conclude that any changes in the total number of basic subscribers, or in the amount of labor required by non-basic subscribers, might alter the span degree in the space of these inputs individually but do not affect the span degree of basic subscribers and labor shared in non-basic subscribers together. In turn, the test shows that the marginal rate of technical substitution between these two particular predictors does not respond to any changes imparted by the new variable, that is, the interactive independent variable. From a policy perspective, the ability to generate such results should be of interest to all regulators, not just those in Canadian CATV. For example, discovering that a change in demand for labor for non-basic CATV services attributable to changes in the

amount of labor needed for basic CATV subscribers has no significant effect on the total cost of operations leads to a better understanding of the relationship between inputs in this industry (i.e., that they are separable). Ultimately, understanding which production inputs are separable in a statistical sense can better inform the choice among regulatory regimes for that industry (for instance, a choice between rate of return vs. price cap regulation) that are more or less likely to influence input choices by the regulated firm.

Finally to properly assess the usefulness of the suggested methods, one must remain mindful of the following theoretical issues: (*i*) not all technologies are separable, and (*ii*) the concept of separability is most easily described in the context of continuously differentiable technologies (see Chambers, 1997, p. 42). We note that in the preceding sections we assumed that the mean response variable would be smooth enough to be twice differentiable, and this is obviously a crucial assumption for performing any non-parametric regression analysis.

*4.3 Discussion*

Our findings indicate that Canadian CATV during this period was characterized by separable inputs. At first glance, this result might seem surprising, given the interconnected nature of CATV provision at that time. However, considering the combination of the underlying technology of CATV service – highly capital-intensive networks – and the nature of the prevailing CRTC regulatory structure, upon further reflection our separability finding is perhaps not so startling after all. It is important to note as well that these general results concur with previous research about the nature of the provision of CATV services in Canada at that time. Ultimately, a better understanding about the prevalence of input separability, especially in network-based utilities, will serve to assist policy makers concerned about the nature of production processes in these often regulated industries, many of which share the capital-intensive characteristics of cable television systems.

## 5. Summary and Conclusions

The purpose of this research was to develop an example of the use of generalized additive models (GAMs) and show that the choice of non-parametric estimation technique need not imply an inability to test structural characteristics of the estimated model. We show that if the research goal is to estimate a cost function within a GAM non-parametric framework, one must proceed by evaluating the additive separability condition for the inputs. In this study, we have shown that labour and capital were additively separable in the production of cable television services in Canada through the early to mid 1990's. While this result that may seem counterintuitive at first given the network nature of CATV provision, other factors (particularly capital intensity) seem to have exerted more influence on technological decisions in Canadian CATV at that time than the particular structure of CATV regulation.

This study suggests several avenues for research, especially in other regulated industries. It would be interesting to test not only the existence of separability, but also substitutability of regulated inputs. Testing whether inputs are used more as complements or substitutes would have particular policy relevance for many industries, including CATV. This is the case, for example, because rate of return regulation can take different forms if returns are earned on an input (capital) that is complementary to, rather than a substitute for, other inputs.

Another area of future research would test separability across output categories. Applied to our particular example, we note that 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 installed for the provision of basic service. CATV firms were permitted to designate assets or declare what fractions of capital expenditures were used for basic service. In fact, our use of the CRTC data assumes that there was no bias introduced by the incentive to overstate the fraction of capital dedicated to the production of basic service. It would also be interesting to test whether basic and non-basic assets in this industry are indeed separable. Such results would certainly have implications for other regulated industries in which artificial reporting requirements create distinct but non-intuitive categories for production, either for inputs or for outputs.

## References

Averch, H., & L. Johnson. (1962). Behaviour of the firm under regulatory constraint. *American Economic Review*, 52(5): 115-121. http://links.jstor.org/sici?sici=0002-8282%28196212%2952%3A5%3C1052%3ABOTFUR%3E2.0.CO%3B2-G

Bowman, A. D., & A. Azzalini. (1997). *Applied smoothing techniques for data analysis: The kernel approach with s-plus illustration*, Oxford: Oxford University Press.

Breiman, L., & J. H. Friedman. (1985). Estimating optimal transformations for multiple regression and correlation. *Journal of American Statistics Association,* 80(391): 580-619. http://dx.doi.org/10.2307/2288473

Chambers, R. G. (1997). *Applied production analysis: The dual approach*, Cambridge: Cambridge University Press.

Chen, R., W. Hardle, O.B. Linton, & E. Sevarance-Lossin. (1996). *Nonparametric estimation of additive separable regression models: statistical theory and computational aspects of smoothing*, New York: Physica Verlag.

Eubank, R. L. (2000). Spline regression, in M. G. Schimek (ed.) *Smoothing and regression: approaches, computation, and application*, New York: John-Wiley and Sons.

Färe, R., & J. Logan. (1983). The rate-of-return regulated firm: Cost and production duality. *Bell Journal of Economics*, 14(2): 405-414. http://www.jstor.org/stable/3003642

Friedman, J. H., & W. Stuetzle. (1981). Projection pursuit regression. *Journal of American Statistics Association,* 76(376): 817-823. http://dx.doi.org/10.2307/2287576

Gozalo, P. L., & O. B. Linton. (2001). Testing additivity in generalized nonparametric regression models with estimated parameters. *Journal of Econometrics,* 104(1): 1-48. http://dx.doi.org/10.1016/S0304-4076(01)00049-5

Haghiri, M., J. F. Nolan, & K. C. Tran. (2004). Assessing the impact of economic liberalisation across countries: A comparison of dairy industry efficiency in Canada and the United States. *Applied Economics,* 36(11): 1233-1243. http://dx.doi.org/10.1080/0003684042000247406

Henderson, J., A. Elek, & C. Gibson. (1992). *A financial and regulatory analysis of the Canadian cable television industry to the year 2001*, Toronto: KPMG Peat, Marwick, Stevenson and Kellog.

Hastie, T. J., & R. J. Tibshirani. (1986). Generalized Additive Models. *Statistical Science*, 11(3): 297-318. http://dx.doi.org/10.1214/ss/1177013604

Hastie, T. J., & R. J. Tibshirani. (1990). *Generalized additive models*, U.K.: Chapman & Hall/CRC.

Ihaka, R., & R. Gentleman. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics,* 5(3): 299-314. http://dx.doi.org/10.2307/1390807

Kneip, A., & L. Simar. (1996). A general framework for frontier estimation with panel data. *Journal of Productivity Analysis,* 7(2-3): 187- 212. http://dx.doi.org/10.1007/BF00157041

Law, S. M. (1997). *Economic policy interactions:* Intellectual property rights and competition policy; exclusive licensing and rate regulation, PhD Dissertation, University of Toronto.

Law, S. M. (2002). The problem of market size for Canadian cable television regulation. *Applied Economics,* 34(1): 87-99. http://dx.doi.org/10.1080/00036840010025092

Law, S.M. (2008). Assessing evidence for the Averch-Johnson-Wellisz effect for regulated utilities, *ACEA Papers and Proceedings*, Atlantic Canada Economics Association. http://www.unb.ca/econ/acea/papers_proceedings.html

Law, S. M., & J. F. Nolan. (2002). Measuring the impact of regulation: A study of Canadian basic cable television. *Review of Industrial Organization,* 21(3): 231-249. http://dx.doi.org/10.1023/A:1020445804113

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,* 53 (4): 343-350. http://dx.doi.org/10.1090/S0002-9904-1947-08796-6

Leontief, W. W. (1947b). Introduction to a theory of the internal structure of functional relationships. *Econometrica,* 15(4): 361-373. http://dx.doi.org/10.2307/1905335

Linton, O. B., & J. P. Nielsen. (1995). A kernel method of estimating structured nonparametric regression based on marginal integration. *Biometrika,* 82(1): 93-100. http://dx.doi.org/10.1093/biomet/82.1.93

Nelson, R.A., & M.E. Wohar. (1983). Regulation, scale economies, and productivity in steam-electric generation. *International Economic Review,* 24(1): 57- 79. http://dx.doi.org/10.2307/2526115

Nelson, R.A., & M.E. Wohar. (1987). A reply to regulation, scale and productivity: A comment, *International Economic Review,* 28(2): 535-539. http://dx.doi.org/10.2307/2526742

Pagan, A., & A. Ullah. (1999). *Nonparametric econometrics*, Cambridge: Cambridge University Press.

Richmond, J. (2000). *Separability and specification tests*, Discussion Paper No.527, University of Essex.

Schimek, M. G., & 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.

Sono, M. (1961). The effect of price changes on the demand and supply of separable goods. *International Economic Review,* 2(3): 239-271. http://www.jstor.org/stable/2525430

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society,* Series B (Methodological) 36(2): 111-147. http://www.jstor.org.libproxy.mta.ca/stable/2984809

Wahba, G. (1990). *Spline models for observational data*, Philadelphia: CBMS-NSF Regional Conference Series SIAM (Society for Industrial and Applied Mathematics).

Wellisz, S.H. (1963). Regulation of natural gas pipeline companies: An economic analysis. *Journal of Political Economy,* 71(1): 30-43. http://dx.doi.org/10.1086/258732

Table 1. Statistical description of the variables

| Variable | Mean | S. D. | Min. | Max. |
|---|---|---|---|---|
| **Total Costs ($ 000)** | 6,975.3 | 16,533.7 | 30.6 | 222,821.3 |
| **Price of Labor Shared in Basic Subscribers ($)** | 40,388.3 | 16,275.1 | 487.6 | 163,872.6 |
| **Price of Labor Shared in Non-Basic Subscribers ($)** | 30,790.8 | 18,458.6 | 90.3 | 177,500.0 |
| **Price of Capital Shared in Basic Subscribers** | 0.3211 | 0.2604 | 0.1267 | 7.6618 |
| **Price of Capital Shared in Non-Basic Subscribers** | 0.6045 | 1.3354 | 0.1202 | 40.8860 |
| **Price of Material Shared in Basic Subscribers ($/km 000 per channel)** | 181.4 | 128.9 | 15.1 | 1,872.6 |
| **Price of Material Shared in Non-Basic Subscribers ($/km 000 per channel)** | 161.3 | 143.6 | 10.2 | 1,655.9 |
| **Number of Basic Subscribers** | 28,004.0 | 58,747.0 | 91.0 | 586,606.0 |
| **Number of Non-Basic Subscribers** | 19,581.0 | 51,744.0 | 14.0 | 573,420.0 |

Source: Sample data.

Table 2. Estimation results from GAM (restricted model)

| Predictors | Estimated DF | Estimated Rank | Non-parametric | |
|---|---|---|---|---|
| | | | F-value | p-value |
| Intercept | 1 | | | |
| Price of labor shared in basic subscribers | 1.000 | 1.000 | 3.302 | 0.0695 |
| Price of labor shared in non-basic subscribers | 1.747 | 9.000 | 2.2530 | 0.0171 |
| Price of capital shared in basic subscribers | 8.330 | 9.000 | 7.2560 | 0.0000 |
| Price of capital shared in non-basic subscribers | 7.806 | 9.000 | 11.1850 | 0.0000 |
| Price of material shared in basic subscribers | 8.068 | 9.000 | 14.6930 | 0.0000 |
| Price of material shared in non-basic subscribers | 7.106 | 9.000 | 11.0810 | 0.0000 |
| Number of basic subscribers | 8.195 | 9.000 | 1720.933 | 0.0000 |
| Number of non-basic subscribers | 1.000 | 1.000 | 98.394 | 0.0000 |
| Residual Deviance | 31.2258 | | Adjusted $R$-squared | 0.986 |
| Residual Degrees of Freedom | 996.7469 | | Number of observations | 1,041 |

Source: Sample data.

Table 3. Estimation results from GAM (unrestricted models)

| New Predictor | Residual Deviance | Residual DF | Computed Value of Deviance |
|---|---|---|---|
| **Model 1:** Basic subscribers & Price of labor shared in non-basic subscribers | 31.0132 | 993.9400 | 0.4252 |
| **Model 2:** Basic subscribers & Price of capital shared in non-basic subscribers | 30.0574 | 994.2724 | 0.3368 |
| **Model 3:** Basic subscribers & Price of material shared in non-basic subscribers | 30.2222 | 985.9330 | 2.0072 |
| **Model 4:** Non-basic subscribers & Price of labor shared in basic subscribers | 30.2170 | 984.7400 | 2.0180 |
| **Model 5:** Non-basic subscribers & Price of capital shared in basic subscribers | 31.2307 | 996.8204 | 0.0000 |
| **Model 6:** Non-basic subscribers & Price of material shared in basic subscribers | 31.2295 | 996.8045 | 0.0000 |

Source: Sample data.