

Nonparametric Demand Analysis

Draft - not for publication

Matthew Aaron Looney

08/22/2017

Abstract

Nonparametric econometric analysis has been a growing field of study over the past several decades. Many new techniques have been developed within a theoretical framework. However, despite the rapid growth of theoretical results, nonparametric applied research has lagged considerably. This paper employs a nonparametric regression analysis within the context of demand theory. Data on prices and quantities of three commodities (meats, dairy products and beans) from the 2006 Ecuadorian consumer expenditure survey will be evaluated to derive Marshallian price and income elasticities. The nonparametric results will be compared with standard parametric demand analysis tools such as the Log-Log demand model and the Almost Ideal Demand system to gauge the effectiveness of using nonparametric techniques to estimate demand elasticities.

1 Introduction

Emperical demand analysis has been dominated by the use of parametric functional form models since the appearance of the Working-Lesser Model in the 1940's (Working 1943) (Leser 1963). A 2015 paper by Clements and Gao (Clements and Gao 2015) provide a citation count of journal articles related to four popular parametric demand models (Linear Expenditure, Rotterdam, Translog and Almost Ideal Demand). The authors considered multiple time periods over the date range of 1974-2013. Their results clearly show an upward trajectory for all four parametric demand models with the Linear Expenditure and Almost Ideal Demand (AIDS) models having the highest citation counts, both in the multiple thousands of citations. This result suggests that parametric demand modeling is still very much a hotly researched area.

While parametric form models have clearly dominated the research landscape there are other empirical techniques available to evalaute consumer expenditure datasets. The goal of empirical demand analysis is to answer basic questions about consumer behavior and in many cases, use these answers to design and implement policy (government) or imporove some aspect of business design to increase market exposure/profits (private industry). Beyond these practical objectives Hal Varian (Varian 1982) suggests that, given a consumers expenditure dataset the demand analyst should ask four basic questions concerning the consumers behavior,

- (1) Consistency? Is the observerd data consistent with the utility maximizing model?
- (2) Structure? Is the observed data consistent with a utility function with some special structure?
- (3) Recoverability? How can the underlying utility function be recovered?

(4) Extrapolation? How can we forecast behavior in other circumstances?

With these goals in mind economists are constantly refining their tools to improve their estimates of consumer behavior.

Nonparametric economic analysis has been a growing field of study over the past several decades. Many new techniques have been developed within a theoretical framework. However, despite the rapid growth of theoretical results, nonparametric applied research has lagged considerably. This may be due in part to the advanced mathematical and statistical exposition presented in many nonparametric research papers. It is often difficult for economist, while well trained in advanced mathematical techniques, to fully grasp the significance and translate from purely theoretical results into applied research. Theoretical researchers working within the field of mathematics and statistics often fail to consider applied economic problems and applied economist are not exploring these newly developed theoretical techniques and refining the theory once they have touched real world data. In addition, until recently, nonparametric techniques were substantially more computationally intensive compared to their parametric counterparts. While the previous statement would seem to imply that nonparametric techniques have now become computationally less intensive, in reality it is our computer hardware and programming techniques which have improved to the point where nonparametric analysis has become a viable alternative, especially where parametric techniques fall short. These improvements in computer hardware have effectively opened an area of research which, until recently, had remained closed.

Regression analysis is the workhorse technique employed in econometrics. However, in linear regression it is assumed the regressors enter the conditional mean in a linear fashion and each regressor is independent of the other which is often a violation of the data under study. Even when we use nonlinear regression techniques we often still

assume we know the functional form for the data generating process (DGP).

The single largest potential issue with using a parametric form model to evaluate consumer demand is the prior imposition of functional form on the demand model. If a demand system is well represented by a functional form and the econometric model is correctly specified with theoretical assumptions met then a parametric estimator is both consistent and efficient. In fact if the previous criteria are met parametric regression is superior in just about every way. However, these requirements are excessively difficult to satisfy when used in the wild to gain insight about real world data and thus the parametric approach can be seriously flawed and worst, can lead the researcher to faulty conclusions which can have serious repercussions if used to implement policy.

Nonparametric modelling affords many advantages over their parametric counterparts. Primary among them is the nonparametric models ability to help us uncover a more accurate representation of the unknown function, conditioned on the actual data in hand. In section 2 I will explore the methods and models used to estimate a nonparametric regression. In nonparametric analysis a critical step in achieving a reliable kernel estimator is choosing the correct bandwidth estimator. Some time will be spent exploring several different bandwidth estimators which are prominent in the current literature. I will also explore the two most popular kernel regression methods in current use, Local-Constant Least Squares (LCLS) and Local-Linear Least Squares (LLLS). I will summarize the data and quickly review the parametric models being used for comparison purposes. Section 3 explores the results obtained from the nonparametric kernel regression and compares these results with the parametric form models used on the same dataset. Section 4 concludes and details direction for additional studies.

2 Models, Methods, and Data

In this paper we seek to explore the validity of using nonparametric techniques to uncover estimates of demand elasticities. To this end we use the Log-Log demand system developed by Working (1943) and Lesser (1963). The benefit of using this type of parametric form is the sheer simplicity of the model. This simplicity allows us to easily compare the effectiveness of our nonparametric econometric technique without getting bogged down in complicated estimation considerations or theory constraints. The only theory constraints available to the Log-Log demand model is homogeneity of degree zero in prices and income. However, even though this constraint is available to the Log-Log model we decide to forego all economic theory constraints to make comparability more consistent.

2.1 Log-Log model

Equation 1 shows the Log-Log model specification for the Marshallian own and cross prices elasticities (e_{ik}) and the income elasticities (e_i). Since we are unable to impose cross equation constraints our study will focus on the own price and income elasticities of this model. While the Hicksian elasticities are of more interest for policy and welfare studies, we focus only on the Marshallian estimates so a direct estimate comparison can be made between the parametric and nonparametric elasticity.

$$\log(q_i) = \alpha_i + e_i \log(x) + \sum_{k=1}^n e_{ik} \log(p_k) \quad (1)$$

It is well known that a correctly specified parametric model is preferred over nonparametric methods. However, we also know that correct model specification prior to study

is a near impossible task. It is important to test the parametric model to assess if the functional form is consistent with the DGP. The test we utilize in this study was developed by Hsiao, Li and Racine and is a test of Consistent Model Specification (CMS-test) (Hsiao et al. 2007).

2.2 Consistent Model Specification Test

The CMS-test is a kernel based evaluation which tests for correct specification of the parametric form model. A short review of the details of the test are given below.

The null hypothesis is given by,

$$H_0 : P[E(y_i|x_i) = m(x_i, \beta)] = 1 \text{ for some } \beta \in \mathcal{B}$$

where,

$m(\cdot, \cdot)$ is a known function with β being a $p \times 1$ vector of unknown parameters,

\mathcal{B} is a compact subset in \mathbb{R}^p .

The alternative hypothesis is given by,

$$H_1 : P[E(y_i|x_i) = m(x_i, \beta)] < 1 \text{ for all } \beta \in \mathcal{B}$$

The test statistic was originally developed by Fan and Li (1996) and Zheng (1996).

The test statistic is given by,

$$I_n = \frac{1}{n^2} \sum_i \sum_{j \neq i} \hat{u}_i \hat{u}_j K_{\gamma_{ij}}$$

where,

$$K_{\gamma_{ij}} = W_{h,ij} L_{\lambda,ij}(\gamma = (h, \lambda)),$$

$\hat{u}_i = y_i - m(x_i, \hat{\beta})$ is the parametric null model's residual,

β is a $\sqrt{(n)}$ -consistent estimator of β under H_0 .

Hsiao et al. (2007) advocate for the use of cross-validation methods for selecting the kernel smoothing parameter vectors, which is the approach we took in this analysis. Under assumptions (A1)-(A3) of their paper we can obtain the cross-validation based (CV-based) test with the new test statistic \hat{J}_n :

$$n(\hat{h}_1, \dots, \hat{h}_q)^{1/2} \hat{I}_n \rightarrow N(0, \Omega) \text{ in distribution under } H_0,$$

where,

$$\Omega = 2E[\sigma^4(x_i)f(x_i)] [\int W^2(v)dv]$$

A consistent estimator of Ω is given by,

$$\hat{\Omega} = n^{-2} 2(\hat{h}_1, \dots, \hat{h}_q) \sum_i \sum_{j \neq i} \hat{u}_i^2 \hat{u}_j^2 W_{\hat{h},ij}^2 L_{\hat{\lambda},ij}^2$$

Which gives the CV-based test statistic as,

$$\hat{J}_n = n(\hat{h}_1, \dots, \hat{h}_q)^{1/2} \hat{I}_n \hat{\Omega}^{-1/2} \rightarrow N(0, 1) \text{ in distribution under } H_0.$$

According to the authors, it can be easily shown that the \hat{J}_n test statistic diverges to $+\infty$ if H_0 is false.

2.3 Bandwidth Selection models

The importance of proper bandwidth selection in nonparametric analysis is reviewed here. However, for a more detailed discussion of the methods and theory used in this paper see the excellent work by the following authors: Cheng et al.(1997), Hall et al.(2007), Hurvich et al.(1998), Li and Racine(2007), Racine and Li(2004), and Li and Racine(2004).

The most popular methods for bandwidth selection in applied research are data driven,

and of the data driven methods Least-Squares Cross Validation (LSCV) and expected Kullback-Leibler cross-validation also known in the literature as Akaike Information Criterion Cross Validation (aic-CV), are the two used in this paper to evaluate the validity of using nonparametric techniques to uncover demand elasticities.

2.3.1 Least-Squares Cross Validation (LSCV)

While there are many methods to determine an optimal bandwidth, the most commonly used in the current literature is the cross-validation method of LSCV which is defined as:

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

where,

$\hat{m}_{-1} - (x_i)$ is the leave-one-out estimator defined by Li and Racine(2004) as,

$$\hat{m}_{-1} - (x_i) = \frac{\sum_{\substack{j=1 \\ j \neq i}}^n y_h(x_i, x_j)}{\sum_{\substack{j=1 \\ j \neq i}}^n k_h(x_i, x_j)} \quad (3)$$

An examination of equation 2 reveals that our algorithm will require n^2 calculations. Several authors have suggested a derivation of the LSCV that side-steps this issue; we have decided to use the classic formula as defined in the original paper by Li and Racine(2004). It should be noted however, as the data sample grows, the calculations become computationally burdensome.

Gaining in popularity is a cross-validation bandwidth selection method developed by

Hurvich et al.(1998) called Akaike Information Criterion Cross Validation (aic-CV). This method is similar to Kullback-Leibler information criterion method which is the foundation for density estimation using likelihood cross-validation. However, the true benefit of using the aic-CV method is the property of reduced variability the tendency for this method to undersmooth.

2.3.2 Akaike Information Criterion Cross Validation (aic-CV)

The aic-CV method is defined as:

$$AIC_c(h) = \ln(\hat{\sigma}^2) + \frac{1 + p/n}{1 - (p + 2)/n} = \ln(\hat{\sigma}^2) + 1 + \frac{2(p + 1)}{n - p - 2} \quad (4)$$

where,

AIC_C is the bias corrected version,

$\hat{\sigma}^2$ is the estimated error variance,

p is the number of regression (or autoregressive) paramaters in the model.

which gives the following formula for smoothing paramter selction:

$$AIC_c(h) = \ln(\hat{\sigma}^2) + \frac{1 + tr(H)/n}{1 - \{tr(H) + 2\}/n} = \ln(\hat{\sigma}^2) + 1 + \frac{2(tr(H) + 1)}{n - tr(H) - 2} \quad (5)$$

where,

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

and H is called the hat matrix or smoother matrix defined by:

$$H = \begin{bmatrix} \frac{K(x_1, x_1)}{\hat{f}(x_1)} & \frac{K(x_1, x_1)}{\hat{f}(x_1)} & \vdots & \frac{K(x_1, x_n)}{\hat{f}(x_1)} \\ \frac{K(x_2, x_1)}{\hat{f}(x_2)} & \frac{K(x_2, x_2)}{\hat{f}(x_2)} & \vdots & \frac{K(x_2, x_n)}{\hat{f}(x_2)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{K(x_n, x_1)}{\hat{f}(x_n)} & \frac{K(x_n, x_2)}{\hat{f}(x_n)} & \dots & \frac{K(x_n, x_n)}{\hat{f}(x_n)} \end{bmatrix}$$

The trace of the H matrix should be thought of as a measure of the number of parameters in the nonparametric model (Hastie and Tibshirani 1993).

2.4 Nonparametric Regression models

While the method of bandwidth selection is critical, the end game of this analysis is to use the bandwidth estimator in a regression analysis to evaluate a comparison between elasticities derived from a parametric model estimated under Least Squares versus elasticities derived from nonparametric regression methods. With this goal in mind we now explore nonparametric regression methods.

2.4.1 Local-Constant Least Squares (LCLS)

The LCLS regression was the first nonparametric regression model developed and is sometimes referred to as the Nadaraya–Watson estimator after Nadaraya (1964) and Watson (1964). This method minimizes the kernel weighted least-squares regression of y_i on a constant c . If we assume some function g s.t. $y_i = g(x_i) + \epsilon_i$ then our objective with LCLS is to approximate $g(\cdot)$ with a constant c . More formally the LCLS can be defined as:

$$LCLS(y; \mathbf{x}) = \frac{1}{nh_y|\mathbf{h}|} \sum_{i=1}^n k\left(\frac{y_i - y}{h_y}\right) K_h(\mathbf{x}_i, \mathbf{x}) \quad (6)$$

where,

$K_h(\mathbf{x}_i, \mathbf{x})$ is the product kernel.

$$a = \arg \min \{\hat{m}(\mathbf{x})\} = \frac{\sum_{i=1}^n K_h(\mathbf{x}_i, \mathbf{x}) y_i}{\sum_{i=1}^n K_h(\mathbf{x}_i, \mathbf{x})} \quad (7)$$

where a can be thought of as our constant c which approximates $g(\cdot)$.

The above estimator can be represented in matrix form as:

$$\hat{m}(\mathbf{x}) = (\mathbf{i}' K(\mathbf{x}) \mathbf{i})^{-1} \mathbf{i}' K(\mathbf{x}) y$$

where,

\mathbf{i} is the identity matrix of dimension $n \times 1$,

and $K(\mathbf{x})$ is the same product kernel from equation 6,

$$K(\mathbf{x}) = \begin{bmatrix} K_h(\mathbf{x}_1, \mathbf{x}) & 0 & 0 & 0 \\ 0 & K_h(\mathbf{x}_2, \mathbf{x}) & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & K_h(\mathbf{x}_n, \mathbf{x}) \end{bmatrix}$$

One of the problems with the LCLS estimator is that it is bias. This may not be of concern depending on the circumstances but it should be noted.

2.4.2 Local-Linear Least Squares (LLLS)

The LLLS estimator has been growing in popularity as an alternative to the LCLS method. In addition to increasing use in applied research the LLLS estimator has been dominated the theoretical literature.

Conceptually, the LLLS estimator is similar to the LCLS with the exception that instead of fitting a constant to our left hand side variable we now consider fitting a locally defined line segment. This allows us to more accurately model the data and is especially useful when we have outliers and nonlinear relationships in our data.

We can think of our local-linear fitting as a local Taylor Series expansion at some point \mathbf{x} . In other words, for the relationship $y = m(x) + u$, we have paired data on $(y_1, x_1), (y_2, x_2), \dots, (y_n, x_n)$. This allows us to take a local Taylor approximation for each of the points \mathbf{x} . It now become clear, with locally fitting with kernel weighting, problematic data (outliers, nonlinearity) can be better modeled by LLLS.

The LLLS estimator as developed and defined by Li and Racine (2004) is given as follows:

We begin with a Taylor expansion of \mathbf{x} for every observation i as,

$$y_i = m(x_i) + u_i \approx m(x) + (x_i - x)\beta(x) + u_i$$

where,

$(x_i - x)$ is a $1 \times q$ vector and $\beta(x)$ is the gradient vector of dimension q .

then,

$$y_i = \alpha + (x_i - x)b + u_i = \begin{bmatrix} 1 & (x_i - x) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + u_i$$

where,

$m(x)$ and $\beta(x)$ are paramaters.

Our minimization problem then becomes,

$$\arg \min_{a,b} \sum_{i=1}^n [y_i - a - (x_i - x)b]^2 K_h(x_i, x)$$

and in matrix notation,

$$\arg \min_{\delta} [y - X\delta]' K(x) [y - X\delta]$$

where,

$$\delta = (a, b)',$$

X is a $n \times (q + 1)$ matrix with the i^{th} row equal $\begin{bmatrix} 1 & (x_i - x) \end{bmatrix}$,

$K(x)$ is $n \times n$ diagonal matrix with the i^{th} element equal to $K_h(x_i, x)$.

If $K(x)$ is an identity matrix then we simply have the OLS estimator.

If $K(x)$ is equal to $\Omega = vcov(\theta)$ then the above relationship reduces to the generalized least squares.

When we minimize the objective function $\hat{\delta}(x)$ we get:

$$\begin{aligned}
\hat{\delta}(x) &= \begin{bmatrix} \hat{m}(x) \\ \hat{\beta}(x) \end{bmatrix} \\
&= \left[\sum_{i=1}^n K_h(x_i, x) \begin{pmatrix} 1 \\ x_i - x \end{pmatrix} \begin{pmatrix} 1 & (x_i - x) \end{pmatrix} \right]^{-1} \sum_{i=1}^n K_h(x_i, x) \begin{pmatrix} 1 \\ x_i - x \end{pmatrix} y_i \\
&= [X'K(x)X]^{-1}X'K(x)y
\end{aligned}$$

where the final equation in matrix form gives our LLS estimators.

The bias of the LLS estimator is less complicated to work with than the LCLS and deals better with nonlinear and irregular data. For these reasons many researchers employ the LLS estimator over the LCLS. However, there are situations where the LCLS estimator will be preferred so an analysis of both methods within the context of the problem is always wise.

Table 1: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
obs	2,066	1,033.500	596.547	1	2,066
p1	2,066	2.610	0.820	0.365	11.351
p2	2,066	1.231	1.510	0.011	23.736
p3	2,066	1.218	0.444	0.010	4.386
q1	2,066	4.498	3.011	0.001	38.574
q2	2,066	6.244	5.635	0.001	63.762
q3	2,066	0.943	0.929	0.0004	14.074
p1q1	2,066	11.395	7.883	0.003	98.101
p2q2	2,066	4.619	4.215	0.001	51.451
p3q3	2,066	1.073	1.088	0.0005	15.250
X	2,066	17.087	10.679	0.199	128.602
w1	2,066	0.657	0.168	0.001	0.984
w2	2,066	0.269	0.159	0.0001	0.945
w3	2,066	0.074	0.070	0.00003	0.884

good 1 = meats
 good 2 = dairy
 good 3 = beans

Table 2: Parametric - Double Log Demand Model

	meats	dairy	pulses	income_elasticity	R-squared
meats_lnq1	-0.9073***	-0.07323***	-0.003844	1.164***	0.74226
dairy_lnq2	-0.05847	-0.8216***	-0.07056*	1.004***	0.64494
pulses_lnq3	-0.3804***	-0.05684*	-0.4394***	0.5658***	0.14743

***Significant at the 1 percent level,
 **Significant at the 5 percent level,
 *Significant at the 10 percent level.

Table 3: Nonparametric Regression using Gaussian Kernel

	meats	dairy	pulses	income_elasticity	R-squared
meats_lnq1	-0.9674***	-0.06514*	-0.04756	1.124***	0.88725
dairy_lnq2	0.002597	-0.843***	-0.0752***	0.9875***	0.71745
pulses_lnq3	-0.3402***	-0.05632**	-0.7777***	0.4965***	0.27586

***Significant at the 1 percent level,
 **Significant at the 5 percent level,
 *Significant at the 10 percent level.

Table 4: Full AIDS - Marshallian

	meats	dairy	pulses	income_elasticity	R-squared
meats_lnq1	-0.9813***	-0.06048***	-0.03157**	1.073***	0.067727
dairy_lnq2	-0.06908***	-0.8626***	-0.03394***	0.9657***	0.052564
pulses_lnq3	0.1063	0.006219	-0.5657***	0.4532***	0.13459

***Significant at the 1 percent level,

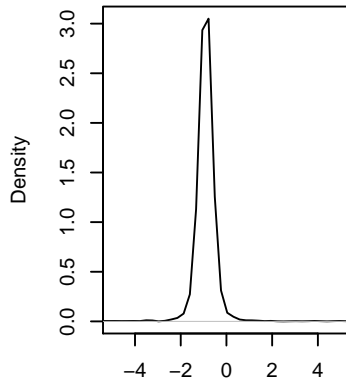
**Significant at the 5 percent level,

*Significant at the 10 percent level.

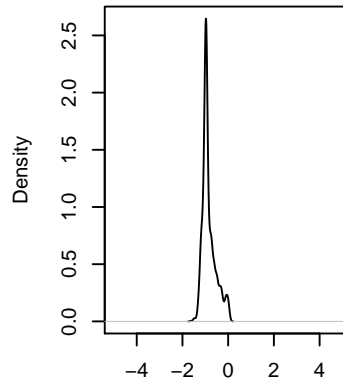
2.5 Methods

2.6 Data

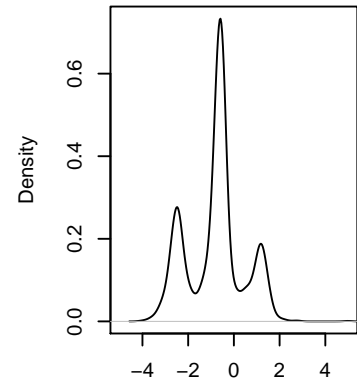
3 Results and Discussion



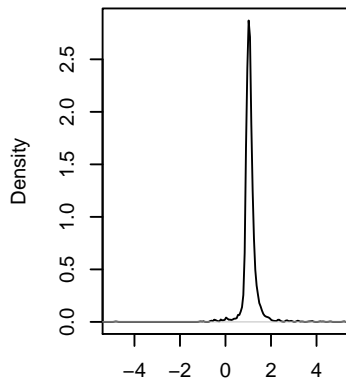
N = 2066 Bandwidth = 0.04266
Meats – own price elasticity



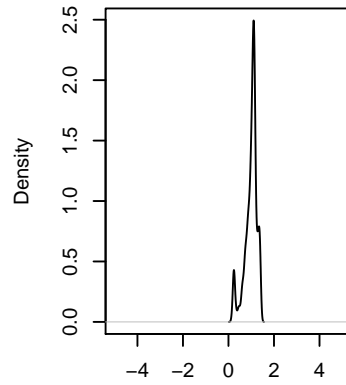
N = 2066 Bandwidth = 0.0456
Dairy – own price elasticity



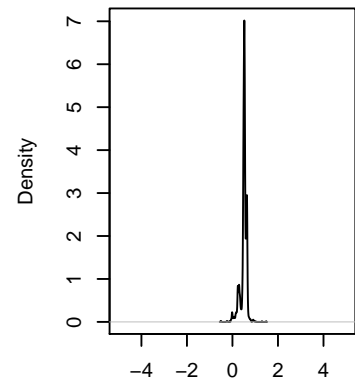
N = 2066 Bandwidth = 0.1788
Beans – own price elasticity



N = 2066 Bandwidth = 0.02809
Meats – income elasticity



N = 2066 Bandwidth = 0.04164
Dairy – income elasticity



N = 2066 Bandwidth = 0.01302
Beans – income elasticity

4 Conclusions

References

- Cheng, M.-Y., Hall, P., & Titterton, D. (1997). On the shrinkage of local linear curve estimators. *Statistics and Computing*, 7(1), 11–17.
- Clements, K. W., & Gao, G. (2015). The Rotterdam demand model half a century on. *Economic Modelling*, 49 IS -, 91–103.
- Fan, Y., & Li, Q. (1996). Consistent model specification tests: omitted variables and semiparametric functional forms. *Econometrica*, 64(4), 865.
- Hall, P., Li, Q., & Racine, J. S. (2007). Nonparametric Estimation of Regression Functions in the Presence of Irrelevant Regressors. *The Review of Economics and Statistics*, 89(4), 784–789.
- Hastie, T., & Tibshirani, R. (1993). Varying-Coefficient Models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 55(4), 757–796.
- Hsiao, C., Li, Q., & Racine, J. S. (2007). A consistent model specification test with mixed discrete and continuous data. *Journal of Econometrics*, 140(2), 802–826.
- Hurvich, C. M., Simonoff, J. S., & Tsai, C.-L. (1998). Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 60(2), 271–293.
- John Xu Zheng. (1996). A consistent test of functional form via nonparametric estimation techniques. *Journal of Econometrics*, 75(2), 263–289.
- Leser, C. E. V. (1963). Forms of Engel Functions. *Econometrica*, 31(4), 694–703.
- Li, Q., & Racine, J. (2004). Cross-Validated Local Linear Nonparametric Regression.

Statistica Sinica, 14(2), 485–512.

Li, Q., & Racine, J. S. (2007). *Nonparametric econometrics: theory and practice*. Princeton University Press.

Nadaraya, E. A. (1964). On Estimating Regression. *Theory Probab. Appl.*, 9(1), 141–142.

Racine, J., & Li, Q. (2004). Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics*, 119(1), 99–130.

Varian, H. R. (1982). Non-parametric methods in demand analysis. *Economics Letters*, 9(1), 23–29.

Watson, G. S. (1964). Smooth Regression Analysis. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 26(4), 359–372.

Working, H. (1943). Statistical Laws of Family Expenditure. *Journal of the American Statistical Association*, 38(221), 43–56.