A Nonparametric Demand Analysis

Recovering elasticities with nonparametric regression

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 quantites 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 guage the effectivness 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 techinques, to fully grasp the significance and translate from purley 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 techinques 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 paramteric counterparts. While the previous statement would seem to imply that nonparametric tecniques have now become computationally less intesive, in reality it is our computer hardware and programing techniques which have improved to the point where nonparametric analysis has become a viable alternative, especially where parametric techniques fall short. These improvemnts in computer hardware have effectivly 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

3

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

The single largest potiential 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 effiecient. In fact if the previous criteria are met parametric regression is superior in just about every way. However, these requirements are excessivly difficult to satisfy when used in the wild to gain insight about real world data and thus the parametric approach can be seroiusly 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 amoung 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 acheiving a reliable kernel estimator is choosing the correct bandwidth estimator. Some time will be spend exploring several different bandwidth estimators which are prominant 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 effectivness 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 Demand 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.

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

It is well known that a correctly specified parametric model is prefered 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 untilize in this study was developed by Hasiao, 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,

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

where,

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

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

The alternative hypothesis is given by,

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

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

The test statistic is given by,

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

where,

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

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

6

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 paramater vectors, which is the approach we took in this analysis. Under assumptions (A1)-(A3) of thier paper we can obtain the cross-validation based (CV-based) test with the new test statistic $\hat{J}_n$:

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

where,

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

A consistent estimator of Ω is given by,

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

Which gives the CV-based test statistic as,

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

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 bandwith 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 know in the literature as Akaike Information Criterion Cross Validation (aic-CV), are the two used in this paper to evaluate the valididty 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:

$$LSCV(h) = \sum_{i=1}^{n} \left[y_i - \hat{m}_{-1} - (\mathrm{x}_i) \right]^2 \tag{2}$$

where,

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

$$\hat{m}_{-1} - (\mathrm{x}_i) = \frac{\sum\limits_{\substack{j \neq i \\ j=1}}^{n} y_h(\mathrm{x}_i, \mathrm{x}_j)}{\sum\limits_{\substack{j \neq i \\ j=1}}^{n} k_h(\mathrm{x}_i, \mathrm{x}_j)} \tag{3}$$

An examination of equation 2 reveals that our algorithm will require n^2 calculations. Several authors have suggested a derevation of the LSCV that side-steps this issue; we have decided to use the classic formuala 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 tendancy for this method to undersmooth.

2.3.2 Akaike Information Criterion Cross Validation (aic-CV)

The aic-CV method is defined as:

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

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:

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

where,

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

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

$$H = \begin{bmatrix} \frac{K(\mathrm{x}_1, \mathrm{x}_1)}{\hat{f}(\mathrm{x}_1)} & \frac{K(\mathrm{x}_1, \mathrm{x}_1)}{\hat{f}(\mathrm{x}_1)} & \vdots & \frac{K(\mathrm{x}_1, \mathrm{x}_n)}{\hat{f}(\mathrm{x}_1)} \\ \frac{K(\mathrm{x}_2, \mathrm{x}_1)}{\hat{f}(\mathrm{x}_2)} & \frac{K(\mathrm{x}_2, \mathrm{x}_2)}{\hat{f}(\mathrm{x}_2)} & \vdots & \frac{K(\mathrm{x}_2, \mathrm{x}_n)}{\hat{f}(\mathrm{x}_2)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{K(\mathrm{x}_n, \mathrm{x}_1)}{\hat{f}(\mathrm{x}_n)} & \frac{K(\mathrm{x}_n, \mathrm{x}_2)}{\hat{f}(\mathrm{x}_n)} & \dots & \frac{K(\mathrm{x}_n, \mathrm{x}_n)}{\hat{f}(\mathrm{x}_n)} \end{bmatrix}$$

The trace of the H matrix should be thought of as a measure of the number of paramaters 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 paramertic model estimated under Least Squares versus elasticities derived from nonparamtertic regression methods. With this goal in mind we now explore nonparamteric regression methods.

2.4.1 Local-Constant Least Squares (LCLS)

The LCLS regression was the first nonparametric regression model developed and is sometimes refered to as the Nadaraya–Watson estimator after Nadaraya (1964) and Watson (1964). This method minimizes the kernel weighted least-squres 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:

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

where,

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

$$a = \arg\min\{\hat{m}(\mathrm{x})\} = \frac{\sum\limits_{i=1}^{n} K_h(\mathrm{x}_i, \mathrm{x}) y_i}{\sum\limits_{i=1}^{n} K_h(\mathrm{x}_i, \mathrm{x})} \tag{7}$$

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

The above estimator can be represented in matrix form as:

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

where,

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

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

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

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 accuratly 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 $(y_1, x_1), (y_2, x_2), ..., (y_n, x_n)$. 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 betetr 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 ith row equal $\begin{bmatrix} 1 & (x_i - x) \end{bmatrix}$,

$K(x)$ is $n \times n$ diagonal matrix with the ith 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:

$$\hat{\delta}(x) = \begin{bmatrix} \hat{m}(x) \\ \hat{\beta}(x) \end{bmatrix}$$

$$= \left[\sum_{i=1}^{n} K_h(x_i, x) \begin{pmatrix} 1 \\ x_i - x \end{pmatrix} \begin{pmatrix} 1 & (x_i - x) \end{pmatrix} \right]^{-1} \sum_{i=1}^{n} K_h(x_i, x) \begin{pmatrix} 1 \\ x_i - x \end{pmatrix} y_i$$

$$= [X'K(x)X]^{-1} X'K(x)y$$

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

The bias of the LLLS 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 LLLS estimator over the LCLS. However, there are situations where the LCLS estimator will be prefered so an analysis of both methods within the context of the problem is always wise.

With parametric regression analysis it is customary to evaluate the overall significance of the regression. In nonparametric regression we employ a analogous test refered to as Kernel Regression Significance Test as developed by Racine et al. (2006) and Racine (1997).

2.5 Kernel Regression Significance Test

The Kernel Regression Significance Test developed by Racine et al. (2006) and Racine (1997) is applicable to mixed nonparametric regressions, that is, the test will work with a combination of discrete and continuous regressors.

The null hypothesis and the test statistic are as follows:

$H_0 : E(y|x, z) = E(y|x)$ almost everywhere (a.e.)

where,

z denotes catagorical or continuous variables,

x denotes the remaining explanatory variables,

y denotes the dependent variable.

Then the test statistic given by Racine et al. (2006):

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^{n} \sum_{l=1}^{c-1} \left[\hat{m}(x_i, z_i = l) - \hat{m}(x_i z_i = 0) \right]^2$$

where,

$$\hat{m}(x_i z_i = l) = \frac{\sum\limits_{j=1}^{n} y_i W_{\hat{h},ij} L_{\hat{\lambda},ij} l_{zj,z=l,\hat{\lambda}z}}{\sum\limits_{j=1}^{2} W_{\hat{h},ij} L_{\hat{\lambda},ij} l_{zj,z=l,\hat{\lambda}z}}$$

$\hat{I}_n \to 0$ in probability under H_0.

2.6 Methods

The goal of this study is to investigate the effectivness of using nonparametric regression analysis to recover demand elasticities for a group of three commodities; meats, dairy, and beans. To facilitate the comparison we will estimate equation by equation Ordinary Least Squares regression on the Log-Log demand model (equation 1) developed by Working and Lesser. The reason for choosing this model is the sheer simplicity of the functional form which allows us to easily estimate the model using OLS and nonparametric regression techinques without the concern of complicated estimation procedures or elaborate constraint requirements. This simplicity will allow us to

compare the two models as closly as possible to guage how effect the nonparametric technique is at recovering the elasticities.

Having assumed the Log-Log demand model functional form we use the Consistent Model Specification Test detailed in section 2.2 to assertain how well the Log-Log functional form does at recovering the true DGP of our date set. An impropper specification of functional form will lead to incorrect paramater estimates. The test is run three times, once for each commodity according to equation 1. In all three cases we preform the test by the following configuration: The distribution was estimated using iid bootstrapping methods with 1000 replications. A pivot table was used so that the distribution approched $N(0,1)$. A random seed was used in the analysis to enable reproducibility and is embedded within the code (available upon request).

The Log-Log demand model is estimated equation by equation using standard OLS regression. The only theory constraint available to the model is homogeneity, however, no constraints were imposed on the data in this study. There was no correction for heterogeneity and all variables are assumed to be exogenous, a faulty assumption in the current context, but an assumption that is ignorable given the goal of this study.

The nonparametric regression requires the selection of a bandwidth estimation technique. In this study we consider two methods of bandwidth estimation, Least-Squares Cross Validation (LSCV), and Akaike Information Criterion Cross Validation (aic-CV). Within both of these bandwidth estimators there are numerious variations available to fine tune the optimal bandwidth selction. Aside from the difference of base technique, ie. LSCV and aic-CV the configuration is identical between the two methods and follows as given: We use a Fixed bandwidth selection method. We ued a gaussian kernel for all studies, althogh the epanechnikov kernel was evaluated and rejected for lack of preformance. The kernel oder was 2 in all studies. We did not employ any

16

bandwidth scalling in any of these studies.

The nonparametric regression techniques used in this study are the Local-Constant Least Squares and the Local-Linear Least Squares. With both of these tehcniques we estimated the residuals of the regression and the gradients, which are of interest when comparing the elasticities obtained by the parametric regression versus the nonparametric regression.

A test of the overall regression significance of the nonparametric regression is preformed using the Kernel Regression Significance Test detailed in section 2.5. The test follows the algorithms detailed in Racine et al. (2006) with the specifics as follows: iid bootstrapping with 1000 replications was preformed. Bootstrap method type 1 was utilized. A pivitol statistic was calculated by dividing the calculated gradients by their respective standard errors. A random seed was embedded within the code to ensure reprodicibility.

2.7 Data

The data were obtained from the 2006 Ecuadorian consumer expenditure survey and consist of three commidties, meats, dairy, and beans. For all threde commidties we have price and quantity data for 2066 observations. For the Log-Log model the natural logarithm for quantities and prices was taken and passed to the model. For consistancy the nonparametric regression uses the same logarithmic values. See table 1 (below) for summary statistics.

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

3 Results and Discussion

The primary focus of this paper is to access the valididyt of using nonparametric regression methods to recover demand elasticities from a consumer consumption data set. In traditional demand analysis a paramatric function form model is often picked for its simplicity of estimation or tractability or both. However, the validity of the paramaters estimated form a misspecified demand model are called into question when the functional form is assumed and assumed incorrectly. Testing methods have been developed to assertain the validity of using a parametric form to uncover the underlying DGP. The test we employed in this analysis is the Consistent Model Specification Test. In all three cases the Null of correct specification is rejected at the 0.1% level or greater. This indicates that the Log-Log specification does a poor job at recovering the true DGP for our data set.

Table 2: Parametric - Double Log Demand Model

	meats	dairy	beans	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
beans_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	beans	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
beans_lnq3	-0.3402***	-0.05632**	-0.7777***	0.4965***	0.27586

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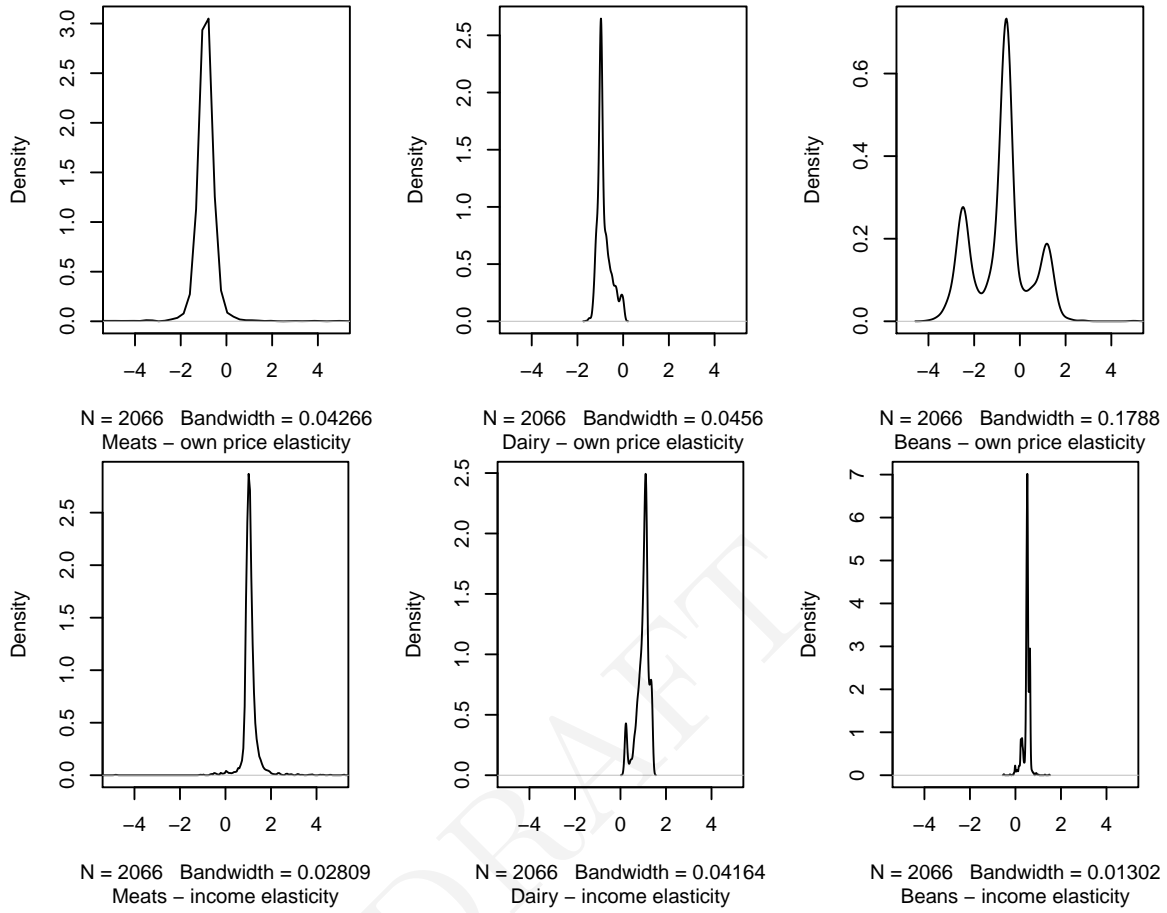*Significant at the 10 percent level.

Table 2 shows the results from the parametric OLS regression of the Log-Log demand model. When we examine the R-squared values for the three equaitons we notice that equation 1 (meats) does have a relativly high R-squared vaule which seems to indicate a resonable fit for this model. However, when we move to evauluate equation 2 (dairy) and equation 3 (beans) we notice progressivly decreasing R-squared values which idicates, especially for equation 3, that our OLS regression does not fit the data well enough to be confident that our paramater estimates are reliable.

Table 3 shows the results from the nonparametric regression. When we evaluate the R-squared values of the nonparamteric regression we notice a systematic increase in value. All three nonparametric regressions give higer R-squared values over the parametric regression models. This indicates an improvement in model fit using nonparametric techniques.

Since the parametric form model we have choosen to work with only permits within equation constrinats it only makes sense to make comparissons of own price and income elasticities. When we compare the own price elasticities of the parametric regression to the elasticities obtained from the nonparamteric regression we notice the nonparametric own price elasticities are all greater in magnitude. When we look at the income elasticities between the two regressions the nonparamteric model has lower values. However, the numeric values are all very close and the signs are all in agreement for the elasticities that are significant.

Due to the mechanism of estimation of the nonparametric regression a one for one paramater estimate is complicated by the fact that nonparametric regressions make a full range of estimates as opposed to paramaertic regressions which make point estimates of elasticities. This means we can get a better sense of how the elasticities are distributed across the data by plotting the gradiants, in our case the actuall elasticity estimates, and viewing the distribution of the elasticities across the entire sample range.

The plots below show the distribution for the own price and income elasticities. It becomes clear that we have relativly tight clustering of own price elasticities for meats and dairy but we an unusual distribution of own price elasticity on beans. We have a primary peak with two smaller peaks of elasticities to either side of the primary peak. With the income elasticities we see relativly tight clustering for all three commidities.

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

The Almost Ideal Demand System (AIDS) developed by Deaton and Muellbauer (1980), (Deaton 1980) has been one of the most highly used demand system specifications since its introduction in the 1980's. Monte Carlo studies have shown the AIDS specification to be effective at recovering the "true"" elasticities of demand under a variety of situations (Barnett and Seck 2008). For this reason, we will use the AIDS model as a reality check against our Log-Log demand model and our nonparametric regression model. We will compare the elasticities estimated between all three model specifications to gain insight about how effective our nonparametric regression isat recovering demand elasticities.

The Full AIDS model takes the following functional form:

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left(\frac{X}{P}\right) + \mu_i$$

where,

$lnP = $ a price index $= \alpha_0 + \sum\limits_{j} \alpha_j p_j + \frac{1}{2}\sum\limits_{j}\sum\limits_{i} \gamma_{ij} \ln p_i \ln p_j$

It should be readily apparent that the Full AIDS model can not be estimated using strictly linear estimation techniques. The non-linear AIDS model has been estimated by full information maximum likelihood (FIML) and by non-linear SUR methods, however, these methods suffer from difficulties with convergence. Several authors have shown that an iterated linear estimator can be consistently and asymptotically efficiently estimated by the Iterated Linear Least Squares Estimator (ILLE) developed by Blundell and Robin (1999). It is this iterated estimation method that is employed in this analysis of the Full AIDS model.

Although the ILLE method is computationally efficient, the algorithm converges more quickly if starting values are passed on the first iteration. In this study, the initial parameter values were obtained by first estimating the LA/AIDS model using Stone's price index, then passing the estimated parameter values to the ILLE algorithm as the initial starting values. Economic theory restrictions were inposed on the AIDS model, homogeneity and symmetry.

The Marshallian own and cross price elasticities are given by the following formula:

$$\varepsilon_{ij} = -\delta_{ij} + \frac{\gamma_{ij}}{w_i} - \beta_i \frac{w_j}{w_i} - \frac{\beta_i}{w_i} \sum_{k} \gamma_{kj} \ln p_k \tag{8}$$

The income elasticity is:

$$\eta_i = 1 + \frac{\beta_i}{w_i} \tag{9}$$

The standard errors of the elasticities are estimated using the Delta Method.

Table 4 shows the Marshallian elasticities obtained from the AIDS model. When we compare these results to our parametric regression and nonparametric regression we see the nonparametric version obtains elasticities closer in values to the AIDS model estimates. We see this trend in both the own price and income elasticities of the nonparametric model and the AIDS model. This result seems to confirm the nonparametric regression does a better job at recovering the demand elasticities in this specific case.

Table 4: Full AIDS - Marshallian

	meats	dairy	beans	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
beans_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.

4   Conclusions

We have explored the fesability of using nonparametric regresision analysis to recover the own price and income elasticities from a consumer expenditure data set. Using the nonparametric techniques we were able to obtain own price and cross price elasticities that were in agreements with those obtained using the Log-Log demand model, a paramteric form model. We confirmed these results by estimating the non-linear AIDS

model which also agreed with the elasticities we obtained by using nonparamtetric regression.

When parametric model specification is taken as known and this assumption is correct, the parametric form regression produces both consistent and efficient estimates. However, we can run into problems when the assumed model does not represent the underlysing data generating process. As a result, the paramater estimates obtained by the model will be faulty and could lead to wrong conclusions which has the potiential to neagtivly effect poilicy.

There are trade-offs with employing nonparametric techniques. When we estimate paramaters using unconstrained nonparametric regression we weaken the conclusions we can draw about commidity substitution and complementarity. In fact, this weakening of economic conclusions can be a pervasive problem when estimating using unconstrained nonparametric techniques. Work has begun to allow the incorporation of economic theory constraints into nonparamteric methods, if so desired. An additional issue with using nonparamteric methods is the much larger computational requirements. This issue is dimenishing as coumputer hardware and efficient algoritm devlopment move forward but a well specified paramteric model requires a small fraction of the computing resourse as a comprably specified nonparamteric model.

Despite these trade-offs nonparametric methods do bring alot to the table.

This type of data evaluation is only available when using nonparamteric regression analysis and the peculiar results on own price elasticity of beans obtained by evaluating the distribution of elasticities would have been obscured with traditional paramaetric regression.

References

Barnett, W. A., & Seck, O. (2008). Rotterdam Model versus Almost Ideal Demand System: Will the Best Specification Please Stand Up? *Journal of Applied Econometrics*, *23*(6), 795–824.

Blundell, R., & Robin, J. M. (1999). Estimation in large and disaggregated demand systems: an estimator for conditionally linear systems. *Journal of Applied Econometrics*, *14*(3), 209–232.

Cheng, M.-Y., Hall, P., & Titterington, 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.

Deaton, A. (1980). *Economics and consumer behavior.* (J. Muellbauer, Ed.). Cambridge ; New York : Cambridge University Press.

Deaton, A., & Muellbauer, J. (1980). An Almost Ideal Demand System. *The American Economic Review*, *70*(3), 312–326.

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. (1997). Consistent Significance Testing for Nonparametric Regression. *Journal of Business & Economic Statistics*, *15*(3), 369–378.

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

Racine, J., Hart, J., & Li, Q. (2006). Testing the Significance of Categorical Predictor Variables in Nonparametric Regression Models. *Econometric Reviews*, *25*(4), 523.

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.