Forecasting of Households Consumption Expenditure with Nonparametric Regression: The Case of Turkey

Noyan Aydin [*]
Taner Akmercan [**]

Abstract: The relationship between household income and expenditure is important for understanding how the shape of the economic dynamics of the households. In this study, the relationship between household consumption expenditure and household disposable income were analyzed by Locally Weighted Scatterplot Smoothing Regression which is a nonparametric method using R programming. This study aimed to determine relationship between variables directly, unlike making any assumptions are commonly used as in the conventional parametric regression. According to the findings, effect on expenditure with increasing of income and household size together increased rapidly at first, and then speed of increase decreased. This increase can be explained by having greater compulsory consumption expenditure relatively in small households. Besides, expenditure is relatively higher in middle and high income levels according to low income level. However, the change in expenditure is limited in middle and is the most limited in high income levels when household size changes.

Keywords: Household Consumption Expenditure; nonparametric regression; LOESS.

JEL Classification: C14, D12.

Introduction

Regression analysis is one of the commonly used methods to examine the relationship between variables. In parametric regression analysis, functional form of the relationship between variables is predetermined or analysis is performed by assuming that the relationship between variables is appropriate for a specific functional form.

---

[*] Noyan Aydin is at Faculty of Economics and Administrative Sciences, Department of Econometrics Kutahya, Turkey.
[**] Taner Akmercan is a PhD Student. This paper is derived from his PhD thesis under the supervision of first author.

In here, the aim is to determine the relationship between variables by estimating the parameters of this specific functional form. However, it cannot always possible to determine how a functional form the relationship between variables has. Besides, it is necessary to be provided a number of assumptions with regard to a predetermined functional form of the relationship between variables in parametric regression methods. This case sometimes can be a major problem for researchers.

Nonparametric regression methods try to uncover the relationship between variables directly without making any assumptions about the functional form of the relationship between variables. Therefore, nonparametric methods provide flexibility for researchers to examine the relationship between variables. Additionally, although most of the basic information contained in the sample seems to be lost at first appearance, it seems not to be so when the theoretical efficiency of the method is analyzed.

Non-parametric methods are less effective than the parametric model that fits a normal distribution. On the contrary, they are more effective in populations that do not meet the normal distribution (Hollander et al., 2014). In this context, nonparametric methods can be used in order to test the validity of the results obtained by the parametric method.

The relationship between household expenditures and income is very important in understanding of the socio-economic structure and consumption patterns of households and in determining the policies to be implemented. Furthermore, some socio-demographic characteristics, which have a significant impact on these two variables, are important to consider when examining this relationship because it will provide information that is more detailed about the structure.

In literature, relationship between expenditure and income has been subject of many studies that use parametric regression method. In these studies, place of residence variable in Sarımeşeli (1999), geographic area variable in Emeç (2001) , place of residence and education level in Beyaz (2007), gender variable on the basis of the Dumlupınar University in Zortuk (2009) and various socio-demographic variables in Avcı (2011) were also given place modeling the relationship between income and expenditure.

In literature, relationship between expenditure and income has been subject of few studies that use nonparametric regression method. Bieren&Pott-Butter (1990) and Delgado&Miles (1997) studied the relationship between Household's compulsory consumption expenditure and disposable income with household-type variable using non-parametric kernel regression method. In another similar study, Gölveren (2012) analyzed this relation with various socio-demographic characteristics using kernel regression method, too.

In this study, the relationship between household compulsory expenditure and household's disposable income was investigated by LOESS (Locally Weighted Scatterplot Smoothing) regression method as nonparametric method. In the litera-

ture, there are many studies examining this relationship using parametric methods. However, there are very few studies using non-parametric methods. Although using parametric methods requires some strict presumptions parametric method, researchers are encouraged to use these methods much more because it is easier to use this method in order to estimate and it can be possible to create scenarios easily in the estimation process in application of these methods. For this reason, a nonparametric method was used in this study in order to enrich the methodology and in order to uncover an unknown structure in relationship if there is. Besides, it is not enough to examine the relationship between expenditure and income alone for a realistic viewpoint how to shape the economic dynamics. Therefore, household size variable, which is defined by OECD, was added to model in order to obtain a more comprehensive information about relationship.

Material and Dataset

Household Budget Survey, which is providing information about households' socio-economic structures, living standards and consumption patterns, is one of the most important sources that are used to test the implementation of socio-economic policy (Gölveren, 2012). In this context, the data set in the study was obtained from Household Budget Survey 2014. This survey was obtained by TSI applying to 13.248 sample households for the period of Jan.1-Dec. 31 in 2014. In this study, compulsory consumption expenditure as dependent variable, household disposable income, and household size defined by OECD as independent variables are used. Compulsory consumption expenditure and household disposable income variables are average annual value and are used taking their logarithm because of having large values.

Non-parametric Regression

In parametric regression methods, functional form of the relationship between dependent and independent variables is predetermined and accordingly it is necessary to provide certain preliminary assumptions. Conversely, in non-parametric regression methods, it is no need to be determined the functional form and to be made any assumptions about the relationship between variables. It is only assumed that the mean of the error term is zero and the variance of the error term is a finite number (Hart, 1997). Therefore, nonparametric regression methods can be used when assumptions are not provided or in the absence of a priori information about relationships between variables. Furthermore, non-parametric methods are preferable because of providing consistent estimates and helping to build a parametric model (Gölveren, 2012).

The nonparametric regression method has four main objectives and these are listed as below (Hardle, 1992: 7-8):
- ✓ To provide a versatile method to reveal the relationship between variables
- ✓ To produce estimates without need for parametric models
- ✓ To reveal fake observations by examining outliers
- ✓ To create a flexible method making an interpolation to replenish the lost values

Let a non-parametric regression model be as below:

$$yi = m(x_i) + ei, \qquad i = 1, 2, \dots, n \tag{1}$$

In here, the aim is to estimate the function $m(x_i)$ and this function is called as smoother. Therefore, the concept of smoothing is important in nonparametric regression method. Smoother is a tool used to express the trend of the dependent variable (Tezcan, 2009).

The most important feature of a smoother is not to assume precisely the form of the relationship between variables. Idea of fitting the data set to a curve and idea of using flexible functions, which can be combination of simpler functions, lies based on the idea of smoothing (Yıldız, 2013). This curve estimation (fitting) process is generally called as smoothing (Hardle, 1992). Besides, the curve that is estimated by smoother is called as smooth curve.

Kernel Regression Method

In non-parametric regression analysis, kernel regression as an estimation method is commonly known and is frequently used. In this method, curve estimation is made by giving local weights to the dependent variable via functions called kernel. Therefore, such non-parametric estimation methods are called as kernel regression method.

Kernel functions used in the kernel regression method are indicated by $K(\cdot)$ and it provides the following features:

- ✓ $\forall u, K(u) \geq 0$
- ✓ $\int_{-\infty}^{+\infty} K(u)du = 1$
- ✓ $K(-u) = K(u),$

The weight that is obtained depending on a kernel function is shown as:

$$W_i(x) = \frac{K\left(\frac{x-x_i}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right)} \tag{2}$$

In here, $W_i(x) = W(x, x_i)$ refers to the weight of observation i depending on the distance of $|x - x_i|$. These weights decrease when the distance of $|x - x_i|$ increase, and vice versa. Besides, $K(\cdot)$ represents the kernel function, h represents bandwidth and $\sum_{i=1}^{n} W_i(x) = 1$. The most important issue in here is the choice of bandwidth and kernel function. However, the choice of kernel function is less important if it compared with the choice of bandwidth. Additionally, change occurred in the form of kernel function has less effect for estimation than effect caused by changes occurred in bandwidth (Gölveren, 2012).

Nadaraya-Watson and Gasser-Müller are estimation methods commonly used in kernel regression method. Nadaraya-Watson method is considerably bias in the border region especially. On the other hand, this bias decreases but variance increases in Gasser-Müller method (Tezcan, 2009).

The K-Nearest Neighborhood Method

The weighted averages are calculated for y values, which are situated in fixed neighborhoods of X is determined by bandwidth h, during the estimation of unknown regression function $m(\cdot)$ in kernel regression method. When the k-nearest neighborhood estimation method is used, weighted averages are found for the y values in neighborhood of X, too. Nevertheless, the most important difference in here is that bandwidth is fixed in the kernel regression but is flexible in k-nearest neighborhood method.

The following function is used in K-nearest neighborhood estimation:

$$\widehat{m}(x) = \frac{1}{n} \sum_{i=1}^{n} W_{ki}(x) Y_i \tag{3}$$

In here, $\{W_{ki}(x)\}_{i=1}^{n}$ represents weights. Besides, providing that $J_x = \{i: Xi\}$ is one of k observations that are situated in the nearest neighborhood of X, W_{ki} is expressed as follows:

$$W_{ki} = \begin{cases} \dfrac{n}{k}, & if\ i \in J_x \\ 0, & in\ other\ cases \end{cases} \tag{4}$$

In this estimation method, k is expressed as the smoothing parameter and the most important issue in the estimation is selection of smoothing parameter k. This parameter should be large enough to obtain a proper estimate by reducing roughness in estimated curve as well as it should be chosen small enough to reduce bias (Hardle et. al., 2004).

LOESS (Locally Weighted Scatterplot Smoothing) Regression Method

Cleveland (1979) has proposed this method firstly, which is also known as locally weighted regression method or local polynomial regression method. This method, which is based on the principle of smoothing of scatterplot, was developed its methodology and was adapted to multivariate data by Cleveland and Devlin (1988). Ezekiel studied the first sample of the scatterplot smoothing in 1941. In this method for the smoothing process, the points on the scatterplot are grouped according to the value of x_i and new values of $(x_{i_{mean}}, y_{i_{mean}})$, which are obtained by averaging the x and y values for each group, are marked on the scatterplot. Subsequently in scatterplot smoothing method suggested by Clark (1977), consecutive points are being interpolated in a straight line and after these points are smoothed by making the convolution with a weighting function (Cleveland, 1979). However, the smoothing in these methods is very sensitive to outliers, whereas the smoothing is made highly robust to outliers because of using iterations in LOESS method.

Let (x_i, y_i) denote the points on scatterplot for $i = 1, 2, ..., n$ and then $(x_i, \hat{y}_i)$ are the points obtained in result of smoothing on this scatterplot. In here, $\hat{y}_i$ values are fitted values on x_i points. On the other hand, let W is a weight function that has the following features:

✓ $W(x) > 0$ for $|x| < 1$
✓ $W(x) = 0$ for $|x| \geq 1$
✓ $W(x) = W(-x)$
✓ $W(x)$ is a non-increasing function for $x \geq 0$

Let $0 < f \leq 1$ and let r be f_n rounded to the nearest integer. For each x, weights are defined as $w_k x_i$ for $k = 1, ...,$ using the weight function. W becomes first 0 is at r_{th} nearest neighbor of x_i. The initial fitted value, $\hat{y}_i$, at each x_i is the fitted value of a d_{th} degree polynomial fit to the data using weighted least squares with weights $w_k x_i$ (Cleveland, 1979: 829-831).

This procedure can be expressed in other way. For each i let h_i be the distance from x_i to the r_{th} nearest neighbor of x_i. That is, h_i is the r_{th} smallest number among $|x_i - x_j|$, for $j = 1, ..., n$. Let $w_k x_i = W(h_i^{-1}(x_k - x_i))$ for $k = 1, ..., n$. Locally weighted regression defined by the following sequence operations (Cleveland, 1979: 829-831):

1. For each i compute the estimates, $\hat{\beta}_j(x_i)$, $j = 0, ... , d$, of the parameters in a polynomial regression of degree d of y_k on x_k, which is fit by weighted least squares with weight $w_k x_i$ for (x_k, y_k). Thus, $\hat{\beta}_j(x_i)$ are the values of β_j that minimize

$$\sum_{k=1}^{n} w_k(x_i)(y_k - \beta_0 - \beta_1 x_k - \cdots - \beta_d x_k^{d})^2 \tag{5}$$

The smoothed point at *x*, using locally weighted regression of degree *d* is $(x_i, \hat{y}_i)$, where $\hat{y}_i$ is the fitted value of the regression at x_i. Thus,

$$\hat{y}_i = \sum_{j=0}^{d} \hat{\beta}_j(x_i) x_i^{\ j} = \sum_{k=1}^{n} r_k(x_i) y_k \tag{6}$$

Where $r_k x_i$ does not depend on y_j, $j = 1, ..., n$,

2. Let C be the bisquare weight function that is defined by

$$B(x) = \begin{cases} (1 - x^2)^2, & for \ |x| < 1 \\ 0 \quad , & for \ |x| \geq 1 \end{cases} \tag{7}$$

and let $e_i = y_i - \hat{y}_i$ be the residuals from the current fitted values and be the median of $|e_i|$. Define robustness weights by $\delta_k = C(e_k/6s)$.

3. Compute new $\hat{y}_i$ for each *i* by fitting a d_{th} degree polynomial using weighted least squares with weight $\delta_k w_k(x_i)$ at (x_k, y_k).

4. Repeatedly carry out steps 2 and 3 a total of *t* times. The final $\hat{y}_i$ are robust locally weighted regression fitted values.

The iterative fitting in steps 2 to 4 is carried out to achieve robust smoothed points in which a small fraction of outliers does not distort the results. The outliers, which can be thought of as arising when e_i has a long-tailed distribution, tend to have small robustness weights, δ_k and therefore do not play a large role in the determination of the smoothed points. The bisquare function is used because other investigations have shown it to perform well for robust estimation of location (Gross, 1976) and for robust regression (Gross, 1977).

Since LOESS regression method has a strong smoothing effect against outliers, it is superior to kernel regression methods by reducing bias estimation. Furthermore, LOESS regression method summarizes the relationship between the dependent variable and one or more independent variables graphically. Therefore, the most distinctive feature of this method is that allowing data to self-talk. A curve estimated by this method is obtained empirically without strict presumptions about any functional relationship may be in data. In this way, complex relationships, which may be overlooked easily by traditional estimation methods, can be uncovered more easily by LOESS method (Jacoby, 2000).

There are four fundamental elements in the LOESS regression method. These are weight function as *W*, polynomial degree as *d*, the number of iterations as *t* and the span value as smoothing parameter. The recommended function for weight function *W* is as follows:

$$W(x) = \begin{cases} (1 - x^3)^3, & |x| < 1 \\ 0 & , & |x| \geq 1 \end{cases} \tag{8}$$

Because this weight function improves chi-square distribution for estimating the error variance.

For t as the number of iterations, numerous studies with real or artificial data sets show that two iterations are adequate for almost all situations (Cleveland, 1979). On the other hand, linear ($d=1$) and quadratic ($d=2$) polynomial are usually used in studies related with choice of the degree of polynomial. Because three and higher degree polynomials do not contribute to estimation. Besides, linear polynomial generates artificial flatness in estimation when there are outliers in the data set. Therefore, quadratic polynomial is preferred because it gives satisfactory results for almost any situation (Keele, 2008; Jacoby, 2000).

Span value is a smoothing parameter value in the 0-1 range. This value should be selected to provide the best balance between the variance and the bias. Because, very large span value will reduce the variance but it will increase the value of bias and vice versa. The easiest and most common method used in the selection of the span value is visual trial and error method. In this method, span value is initially determined as 0,5 and if obtained curve is rough then span value is increased by 0,1. By continuing in this way, the most appropriate and the smoothest curve for data is tried to achieve (Keele, 2008).

The Goodness of Fit Measure in LOESS Regression Method

R-squared as a measure of goodness of fit in LOESS regression method is calculated as follows:

$$R^2_{LOESS} = \frac{\sum_{i=1}^{n}\left(\hat{g}(x_i) - \overline{\hat{g}(x_i)}\right)^2}{\sum_{i=1}^{n}(y_i - \bar{Y})^2} \tag{9}$$

In here, $\hat{g}(x_i)$ is estimated (fitted) value for observation; $\overline{\hat{g}(x_i)}$ is the mean of the estimated values; y_i is the actual value of the dependent variable for i_{th} observation and $\bar{Y}$ is the sample mean for dependent variable.

In LOESS regression method, R^2_{LOESS} cannot be interpreted as explained variance in parametric regression. Because, R^2_{LOESS} expresses the variance explained by the estimated values instead of the variance explained by the independent variables. For all that, false or meaningless R^2_{LOESS} values may occur when robust weights are used in the estimation process. Besides, the variance of the dependent variable is less im-

portant in LOESS regression like other non-parametric methods that do not try to fit specific model to the data set. Because of all these reasons, R^2_{LOESS} is rarely reported in studies with LOESS method (Jacoby, 2000).

Findings

Univariate models were built firstly for the desired relationship for variables. Thus, it was provided a preliminary opinion about the bilateral relations between variables. Then other variable was added to model, and the statistical significance of each independent variable was tested. Based on these results, three-dimensional graph of the variables were plotted and the results were evaluated.

In this study, quadratic polynomial weighted function and two iterations were used to obtain optimum results considering the criteria suggested in the literature. On the other hand, the value of span was determined as 0,5 because data set used is too large and there is no visible visual difference when other span values are selected. LOESS smoothing curves for univariate models are given below:

Figure 1: LOESS smoothing curves for univariate models
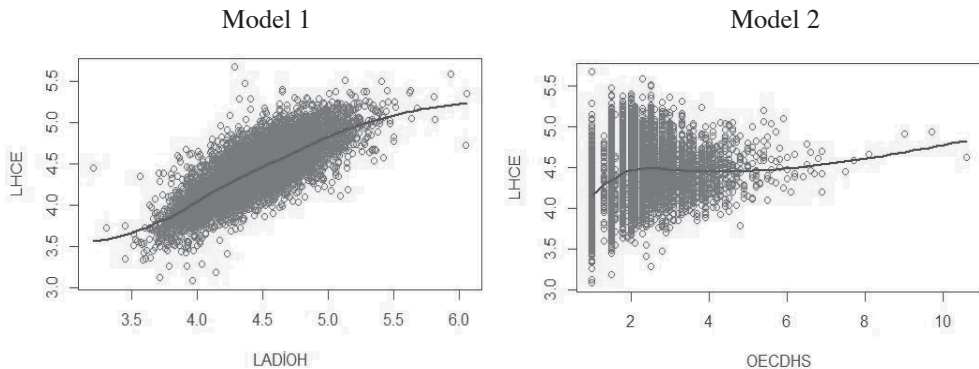
In here, LHCE indicates logarithm of household compulsory consumption expenditure, LADIOH indicates logarithm of household annual disposable income, and OECDHS indicates household size variable defined by OECD. When the model 1 is examined, the relationship between the two variables is almost linear in parts of intensive data, but the relationship has curvilinear form deviating from linearity in the end parts that have less data. The curvilinearity occurring in this end parts shows that household annual disposable income has less effect on household compulsory consumption expenditure in this parts. When the model 2 is examined, a rapid increase seems in household compulsory consumption expenditure while household size in-

creases toward value of about 2. Afterwards, this increase decreases while household size increases toward value of about 4 and this increase stops while household size is in between about 4 and 6. On the other hand, this increase increases again later, but this increase is not a rapid increase as in the first portions. LOESS regression results for both models are given in Table 1.

Table 1: LOESS regression results for univariate models

	Model 1	Model 2
Span Value	0,5	
Number of Observations	10117	
Standard Errors of Residuals	0,1811	0,2737
Number of equivalent parameters (Degree of freedom of model)	8,02	7,54

Referring to table 1, it can be seen that the degree of freedom of the model is 8,02. In LOESS regression method, degree of freedom is not required an integer unlike the parametric regression (Cleveland, 1979: 835). In here, degree of freedom, which is also expressed number of equivalent parameters, is a non-negative number and it refers to the number of parameters that are required when a similar curve like LOESS curve is desired to obtain by parametric method (Jacoby, 2000).

In the built univariate model in the analysis, F test was used to test whether the variable is statistically significant. The F test statistics is calculated as follows:

$$F = \frac{(RSS_1 - RSS_2)/(df_2 - df_1)}{RSS_2/(n - df_2)} \tag{10}$$

In here, RSS_1 is the sum of squared errors of first model, RSS_2 is the sum of squared errors of second model, df_1 is degree of freedom of first model, df_2 is degree of freedom of second model, and n is total number of observations (Jacoby, 2000).

The purpose of performing this test for each variable is to compare the model has related variable and the model has not. Thus, it can be determined whether there is a statistically significant difference between the models. In other words, it can be seen by using F test whether variable has a significant contribution to the model. F test results for household disposable income variable are given in Table 2:

Table 2: Results of F test for household annual disposable income variable

	Number of equivalent parameters	Sum of Squared Errors	F value	P value
LHCE ~ OECDHS	7,54	757,44	944,48	2,2e-16*
LHCE ~ OECDHS + LADIOH	16,83	328,51		

*…is statistically significant at 0,001.

According to the results analysis, it is seen that the change in sum of squared errors is statistically significant at 0,001 with the addition of household disposable income variable to the model. In this case, it is understood that household disposable income has a statistically significant effect on household compulsory consumption expenditure.

Similarly, F test results for household size variable by defined OECD are given below.

Table 3: Results of F test for household size variable

	Number of equivalent parameters	Sum of Squared Errors	F value	P value
LHCE ~ LADIOH	8,02	331,57	7,0289	7,995e-13*
LHCE ~ LADIOH + OECDHS	16,83	328,51		

*…is statistically significant at 0,001.

According to the results analysis, it is seen that the change in sum of squared errors is statistically significant at 0,001 with the addition of household size variable by defined OECD to the model. In this case, it is understood that household size variable by defined OECD has a statistically significant effect on household compulsory consumption expenditure.

The results of tests show that both independent variable are statistically significant effect on dependent variable. It is seen that the p value of F test statistics for household size variable is bigger than the p value of F test statistics for household disposable income variable. This is because household size is less effective on household compulsory consumption expenditure in comparison with household disposable income.

Multivariate model was built by adding both independent variables to the model and the results were given in Table 4.

Table 4: LOESS regression results for multivariate models

	LHCE ~ LADIOH + OECDHS Model 2
Span Value	0,5
Number of Observations	10117
Standard Errors of Residuals	0,1804 0,2737
Sum of Squared Errors	328,51
Number of equivalent parameters (Degree of freedom of model)	16,83

Three-dimensional graphic is shown in Fig. 2 in order to make inferences about the multivariate model based on the results obtained.

Figure 2: Three-dimensional graphic of the multivariate model
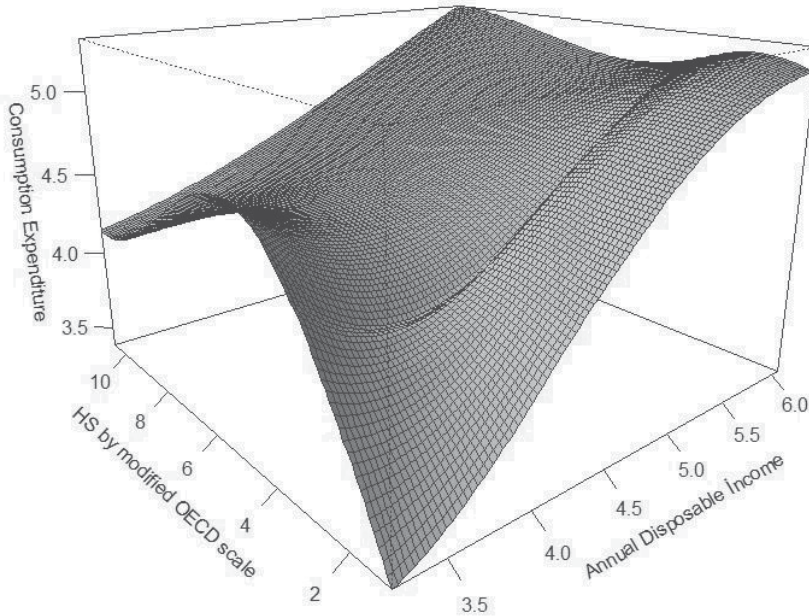
It is seen that the relationship between household disposable income variable and household compulsory consumption expenditure variable is non-linear but almost linear when Figure 2 is examined. Similarly, the relationship between household size variable by defined OECD and household compulsory consumption expenditure is non-linear, too. On the other hand, household compulsory consumption expenditure increases when household disposable income increases in general but it is not valid for household size variable by defined OECD. When Figure 2 is examined, it is seen that a high increase occurs in household compulsory consumption expenditure while household size increases toward to the value of about 6. After this point, household compulsory consumption expenditure stops decreasing while household size variable goes on towards to value of 10.

The effect of household size on household compulsory consumption expenditure decreases with the decrease of household disposable income when household disposable income increases toward to the value of about 3,5. However, this effect decreases and reverses after the value of about 6 of household size. Similarly, the effect of household size on household compulsory consumption expenditure increases toward to the value of about 2 when household disposable income is at the highest level. Then, this increase decreases until the value of about 4,5 and then increasing continues again.

Conclusion

Many researchers studied the relationship between household consumption expenditure and household disposable income frequently and in these studies parametric methods often are used. Therefore, the relationship between household consumption expenditure and household disposable income were examined with household size as socio-demographic variable by LOESS regression as a nonparametric method. The aim in this study is to enrich the related literature by providing a different perspective to reveal hidden structures contained in the data structure if there is. In this manner, it is thought that this study will be shed light to more comprehensive household's studies. In this context in this study at firstly, the relationships between variables were examined with a binary comparison. Then, obtained findings were evaluated by building multivariate models.

When household compulsory consumption expenditure, household annual disposable income, and household size defined by OECD were examined together, it is seen that the effect on household expenditure with increasing of household income and household size together increased rapidly at first, and then the speed of increase decreased. This increase occurred in the first part can be explained by having greater compulsory consumption expenditure relatively in small household as it is seen in the binary comparisons. In other words, compulsory consumption expenditure, which is independent from income, is relatively high in small household. Especially, compulsory consumption expenditure increases rapidly while the number of people is 1 to 6 in the household increases in households in the low-income group.

When looking at the relationship in terms of the middle-income level, independent expenditure is relatively higher according to low-income group. However, the change in independent expenditure is limited when household size changes according to low-income group.

On the other hand, when looking at the relationship in terms of the high-income level, independent expenditure from income is relatively higher according to low-income and middle-income groups. However, the change in independent expenditure is the most limited when household size changes according to low-income and middle-income groups. That is to say, expenditure is not so much associated with the level of household size in high-income households.

In brief, the findings obtained by nonparametric models are quite rich according to parametric models because functional form of the relationship is not limited with a specific pattern. In this regard, examining of the relationship between household consumption expenditure and income can be help to better understanding the relationship, moreover, in this context, it is understood more clearly, how the economic dynamics act and so more effective economic policies implement by relevant public authorities.

REFERENCES

Avcı, G.M. (2011). *The effect of working women on patterns of household expenditures in Turkey*. Unpublished master thesis, Zonguldak Karaelmas University, Zonguldak, Turkey.

Beyaz, F.B. (2007). *Household consumption expenditures and demand estimation in Turkey*. Unpublished master thesis, Akdeniz University, Antalya, Turkey.

Bierens, H.J. & Pott-Buter, H.A. (1991). Specification of household Engel curves by nonparametric regression. *Econometric Reviews*. 9(2), 123-184. DOI: 10.1080/07474939008800185.

Clark, R.M. (1977). Non-parametric estimation of a smooth regression function. *Journal of the Royal Statistical Society*, *Series B (methodological)*. 39(1), 107-113.

Cleveland, W.S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*. 74(368), 829-836. DOI: 10.1080/01621459.1979.10481038

Cleveland, W.S. & Devlin, S.J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*. 83(403), 596-610. DOI: 10.1080/01621459.1988.10478639.

Delgado, M.A. & Miles, D. (1997). Household characteristics and consumption behavior: a nonparametric approach. *Empirical Economics*. 22(3), 409-429.

Emeç, H. (2001). Interregional Consumption Expenditures in Turkey: Tobit Model Approach. Dokuz Eylul University, Faculty of Economics and Administrative Sciences Journal. 16(2), 61-73.

Gölveren, E. (2012). *Kernel regression which is nonparametric regression model and an application in regard to household expenditure in Turkey*. Unpublished master thesis, Ataturk University, Erzurum, Turkey.

Hollander, M., Wolfe, D.A. & Chicken, E. (2014). *Nonparametric statistical methods* (3rd ed.). New York: John Wiley & Sons, Inc.

Hardle, W. (1992). *Applied nonparametric regression*, Cambridge: Cambridge University Press.

Hardle, W., **Müller**, M., Sperlich, S. & Werwatz, A. (2004). *Nonparametric and semiparametric models*, Berlin: Springer.

Hart, J.D. (1997). *Nonparametric smoothing and lack-of-fit tests*. New York: Springer.

Jacoby, W.G. (2000). Loess: a nonparametric, graphical tool for depicting relationships between variables. *Electoral Studies*. 19(4), 577-613. DOI: 10.1016/S0261-3794(99)00028-1.

Keele, L. (2008). *Semiparametric regression for the social sciences,* UK: John Wiley & Sons Ltd.

Sarımeşeli, M. (1999). Household Spending Trends. *Gazi University, Faculty of Economics and Administrative Sciences Journal*. 1(2), 41-50.

Tezcan, N. (2009). *Approach to estimation with non-parametric regression method*. Unpublished PhD thesis, Istanbul University, Istanbul, Turkey.

TUIK, (2014). *Household Budget Survey*. Ankara: TUIK.

Yıldız, M. (2013). *Analysis of nonparametric fuzzy regression models*. Unpublished PhD thesis, Anadolu University, Eskisehir, Turkey.

Zortuk, M. (2009). In the context of gender variable, the relationship between income and the expenditure sub-groups the application on Dumlupinar University Student. *Dumlupınar University, Journal of Social Sciences*. 23(23), 25-32.