Testing Moderation Effects using Non-Parametric Regressions

A. George Assaf[1]
Professor
Isenberg School of Management
University of Massachusetts-Amherst
90 Campus Center Way, 204A Flint Lab, Amherst, MA, 01003
Tel :(+1) 4135451492
assaf@isenberg.umass.edu

Mike G. Tsionas
Professor of Econometrics
Lancaster University Management School
m.tsionas@lancaster.ac.uk

and

Athanasios Andrikopoulos
Lecturer in Finance
Hull University Business School
A.Andrikopoulos@hull.ac.uk.

[1] Corresponding author

Testing Moderation Effects using Non-Parametric Regressions

Abstract

Testing moderation effects is highly common in the hospitality literature. Most theories in the field depend on variables that alter the nature and direction of the relationship between two variables. While moderation continues to be heavily used, methods for testing moderation effects are not always robust. One common problem that researchers face is the need to pre-assign a particular functional form when there is actually no guarantee that such functional form is necessarily correct. The aim of this note is to address this problem. We describe three different non-parametric models that offer more flexibility in testing moderating effects without a need to pre-impose a specific functional form. We test the three models on an interesting application involving the moderating role of corporate social responsibility (CSR) on the relationship between advertising and firm value. The results revealed interesting moderating effects that go beyond the simple linear moderation.

1. Introduction

It is highly common in the tourism and hospitality literature to test moderating hypotheses using a parametric regression model of the form $y = \beta_1 x + \beta_2 z + \beta_3 xz$, where x is an explanatory variable, z is a moderator, and the moderating effect is simply tested with β_3 (one can derive the moderating effect mathematically with the cross-derivative of y with respect to (x, z): $\frac{\partial^2 f(x,z)}{\partial x \partial z} = \beta_3$). One potential problem with such an approach is that one needs to pre-assign a functional form (in this case linear) prior to testing for moderation effects.

This note takes up a different approach to relax such assumption. We consider non-parametric estimation based on local linear likelihood methods in panel data, where firm effects are necessarily present. To wipe out these effects we use a differencing estimator. As many models have been proposed in the literature, it is not entirely clear what functional forms other than $y = \beta_1 x + \beta_2 z + \beta_3 xz$ are allowed to test for moderation effects. In fact, there is no guarantee that such a functional form is necessarily true. In other words, simply relying on such model runs the risk of mispecifying the functional form when testing for moderation. The issue of misspecification of functional forms has increased the need for non-parametric estimation (see Su and Ullah, 2011). In this paper, we extend the use of such models to the context of moderated regression. We describe three different models which are open for testing and the best model should be selected based on the data at hand. It is also possible that these models may reveal a simple linear moderation, but the goal is not to make this assumption *a priori*. With the use of non-parametric regression and the models we are describing one would let the data speak for itself and determine the best fitting moderation, if any.

2. Models

In this section, we propose three different formulations that can be estimated in a non-parametric framework. These models do not require pre-assigning a specific functional form and hence, are far more flexible than the traditional linear regression commonly used in the literature to test for moderation effects (for example see Ro, 2012; Lee et al., 2013; Sun and Lee, 2018; and Yoon et al., 2016).

Model A

In the first model (Model A) we suggest a formulation of the following form:

$$y_{it} = \alpha_0 + \alpha_1 x_{it} + \alpha_2 z_{it} + w'_{it}\gamma + x_{it}\beta(z_{it}) + \mu_i + v_{it}, \qquad (1)$$

where x_{it} is the single regressor, z_{it} is the single moderator, w_{it} is a vector of predetermined variables, and $\beta(z_{it})$ is an unknown function, μ_i is a firm specific effect and v_{it} is an error term with zero mean.

To simplify further, omit the linear terms for simplicity and consider:

$$f(x, z) = x\beta(z). \tag{2}$$

Suppose we expand $\beta(z_{it})$ in a Taylor series expansion. This gives:

$$f(x, z) = \beta_1 xz + \beta_2 xz^2 + \ldots \tag{3}$$

Other than w, no linear terms in x, z need to be present. Therefore, in this case z can moderate (nonlinearly) a linear relation between y and x.

Model B

For Model B we suggest adopting a more flexible formulation of the following form:

$$y_{it} = \alpha_0 + \alpha_1 x_{it} + \alpha_2 z_{it} + w_{it}'\gamma + g(x_{it})z_{it} + \mu_i + v_{it}, \tag{4}$$

where x_{it} is the single regressor, z_{it} is the single moderator, w_{it} is a vector of predetermined variables, and $g(x_{it})$ is an unknown function, μ_i is a firm effect and v_{it} is an error term with zero mean.

To illustrate further, omit the linear terms for simplicity and consider:

$$f(x, z) = g(x)z. \tag{5}$$

Suppose we expand $\beta(z_{it})$ in a Taylor series expansion. This gives:

$$f(x, z) = \alpha_1 xz + \alpha_2 x^2 z + \ldots \tag{6}$$

Other than w no linear terms in x, z need to be present. In this model, z moderates linearly a nonlinear relationship between y and x.

The marginal effect of x in Models A and B is as follows:

$$\begin{aligned}
\frac{\partial f(x, z)}{\partial x} &= \beta(z), \text{for\ \ model\ \ A,} \\
\frac{\partial f(x, z)}{\partial x} &= g'(x)z, \text{for\ \ model\ \ B.}
\end{aligned} \tag{7}$$

Therefore, the moderation effect of z is:

$$\frac{\partial^2 f(x,z)}{\partial x \partial z} = \beta'(z), \text{for model A,}$$
$$\frac{\partial^2 f(x,z)}{\partial x \partial z} = g'(x), \text{for model B.}$$

(8)

Therefore, the two models have different implication in that the moderation effect depends only on z in model A and only on x in model B.

Model C

Of course, there is the more general Model C in which, apart from linear terms, we have:

$$y = F(x,z) = \delta_1 x^2 + \delta_2 z^2 + \delta_3 xz + \cdots.$$

(9)

One should be able to determine which of three models A, B, C receives more support in the light of the data. Focusing on Model A, and using the so-called local linear approach, we have:

$$\beta(z) = \beta_0 + \beta_1 \cdot (z - z_o),$$

(10)

so that the value of β at $z = z_o$ is β_0 and its first derivative is β_1. Of course, β_0, β_1 will be different for different z_o. The same holds for Model B if we interchange the roles of x_{it} and z_{it}.

3. Model Estimation

We focus on estimating (1) as the treatment of (5) is the same if we interchange the roles of x_{it} and z_{it}. Model C is somewhat different. The main problem is the presence of firm specific effects μ_i. One approach is to use temporal differences to obtain:

$$\triangle y_{it} = \alpha_1 \triangle x_{it} + \alpha_2 \triangle z_{it} + \triangle w_{it}'\gamma$$
$$+ \{x_{it}\beta(z_{it}) - x_{i,t-1}\beta(z_{i,t-1})\} + v_{it} - v_{i,t-1},$$

(11)

where $\triangle y_{it} = y_{it} - y_{i,t-1}$ etc. Although the linear terms do not introduce any special problems, there is a problem with the non-parametric part which now is $\{x_{it}\beta(z_{it}) - x_{i,t-1}\beta(z_{i,t-1})\}$. If we omit the linear terms for simplicity in notation, we have:

$$\triangle y_{it} = x_{it}\beta(z_{it}) - x_{i,t-1}\beta(z_{i,t-1}) + u_{it},$$

(12)

where $u_{it} = v_{it} - v_{i,t-1}$. Henderson et al. (2008) suggested using an iterative procedure with a profile likelihood approach. Mammen et al. (2009), on the other hand, suggested using a smoothed backfitting algorithm.

In non-parametric regression, we are interested in obtaining the value of the unknown function at a given point. For example, if the model is $y_{it} = F(x_{it}, z_{it}) + u_{it}$, we are interested in the value of $F(x,z)$ where x,z are given. Similarly, in our case, we are interested in the value of $f(x,z) =$

$x\beta(z)$ where x, z are given. Specifically, we are interested in the value of $\beta(z)$ which gives the marginal effect of x and, of course, in the moderation effect, which is the first derivative of $\beta(z)$, viz. $\beta'(z)$. To obtain the value of the function at a point (x, z) agreeing $z_{it} = z_{i,t-1} = z$, the usual approach is to solve the problem (Rodriguez-Poo and Soberon, 2014, 2015):

$$\min_{\beta(z)} \sum_{i=1}^{n} \sum_{t=1}^{n} \{\triangle y_{it} - x_{it}\beta(z) + x_{i,t-1}\beta(z)\}^2 K_H(z_{it}, z_{i,t-1}). \tag{13}$$

where $K_H(x, x')$ denotes a bivariate kernel with scale matrix H. The novelty is that the kernel function $K_H(z, z')$ depends on both z_{it} and its lag. It is common practice to use a product kernel of the form: $K_H(z, z') = h^{-2}K(z/h)K(z'/h)$, where $K(z)$ is a univariate kernel such that $K(z) \geq 0$, and $\int_{-\infty}^{\infty} K(z)dz = 1$; for example $K(z) = (2\pi)^{-1/2}e^{-x^2/2}$, and $h > 0$ is a bandwidth parameter. (see Fan and Gijbels, 1995a,b; Ruppert and Wand, 1994; or Zhan-Qian, 1996).

For estimation purposes we need to define the objective function at two different points $z_{it} = z_o$ and $z_{i,t-1} = z_o'$ so that the objective becomes:

$$\min_{\beta(z)} Q(z_o, z_o'; D) : \sum_{i=1}^{n} \sum_{t=1}^{n} \{\triangle y_{it} - x_{it}\beta(z_o) \\ + x_{i,t-1}\beta(z'_o)\}^2 h^{-2}K(z_{it}/h)K(z_{i,t-1}/h), \tag{14}$$

where D denotes all available data. Of course, terms that are linear in x_{it}, z_{it}, w_{it} may appear in the term inside the brackets with constant coefficients. In the so-called local linear approach, we have:

$$\beta(z) = \beta_0 + \beta_1 \cdot (z - z_o), \tag{15}$$

so that the value of β at $z = z_o$ is β_0 and its first derivative is β_1. Therefore, the objective function becomes:

$$\min_{\beta(z)} Q(z_o, z_o'; D) : \sum_{i=1}^{n} \sum_{t=1}^{n} \{\triangle y_{it} - x_{it}[\beta_0 + \beta_1(z_{it} - z_o)] + x_{i,t-1}[\beta_0 + \beta_1 \cdot (z_{i,t-1} - z'_o)]\}^2 h^{-2}K\left(\frac{z_{it}}{h}\right)K\left(\frac{z_{i,t-1}}{h}\right) =$$

$$\min_{\beta_0, \beta_1} \sum_{i=1}^{n} \sum_{t=1}^{n} \{\triangle y_{it} - \triangle x_{it}\beta_0 - \triangle(x_{it}z_{it})\beta_1 + \beta_1 \cdot [x_{it}z_o - x_{i,t-1}z_o']\}^2 h^{-2}Kz_{it}/h)K(z_{i,t-1}/h), \tag{16}$$

where $\triangle(x_{it}z_{it}) = x_{it}z_{it} - x_{i,t-1}z_{i,t-1}$. Effectively, this is a weighted least squares problem. When $z_o = z'_o$ it simplifies to:

$$\min_{\beta_0, \beta_1} Q(z_o; D, \beta)$$

$$= \sum_{i=1}^{n} \sum_{t=1}^{n} \{\triangle y_{it} - \triangle x_{it}(\beta_0 - \beta_1 z_o) - \beta_1 \qquad (17)$$
$$\triangle (x_{it} z_{it})\}^2 h^{-2} K(z_{it}/h) K(z_{i,t-1}/h),$$

where $\beta = [\beta_0, \beta_1]'$. The form (17) can be used to determine the value of the function $\beta(z) = \beta_0 + \beta_1 \cdot (z - z_o)$ through β_o and its derivative β_1 which are, of course, specific to a particular value z_o, given the bandwidth parameter h.

The bandwidth parameter is determined using leave-one-out cross-validation, see for example Henderson et al. (2008) and Henderson and Ullah (2005).

4. Application

We test Models A, B and C using an interesting sample involving both restaurant and hotels firms from the United States. In our application, we test the impact of advertising on firm value, while using corporate social responsibility (CSR) as a moderator. While several studies tested the impact of advertising on firm value, the findings remain inconsistent. More particularly, only a few studies have included moderating variables in testing such a relationship. In line with recent research, we argue that "the positive reputation created by CSR activities, helps to maximize the effectiveness of advertising on sales and firm value due to more favorable attitudes toward the firm. This effectiveness will directly influence sales, and in turn, the value of the firm" (Roberts and Dowling 2002, p1079).

In this paper, we aim to test the moderating effect of CSR using the flexible models we are proposing. As mentioned, we have a combined sample of hotels and restaurants. In total, we have a sample of 232 observations covering an unbalanced panel of 22 publicly traded restaurants from 2001 to 2012, and 82 covering an unbalanced panel of 9 publicly traded hotels from 2001 to 2012. We collected the advertising spending and firm value data using the COMPUSTAT database. For CSR data, we used the KLD Research and Analytics' KLD STAT, which covers seven major types of CSR initiatives. We measure CSR as total CSR strengths minus total CSR concerns in line with Rekker et al. (2014). Finally, we measure firm value using the Market Value Added (MVA)= market value-capital, where "market value reflects the equity market valuation of the firm and capital reflects the debt and equity invested in the firm" (Assaf e al. 2017, p. 1488).

5. Results

5.1. Model Comparison

Using our application[2], we first test which of the three models we are proposing fits the data better[3]. Model selection is based on the error sum of squares (ESS) from models A, B and C which can be converted to an R^2-like measure, as $R^2 = 1 - \frac{ESS}{TSS}$, where TSS is the total sum of squares. As one-off model selection is a dangerous practice, we try to examine whether R^2-like measures are sensitive to particular firms. Therefore, we omit M firms from the sample (viz. all its temporal observations) and we solve (17) using cross-validation in the resulting sample. Here, M is randomly selected between 1 and 5 and the firms are randomly chosen as well. We repeat this exercise 1,000 times. In other words:

- o   we omit M firms from the sample
- o   we solve (17) for the sample of remaining firms
- o   we compute R^2
- o   we repeat 1,000 times to determine the sample distribution of R^2.

The results are presented in Figure 1. From the results it seems that models B and C do roughly the same, model A does roughly the same in some instances but, most of the time, it does *considerably better* as the R^2 has a dominant mode near 0.55[4]. We proceed, therefore, on the assumption that model A does best, in this data set, and we report results in Figures 2 and 3.

It is important to note that we also run specification tests of each of the non-parametric models against two different parametric models. The null hypothesis is that a parametric model is correctly specified. The results are reported in Table 2 where the p-value without a parenthesis corresponds to a simple parametric model which includes the following variables: constant, advertising, CSR, size and advertising*CSR. The p-value in the parenthesis corresponds to a more involved parametric model which also includes the square of advertising multiplied by CSR. All parametric models are estimated in a panel framework and include firm specific effects. As we can see from the results, all non-parametric models seem to outperform each of the two parametric specifications. For instance, the p-value in each case indicates a rejection of the null hypothesis that a parametric model is correctly specified

---

[2] All estimations were performed in Fortran 77. Codes can be provided upon request.

[3] Our purpose here is to illustrate the performance of all three models. However, we also believe that the selection of a particular specification should be driven by theoretical arguments.

[4] Following Ullah and Wang (2013) and Hurvich et al. (1998) we also calculated the Akaike information criterion (AIC), which again proved that Model A performs best.

5.2. Moderating Effect

Using Model A, we show in Figure 2 the moderation effect of z when we fix x at different values. Also reported in Figure 2 is the 95% confidence interval to test for significance. The moderation effect is, of course, given by the cross-derivative of y with respect to (x, z). We can see that the moderation effect is clearly non-linear. As the moderator "z" increases from the lowest 10% through the sample median to the upper 10%, the effect becomes more stable, implying that at higher values of the moderator, there tends to be more convergence. In our case, this would mean that heterogeneity is less pronounced at values of z near the upper 10% of y-values. Practically, this would mean the moderating effect of CSR tends to be more stable at a higher level of advertising spending.

The link between advertising and firm value has been researched extensively, but the findings have been inconclusive or even conflicting (see Luo and de Jong, 2012 and Assaf et al. 2017). As emphasized by Luo and de Jong (2012), there is a need for a contingency approach. The link between "advertising spending and firm value is not simple, and assuming that there are no variables interacting with this relationship is unrealistic" (Assaf et al., 2017, p.1485). Yet, research on potential moderators is scant. Results from this study seem to confirm the importance of CSR for such relationship. Practically, the results indicate that for CSR to have a more effective role, managers need to invest more aggressively in advertising as the effect of CSR is more pronounced at a higher level of advertising.

6. Concluding Remarks

This note has described three different models for testing moderation effects in a non-parametric fashion. Our main/central argument is that simply testing moderation effects using a pre-assigned functional form may not be necessarily true. To avoid running the risk of misspecification one can use a non-parametric approach and let the data speak for itself. The three models proposed in this paper are flexible and do not restrict the moderation effects to simple linear moderation. We tested the models on an interesting data set involving the moderating effect of CSR on the relationship between advertising and firm value. The results clearly indicated that all three non-parametric models fit the data better than their parametric counterpart did. It was also clear that the moderation effect is non-linear ranging from non-significant to significant effect at a low value of advertising to become more stable as advertising increases. Of course, the models we are proposing are open for testing and the researcher is encouraged to compare between them before resorting to the correct specification.

Table 1. Descriptive Statistics

Variable	Mean	SD
MTA	1877.33	8244.15
Advertising	92.58	163.88
CSR	0.80	5.88
Size	3.08	0.61

Figure 1. Cross-validated R^2
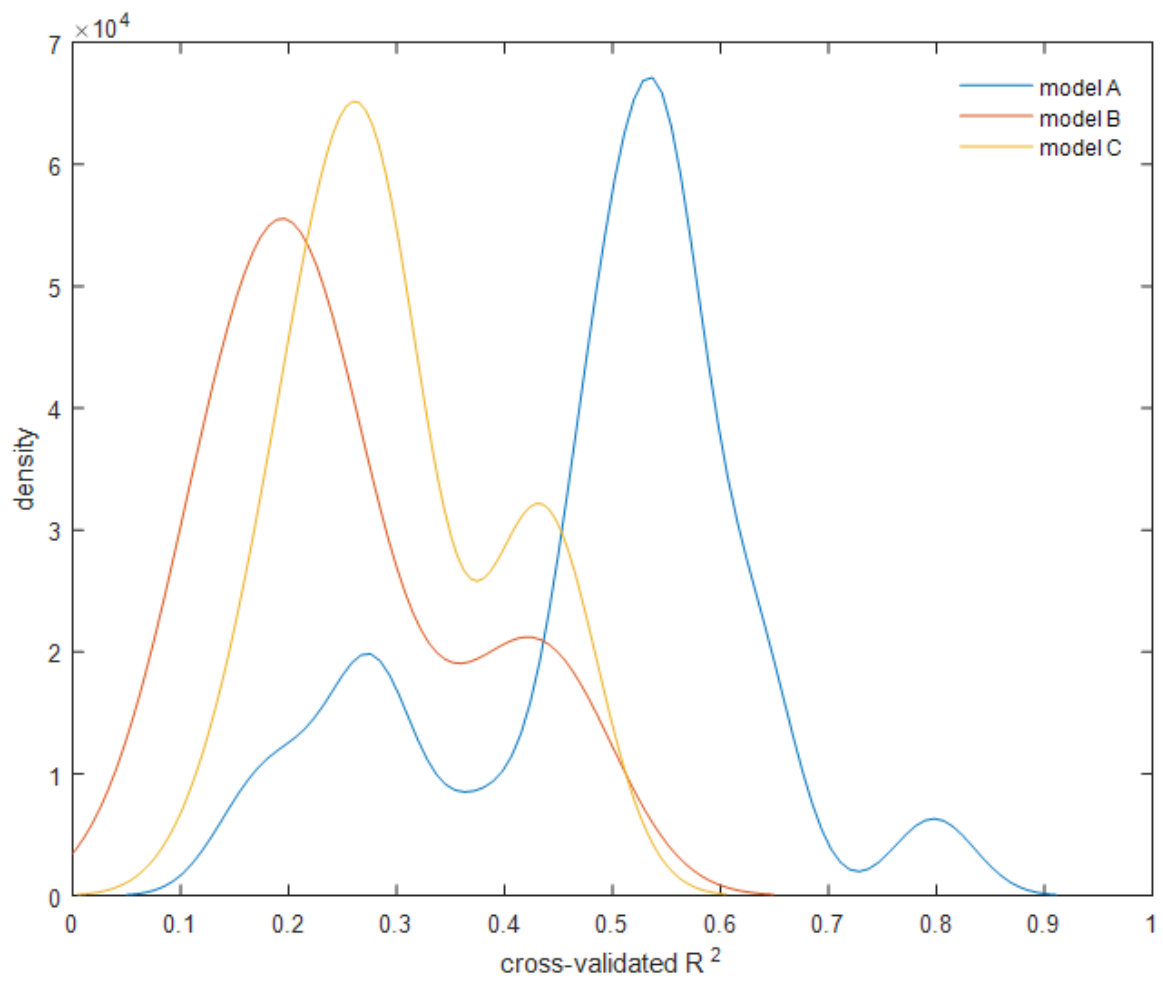
Table 2. Misspecification Tests

	misspecification test p-value
Model A	0.000
	(0.000)
Model B	0.032
	(0.001)
Model C	0.046
	(0.001)

Figure 2. Moderation effect and 95% confidence intervals (in elasticity form)
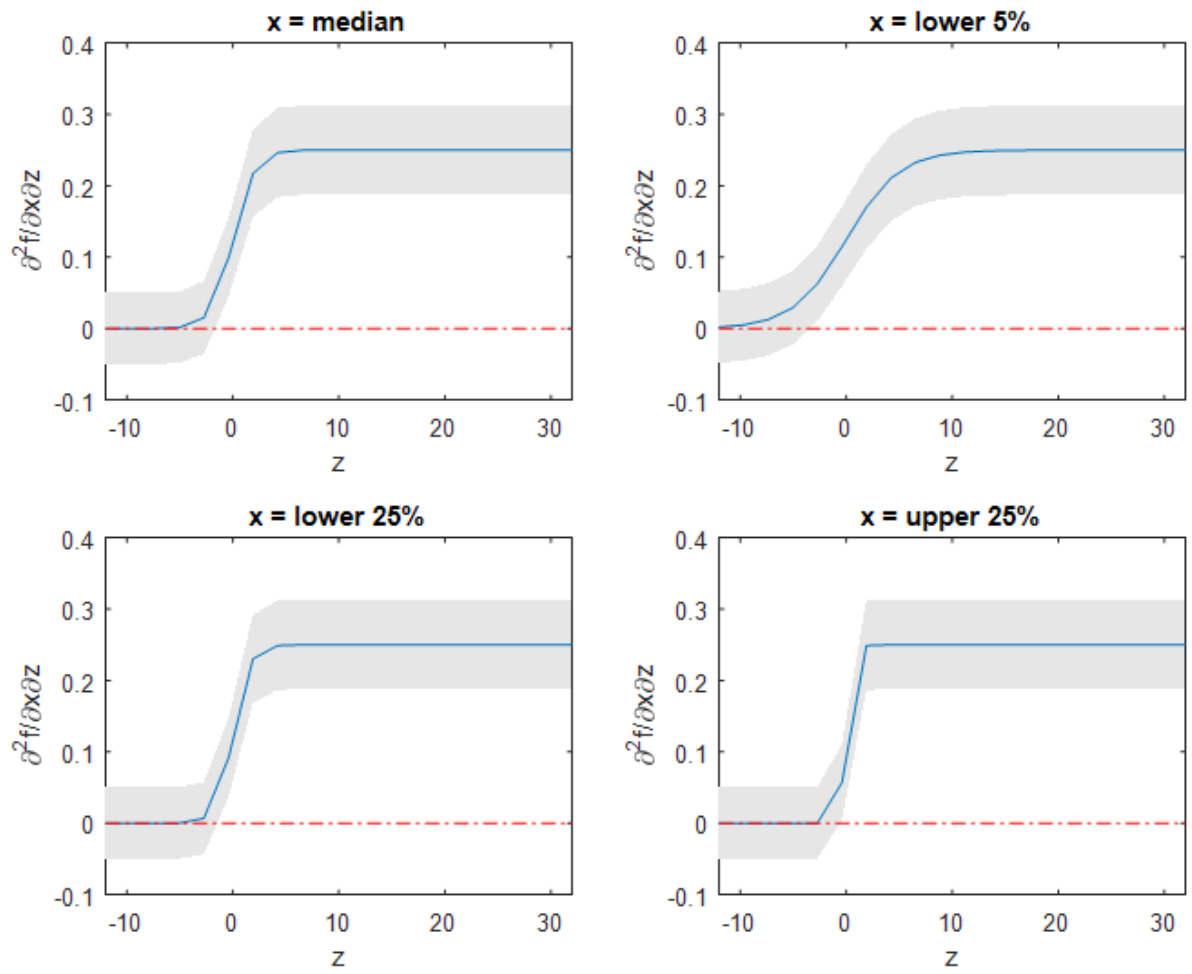
Notes: Elasticity is $\varepsilon = \left(\dfrac{\partial y}{\partial x}\right)\left(\dfrac{x}{y}\right)$.

7. References

Assaf, A. G., Josiassen, A., Ahn, J. S., & Mattila, A. S. (2017). Advertising spending, firm performance, and the moderating impact of CSR. *Tourism Economics*, 23(7), 1484-1495.

Assaf, A. G., & Tsionas, M. G. (2019). Revisiting shape and moderation effects in curvilinear models. *Tourism Management*, 75, 216-230.

Fan, J., & I. Gijbels (1995a). Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. *Journal of the Royal Statistical Society, Series B*, 57, 371–94.

Fan, J., & I. Gijbels (1995b). Local Polynomial Modelling and its Applications. London: Chapman & Hall.

Henderson, D. J., Carroll, R. J., & Li, Q. (2008). Nonparametric estimation and testing of fixed effects panel data models. *Journal of Econometrics*, 144(1), 257-275.

Henderson, D. J., & Ullah, A. (2005). A nonparametric random effects estimator. *Economics Letters*, 88(3), 403-407.

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.

Mammen, E., B. Støve., & D. Tjøstheim (2009). Nonparametric additive models for panels of time series. *Econometric Theory*, 25, 442–81.

Lee, S., Singal, M., & Kang, K. H. (2013). The corporate social responsibility–financial performance link in the US restaurant industry: do economic conditions matter?. *International Journal of Hospitality Management*, *32*, 2-10.

Luo, X., & de Jong, P. J. (2012). Does advertising spending really work? The intermediate role of analysts in the impact of advertising on firm value. *Journal of the Academy of Marketing Science*, 40(4), 605-624.

Rekker, S. A., Benson, K. L., & Faff, R. W. (2014). Corporate social responsibility and CEO compensation revisited: Do disaggregation, market stress, gender matter?. *Journal of Economics and Business*, 72, 84-103.

Ro, H. (2012). Moderator and mediator effects in hospitality research. *International Journal of Hospitality Management*, 31(3), 952-961.

Roberts, P. W., & Dowling, G. R. (2002). Corporate reputation and sustained superior financial performance. *Strategic management journal*, 23(12), 1077-1093.

Rodriguez-Poo, J. M., & Soberon, A. (2014). Direct semi-parametric estimation of fixed effects panel data varying coefficient models. *The Econometrics Journal*, *17*(1), 107-138.

Rodriguez-Poo, J. M., & Soberón, A. (2015). Nonparametric estimation of fixed effects panel data varying coefficient models. *Journal of Multivariate Analysis*, 133, 95-122.

Ruppert, D., & M. P. Wand (1994). Multivariate locally weighted least squares regression. *Annals of Statistics*, 22, 1346–70.

Su, L., & A. Ullah (2011). Nonparametric and semiparametric panel econometric models: estimation and testing. In A. Ullah and D. E. A. Giles (Eds.), *Handbook of Empirical Economics and Finance*, 455–97. New York: Taylor and Francis Group.

Sun, K. A., & Lee, S. (2018). Effects of franchising on industry competition: The moderating role of the hospitality industry. *International Journal of Hospitality Management*, 68, 80-88.

Ullah, A., & Wang, H. (2013). Parametric and nonparametric frequentist model selection and model averaging. *Econometrics*, *1*(2), 157-179.

Yoon, H., Song, S., Lee, S., & Kim, J. H. (2016). Does the restaurant type matter for investment in corporate social responsibility? *International Journal of Hospitality Management*, 58, 24-33.

Zhan-Qian, L. (1996). Multivariate locally weighted polynomial fitting and partial derivative estimation. *Journal of Multivariate Analysis*, 59, 187–205.