

ISSN 0819-2642
ISBN 978 0 7340 3725 1



THE UNIVERSITY OF MELBOURNE
DEPARTMENT OF ECONOMICS

RESEARCH PAPER NUMBER 1015

November 2007

A Reinterpretation of Interactions in Regressions

by

J. Hirschberg & J. Lye

Department of Economics
The University of Melbourne
Melbourne Victoria 3010
Australia.

A Reinterpretation of Interactions in Regressions

J. Hirschberg and J. Lye

Department of Economics
University of Melbourne

November 2007

Abstract

Regression specifications in applied econometrics frequently employ regressors that are defined as the product of two other regressors to form an interaction. Unfortunately, the interpretation of the results of these models is not as straight forward as in the linear case. In this paper, we present a method for drawing inferences for interaction models by defining the partial influence function. We present an example that demonstrates how one may draw new inferences by constructing the confidence intervals for the partial influence functions based on the traditional published findings for regressions with interaction terms.

Key Words: Interaction effects; dummy variables; linear transformation, Fieller method

JEL classification: C12; C51

Corresponding author, jnlye@unimelb.edu.au, Department of Economics, University of Melbourne, Melbourne, Victoria 3010, Australia.

1 Introduction

Regression models with interaction effects are often proposed to allow for the marginal effect of one explanatory variable to depend on another. The interpretation of the marginal effects depends on the choice of a value for the interacted regressors. We demonstrate that the graph of the partial influence function and its confidence interval can be used to both establish the level of the effect of these regressors as well as provide the appropriate tests of hypotheses. Section 2 defines the partial influence function. Section 3 describes how confidence intervals for the partial influence function are equivalent to the confidence bounds found using the Fieller technique and can be used to draw appropriate inference from a regression with interaction terms. An application of this technique using the information from a previously published paper is presented in Section 4. Our conclusions are presented in Section 5

2 The Partial Influence of Interacted Regressors

Consider the model:

$$\begin{aligned} \mathbf{y} &= \beta_1 + \beta_2 x + \beta_3 w + \beta_4 xw + \beta_5 z_1 + \cdots + \beta_k z_{k-4} + \boldsymbol{\varepsilon} \\ &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \end{aligned} \quad (1)$$

where \mathbf{y} is the $(T \times 1)$ vector of observations on the dependent variable, \mathbf{X} is the $(T \times k)$ matrix of explanatory variables consisting of a constant, x , w , xw , and other regressors $\sum_{j=5}^k z_{j-4}$ and $\boldsymbol{\varepsilon}$ is the $(T \times 1)$ vector of disturbance terms where

$\boldsymbol{\varepsilon} \sim (\mathbf{0}_{T \times 1}, \sigma^2 \mathbf{I}_{T \times T})$. The OLS estimators for the parameters are $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}$,

$\hat{\sigma}^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}/(T - k)$. Under the usual assumptions, the parameter estimates are asymptotically normally distributed according to $\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sigma^2 (\mathbf{X}'\mathbf{X})^{-1})$.

When x and w are both continuous variables, the marginal effect of a change in x is defined as:¹

$$\frac{\partial E(y|x,w)}{\partial x} = \beta_2 + \beta_4 w \quad (2)$$

If x is a dummy variable and w is a continuous variable, we define the difference in the regression equation when $x=1$ and when $x=0$ as:

$$\frac{\Delta E(y|x,w)}{\Delta x} = \beta_2 + \beta_4 w \quad (3)$$

The equations defined in (2) and (3) are the partial influence (PI) functions.

Typically when estimating a model such as (1) a researcher would test the significance of the interaction term using the t-test of $H_0 : \beta_4 = 0$, and a joint F -test of $H_0 : \beta_2 = \beta_4 = 0$. Alternatively, one could construct the partial influence function (as (2) or (3)) using the estimated values of the coefficients. However, using the partial influence requires the choice of an appropriate value of the other regressor (w) at which to evaluate it. One approach is to set the value of this other regressor to a particular value, such as the mean. An alternative approach is to determine the value of w at which $\frac{\partial E(y|x,w)}{\partial x} = 0$ or changes sign.

From (2), the value of w that results in $\frac{\partial E(y|x,w)}{\partial x} = 0$ is $w^* = -\frac{\beta_2}{\beta_4}$ and is estimated by $\hat{w}^* = -\frac{\hat{\beta}_2}{\hat{\beta}_4}$, where $\hat{\beta}_i$ are the OLS estimates of β_i .² In the next section we show how a confidence interval for w^* can be constructed using the Fieller method and how the bounds of this interval are related to the usual regression parameter hypothesis tests.

3 Ratios of Parameters in Linear Regression

The Fieller method (Fieller 1932, 1954) provides a general procedure for constructing confidence limits for statistics defined as ratios. Zerbe (1978) defines a version of Fieller's method in the regression context.

For $w^* = -\frac{\beta_2}{\beta_4}$, a $100(1-\alpha)\%$ confidence interval is determined by solving the quadratic equation $aw^2 + 2bw + c = 0$, where $a = \hat{\beta}_4^2 - t_{\alpha/2}^2 \hat{\sigma}_4^2$, $b = -\left(t_{\alpha/2}^2 \hat{\sigma}_{24} - \hat{\beta}_2 \hat{\beta}_4\right)$,

$c = \hat{\beta}_2^2 - t_{\alpha/2}^2 \hat{\sigma}_2^2$; $\hat{\sigma}_i^2$ are the estimated variances of $\hat{\beta}_i^2$ ($i=2,4$), $\hat{\sigma}_{24}$ is the estimated covariance between $\hat{\beta}_2$ and $\hat{\beta}_4$, and $t_{\alpha/2}^2$ is the $\alpha\%$ critical value of the t-distribution with $T-k$ degrees of freedom. The two roots of the quadratic equation $(w_L, w_U) = \frac{-b \pm \sqrt{b^2 - ac}}{a}$, define the confidence bounds of w^* . The parameters a , b and c are equivalent to: $a = \hat{\sigma}_4^2 (t_4^2 - t_{\alpha/2}^2)$, $b = -\hat{\sigma}_2 \hat{\sigma}_4 (rt_{\alpha/2}^2 - t_2 t_4)$ and $c = \hat{\sigma}_2^2 (t_2^2 - t_{\alpha/2}^2)$, where $t_i = \hat{\beta}_i / \hat{\sigma}_i$; $\hat{\sigma}_{24} = r \hat{\sigma}_2 \hat{\sigma}_4$ and r is the correlation between $\hat{\beta}_2$ and $\hat{\beta}_4$.

Using the expressions for a , b and c

$$(b^2 - ac) = t_{\alpha/2}^2 \hat{\sigma}_2^2 \hat{\sigma}_4^2 \left[(t_2 - rt_4)^2 + (1 - r^2)(t_4^2 - t_{\alpha/2}^2) \right]. \quad (4)$$

From (4) if $a > 0$, that is, $t_4^2 > t_{\alpha/2}^2$ then $b^2 - ac > 0$, which implies that one can reject $H_0 : \beta_4 = 0$ at the $\alpha\%$ level of significance. In this case a finite confidence interval can be constructed (see also Buonaccorsi 1979). Alternatively, the resulting confidence interval may be the complement of a finite interval when $((b^2 - ac) > 0, a < 0)$, or the interval may be the whole real line when $((b^2 - ac) < 0, a < 0)$.

It is also interesting to note that when $(a < 0)$ and

$$F > \frac{1}{2} F_{(1, T-k)} \quad (5)$$

where F is the value of the test statistic from the hypothesis: $H_0 : \beta_2 = \beta_4 = 0$, the resulting confidence interval will be a complement of a finite interval. (See also Scheffé 1970 and Zerbe 1982). This is the case when the coefficient on the interaction term is not significant (since $a < 0$ implies that $t_4^2 < t_{\alpha/2}^2$) but there may still be regions of w over which the partial influence function is significantly different from 0.

Hirschberg and Lye (2007) show that the confidence bounds can also be obtained by plotting the $100(1-\alpha)\%$ confidence bounds for the partial influence function defined in (2) and (3) defined by:

$$CI = (\hat{\beta}_2 + \hat{\beta}_4 w) \pm t_{\alpha/2} \sqrt{(\hat{\sigma}_2^2 + 2w\hat{\sigma}_{24} + w^2\hat{\sigma}_4^2)} \quad (6)$$

An estimate of the value of w^* (\hat{w}^*) satisfies the relationship $\hat{\beta}_2 + \hat{\beta}_4 \hat{w}^* = 0$. Similarly, the bounds defining the $100(1-\alpha)\%$ confidence interval on w^* are found by solving:

$$(\hat{\beta}_2 + \hat{\beta}_4 w)^2 - t_{\alpha/2}^2 (\hat{\sigma}_2^2 + 2w\hat{\sigma}_{24} + w^2\hat{\sigma}_4^2) = 0 \quad (7)$$

The solution for (7) is equivalent to the solution of the quadratic equation, $(aw^2 + bw + c = 0)$, where a , b and c are defined as above.

4 Class Sizes Example

Numerous studies have examined the effect of class size on student performance. Using interaction terms, the recent studies by Akerhielm (1995), Angrist and Lavy (1999), Bonesrønning 2003 and Wößmann and West (2006) have examined how class size effects on student performance differ across subgroups, such as gender, race, and socioeconomic class.

In this example, we follow Bonesrønning (2003) who estimates an education production function using individual student data for Norwegian children aged 13 to 16. To allow for possible gender differences an interaction between gender and class size is used. In estimation, predicted class size is used as an instrument for class size.

Bonesrønning finds that the interaction term is significant at the 10% level of significance (Table 2). In this case the partial influence function for the predicted performance difference between males and females is:

$$\frac{\Delta E(y|x,w)}{\Delta x} = -2.53 + 0.11 CS \quad (8)$$

where CS is class size. We find that this relationship changes sign when

$$CS^* = 2.53/0.11 = 23.$$

The upper and lower bounds of the Fieller interval corresponding to $\frac{\Delta E(y|x,w)}{\Delta x} = 0$ are

found by solving the equivalent expression to (7),

$$(-2.53 - 0.11CS)^2 - t_{\alpha/2}^2 \left((1.45)^2 + 2r(1.45)(0.06)CS + (0.11)^2 CS^2 \right) = 0. \quad (9)$$

In order to apply (9) we require a value for r the correlation between the parameter for *class size* as well as for the interaction of *class size* and *gender*. From the reported difference in the \bar{R}^2 between models with and without the interaction term we can estimate the value of the F -statistic for the joint test of the coefficients on *gender* and *gender* \times *class size* = 0. Although the model reported also includes a gender term, the \bar{R}^2 is appropriate to use as the gender t-stat equals 1. Then r is calculated from the formula for the F -statistic for two restrictions given by the following relationship based on the t-statistics of the two parameters:

$$F = \frac{1}{2} \left(\frac{(t_1^2 + t_2^2 - 2rt_1t_2)}{(1-r^2)} \right) \quad (10)$$

in this case we use the negative root $r = -0.996$. In simulations using similar data we have found that correlation between the coefficient for the interaction and the coefficients for the variables with which they are interacted are found to be negative in almost all circumstances.

Figures 1 and 2 plot the $\frac{\Delta E(y|x,w)}{\Delta x}$ function and the corresponding 90% and 95% confidence intervals. In Figure 1 we find that the 90% Fieller interval has finite bounds (13, 24). However, from Figure 2 we note that the 95% interval is a complement of a finite interval, with upper bound equal to 24. The 95% confidence interval indicates that female students are significantly affected by class sizes above 24, which contrasts with Bonesrønning (2003, pg 959) who concludes that "... females are unaffected by

class size”.. Similar results are obtained using the 90% confidence interval, which also indicates that there is a significant benefit for males for class sizes smaller than 13.

(Figures 1 and 2 here)

5. Conclusion

In this paper we have demonstrated that the partial influence function, defined when an interaction is present in a regression, can be examined by plotting the function at appropriate values of the interacted variable. The new material in this paper is the development of the relationship between the Fieller Method for the construction of confidence intervals for normally distributed random variables and the confidence intervals we can construct for the partial influence function.

We also demonstrate how it may be possible to construct a partial influence function and the corresponding confidence bounds for the influence of the interaction effect from previously published results. These bounds can then be used to consider new inferences that are not available from the traditional approach.

Acknowledgements

The research carried out in this paper was partially funded by a grant from the Faculty of Economics and Commerce at The University of Melbourne. We wish to thank David Moreton for helpful comments.

References

- Akerhielm, K. (1995) "Does Class Size Matter?", *Economics of Education Review*, 14, 229-241.
- Angrist, J. and V. Lavy (1999), "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement", *Quarterly Journal of Economics*, 114, 533-575.
- Bonesrønning, H. (2003), "Class Size Effects on Student Achievement in Norway: Patterns and Explanations", *Southern Economic Journal*, 69, 952-965.
- Buonaccorsi, J. P. (1979), "On Fieller's Theorem and the General Linear Model", *The American Statistician*, 33, 162.
- Fieller, E. C. (1932), "The Distribution of the Index in a Normal Bivariate Population", *Biometrika*, 24, 428-440.
- Fieller, E. C. (1954), "Some Problems in Interval Estimation", *Journal of the Royal Statistical Society. Series B*, 16, 175-185.
- Hirschberg, J. and J. Lye (2007), "Providing Intuition to the Fieller Method with two Geometric Representations using Stata and Eviews", Department of Economics Working Paper #952, University of Melbourne.
- Scheffé, H. (1970), "Multiple Testing versus Multiple Estimation. Improper Confidence Sets. Estimation of Directions and Ratios", *The Annals of Mathematical Statistics*, 41, 1-29.
- Zerbe, G., (1978), "On Fieller's Theorem and the General Linear Model", *The American Statistician*, 32, 103-105.
- Zerbe, G. (1982), "On Multivariate Confidence Regions and Simultaneous Confidence Limits for Ratios", *Communications in Statistics Theory and Methods*, 11, 2401-2425.

Wößmann, L. and M. West (2006), "Class-Size Effects in School Systems around the World: Evidence from between-Grade Variation in TIMSS", *European Economic Review*, 50, 695-736.

Footnotes

1. Note that a similar analysis can be performed on $\frac{\partial E(y|x,w)}{\partial w}$
2. We use the case of the continuous regressor w it can be show that these results are the same for the discrete case.

Figure 1: $\frac{\Delta E(y|x,w)}{\Delta x}$ and the 90% Confidence Bounds

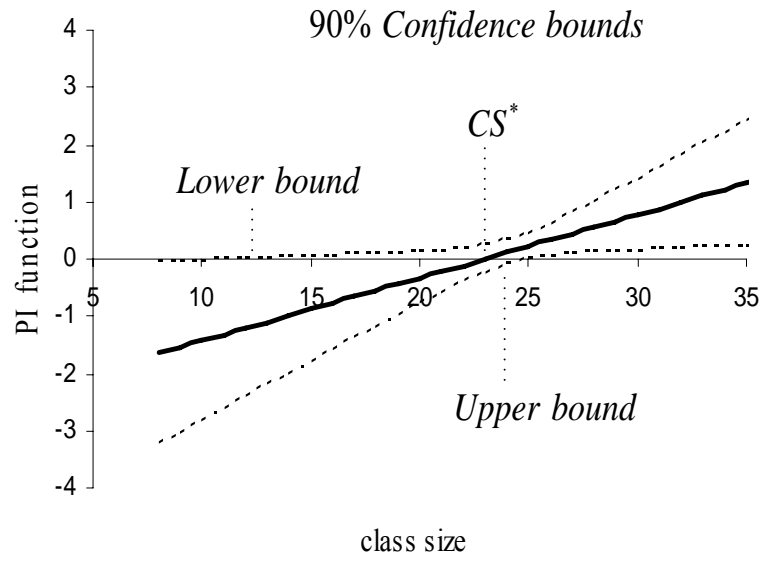


Figure 2: $\frac{\Delta E(y|x,w)}{\Delta x}$ and the 95% Confidence Bounds

