

*UCD CENTRE FOR ECONOMIC RESEARCH*

*WORKING PAPER SERIES*

*2007*

**Exploring nonlinearity with random field regression**

Derek Bond, University of Ulster,  
Michael J Harrison, University College Dublin and  
Edward J O'Brien, European Central Bank

WP07/17

November 2007

**UCD SCHOOL OF ECONOMICS  
UNIVERSITY COLLEGE DUBLIN  
BELFIELD DUBLIN 4**

---

# Exploring nonlinearity with random field regression<sup>\*</sup>

DEREK BOND, MICHAEL J. HARRISON AND EDWARD J. O'BRIEN<sup>†</sup>

*University of Ulster, Coleraine, Co. Londonderry, BT52 1SA, United Kingdom*

*School of Economics, University College Dublin, Belfield, Dublin 4, Ireland*

*European Central Bank, Kaiserstraße 29, D-60311 Frankfurt am Main, Germany*

---

## ABSTRACT

Random field regression models provide an extremely flexible way to investigate nonlinearity in economic data. This paper introduces a new approach to interpreting such models, which may allow for improved inference about the possible parametric specification of nonlinearity.

---

<sup>\*</sup>This paper is forthcoming in *Applied Economics Letters*.

<sup>†</sup>Corresponding author. Email: edward.obrien@ecb.int. The views expressed in this paper do not necessarily reflect those of the European Central Bank or its members.

## **I. Introduction**

Hamilton (2001) proposed a novel approach to nonlinear modelling that provides a single flexible parametric framework for testing for nonlinearity, drawing inference about its form and assessing the adequacy of resultant models. The methodology, based on the concept of random field regression, has been shown to have excellent potential; see Dahl (2002). From the practitioner's viewpoint, the importance of Hamilton's approach lies in the valuable insights it can provide for model construction.

By identifying variables that are nonlinearly significant, the form of nonlinearity *may* be inferred from a plot of a conditional expectation function against those variables. However, this approach may not be widely applicable: few studies have adopted this method.<sup>1</sup> Two significant practical problems are immediately apparent. If more than two variables are found to be nonlinearly significant, plotting the conditional expectation against these variables is not possible. Even if plotting is possible, it may not provide sufficient information to infer a specification.

This paper proposes a simple-to-implement approach to inference based on an alternative view of nonlinearity that exploits the flexible nonlinear framework. The results of empirical examples suggest the method holds much promise, offering a complementary view when the conditional expectation approach is viable, and a robust approach when it is not.

## **II. Flexible Nonlinear Inference**

Hamilton (2001) treats functional form as the outcome of a stochastic process that is part of the data-generating process; that is, the conditional expectation function of a regression model is thought of as being generated randomly prior to the generation of the data. This latent process is modelled using a Gaussian random field, the parameters of which are estimated by maximum likelihood.

If  $y_t$  is stationary,  $\mathbf{e}_t \sim N(0, \mathbf{S}^2)$  and  $\mathbf{x}_t$  is a  $k$ -vector of explanatory variables, then the model is

$$y_t = \mathbf{m}(\mathbf{x}_t) + \mathbf{e}_t, \quad (1)$$

where the form  $\mathbf{m}(\mathbf{x}_t)$  is unknown and assumed to be the outcome of a random field. Hamilton (2001) proposes a conditional mean function,  $\mathbf{m}(\mathbf{x}_t)$ , with the usual linear component and a stochastic, and hence unobservable, nonlinear component. The conditional mean function can be written as

$$\mathbf{m}(\mathbf{x}_t) = \mathbf{a}_0 + \mathbf{a}_1' \mathbf{x}_t + \mathbf{I} m(\mathbf{g} \circ \mathbf{x}_t), \quad (2)$$

where  $\mathbf{g}$  is a  $k$ -vector of parameters and  $\circ$  denotes the Hadamard product.

A simple method of testing for nonlinearity is to check if  $\mathbf{I}^2$  is zero and Hamilton (2001) proposes an Lagrange Multiplier-type test based on this. Inference hinges on estimation of the parameter vector  $\mathbf{g}$ , the estimated values of the elements of which are used to infer the form of nonlinearity. A highly significant  $g_i$ , where  $i = 1, \dots, k$ , suggests that the corresponding variable plays an important role in any nonlinearity. Hamilton suggests plotting the conditional expectation function  $\tilde{E}[\mathbf{m}(\mathbf{x}) | \mathbf{Y}_T]$ , where

$$\mathbf{Y}_T = (y_T, \mathbf{x}_T', y_{T-1}, \mathbf{x}_{T-1}', \dots, y_1, \mathbf{x}_1'), \quad (3)$$

against the significant nonlinear variables.

This may not always be possible, however, as plotting is limited to two significant explanatory variables. Huang and Lin (2006) present a case in point. In exploring Okun's relationship with flexible nonlinear inference (FNI), they find three significantly nonlinear explanatory variables, and resort to plotting the conditional expectation function against each significant variable. It is not surprising, therefore, that they fail to specify a nonlinear form, and only draw attention to the relative nonlinearities between pairs of variables.

### III. An Alternative Approach to Inference

The random field regression model consists of a linear and a nonlinear component,  $Im(\bullet)$ . Assuming that the error terms are small, a good approximation to the nonlinear term may be obtained as  $y_t - (\mathbf{a}_0 + \mathbf{a}'_1 \mathbf{x}_t)$ , from equations (1) and (2). To understand how the evaluation of the linear and nonlinear components may aid in inferring a nonlinear specification, consider the following. Assume that the linear component represents an approximation to the underlying data generating process and that the nonlinear component represents the deviation from the linear approximation to the true form, plus an error component. This view may be consistent with a range of nonlinear models. In this context, the linear term can be seen as a long-run component and the nonlinear term as a short-run dynamic component. Viewing FNI in this way could be important, as evaluating the linear term alone may represent a novel and satisfactory method of obtaining a linear approximation to a nonlinear model.

In terms of inference, viewing the random field regression in this way could be very instructive. Assuming the linear term is a good approximation, the best approach to inferring a nonlinear specification is to plot the nonlinear component against the explanatory variables. It is intuitive that the deviations from this approximation will capture information regarding the form of nonlinearity, such as structural breaks, turning points or thresholds. To illustrate this approach, three examples are considered.

### IV. Examples

Using a data generating process

$$y_t = 0.6x_{1t}1_{[x_{1t}>0]} + 0.2x_{2t} + \mathbf{e}_t, \quad (4)$$

where  $x_{it} \sim N(0,100)$  and  $\mathbf{e}_t \sim N(0,1)$ , Hamilton (2001) illustrated the effectiveness of FNI. To illustrate the proposed alternative approach, a fixed sample of  $x_{it}$  was generated, and the following random field regression was estimated:

$$\hat{y}_t = 4.75 + \underset{(1.10)}{0.31}x_{1t} + \underset{(0.06)}{0.20}x_{2t} + \underset{(0.01)}{1.03} \left[ \underset{(0.40)}{1.76}m \left( \underset{(0.006)}{0.07}x_{1t}, \underset{(0.002)}{2.1E-12}x_{2t} \right) \right]. \quad (5)$$

Only  $g_1$  is significant, suggesting that  $x_{1t}$  alone plays a significant role in the nonlinearity found in  $y_t$ .

Figure 1 plots  $\hat{l}_t = 4.75 + 0.31x_{1t} + 0.20x_{2t}$ , the linear component, and  $y_t$ . Clearly,  $\hat{l}_t$  is a good linear approximation to  $y_t$ . Rather than plotting  $\tilde{E}[\mathbf{m}(x_{1t}, \bar{x}_{2t}) | \mathbf{Y}_T]$  against  $x_{1t}$ , Figure 2 plots  $y_t - \hat{l}_t$  against  $x_{1t}$ ; that is, the deviation from the linear approximation is plotted against the nonlinearly significant explanatory variable. It can be seen that  $y_t - \hat{l}_t$  is linear in  $x_{1t}$ , with a threshold at  $x_{1t} = 0$ . These findings lead easily to the same conclusion as Hamilton (2001), i.e., that the data generating process for  $y_t$  is a threshold regression.

These results bear further consideration, however. From Figure 2, the deviation from  $y_t$  was estimated by OLS to be

$$\widetilde{y_t - \hat{l}_t} = \begin{cases} -4.74 - 0.31x_{1t}, & \text{for } x_{1t} < 0 \\ -4.75 + 0.29x_{1t}, & \text{for } x_{1t} > 0 \end{cases} \quad (6)$$

Examination of the plot of  $x_{1t}$  along with figures 1 and 2 shows that only values of  $x_{1t} > 0$  are relevant. This leads to

$$\begin{aligned} \hat{y}_t &= 4.75 + 0.31x_{1t} + 0.20x_{2t} - 4.75 + 0.29x_{1t}, \quad \text{for } x_{1t} > 0, \\ &= 0.60x_{1t}1_{[x_{1t}>0]} + 0.20x_{2t}, \end{aligned} \quad (7)$$

which is the underlying data generating process. While this is a straightforward example, it outlines, nevertheless, both the powerful nature of random field inference and the usefulness of the approach suggested here.

Hamilton's (2001) Example 3 concerning the US Phillips curve, was also revisited. Using data for the period 1949 to 1997, Hamilton estimates the following regression:

$$\hat{\mathbf{p}}_t = -88 - \underset{(127)}{0.92}u_t + \underset{(0.46)}{0.44}\mathbf{p}_{t-1} + \underset{(0.23)}{0.049}t + \underset{(0.07)}{1.24} \left[ \underset{(0.44)}{2.05}m \left( \underset{(1.29)}{0.14}u_t, \underset{(0.08)}{0.16}\mathbf{p}_{t-1}, \underset{(0.03)}{0.14}t \right) \right], \quad (8)$$

where  $p_t$  is the inflation rate and  $u_t$  is the unemployment rate in year  $t$ . This suggests that time is the only significant nonlinear variable. The examination of  $\tilde{E}[\mathbf{m}(\bar{u}_t, \bar{p}_t, t) | \mathbf{Y}_T]$  against  $t$  reveals evidence of distinct periods of high inflation. Using the alternative approach, i.e., plotting  $p_t - \hat{l}_t = p_t - (-88 - 0.92u_t + 0.44p_{t-1} + 0.049t)$ , leads to Figure 3. Three clear phases are seen when  $p_t - \hat{l}_t$  is plotted against  $t$ , the significant variable. These differ slightly from those reported by Hamilton, but the overall message is the same.

To illustrate how this approach could encapsulate a wide range of nonlinear models such as regime-switching and transition models, consider the results of Bond et al. (2007), relating to purchasing power parity (PPP) between Ireland and Germany.<sup>2</sup> They used random field regression to analyse a PPP specification including short-term interest rates and estimated the following random field regression:

$$\hat{s}_t = \underset{(1.12)}{0.77} - \underset{(0.15)}{0.84} p_t + \underset{(0.39)}{0.72} p_t^* - \underset{(0.002)}{0.0004} i_t + \underset{(0.01)}{0.01} i_t^* + \underset{(0.004)}{0.01} \left[ \underset{(2.55)}{5.86} m \left( \underset{(1.10)}{4.61} p_t, \underset{(3.02)}{16.97} p_t^*, \underset{(0.02)}{-0.03} i_t, \underset{(0.05)}{-0.15} i_t^* \right) \right], \quad (9)$$

where  $s_t$  is the log of the nominal exchange rate,  $p_t$  and  $p_t^*$  are the logs of the domestic and foreign price levels, and  $i_t$  and  $i_t^*$  are the respective short-term interest rates. Strong evidence of nonlinearity was found. The variables found to be nonlinearly significant were  $p_t$ ,  $p_t^*$  and  $i_t^*$ . Since three variables were found to be significant, the conditional expectation function could not be plotted.

Following the alternative approach, the linear or long-run term was found to be  $\hat{l}_t = 0.77 - 0.84p_t + 0.72p_t^* - 0.0004i_t + 0.01i_t^*$ . Figure 4 plots  $\hat{l}_t$ ,  $s_t$  and  $s_t - \hat{l}_t$  against time. The fit here is striking, but perhaps not surprising given the coefficient estimates, which are well within the confidence bounds of the (1 -1 1) parameters predicted by PPP theory. The deviation term shows breaks at 1979, 1987 and 1995, clearly related to important monetary developments in Ireland and Germany. Plotting  $s_t - \hat{l}_t$  against  $p_t$  and

$p_t^*$  reveals a similar pattern, with three or possibly four apparent regimes. Plotting  $s_t - \hat{l}_t$  against  $i_t^*$ , the remaining significant variable reveals no obvious pattern, however. The conclusion from this alternative view of FNI suggests that the PPP relationship between Ireland and Germany should be characterised by a regime-switching model, an approach successfully applied by Bond et al. (2007).

## **V. Conclusion**

This paper has introduced an alternative approach to inference in random field regression. Hamilton's flexible nonlinear framework is a powerful tool in nonlinear modelling and the approach to inference outlined here will be of benefit to those contemplating its use. Using three illustrative examples, the approach is shown to be both simple to implement and powerful. The results show that viewing nonlinearity in terms of linear and nonlinear components, in this context, may offer insights into the underlying data generating process, thereby facilitating the specification of an appropriate model.

## **Endnotes**

1. Notable exceptions include Hamilton (2003) and Kim et al. (2005).
2. For an excellent survey of issues relating to PPP, see Taylor and Taylor (2004) and Taylor (2006).

## References

- Bond, D., Harrison, M. J. and O'Brien, E. J. (2007) Modelling Ireland's exchange rates: from EMS to EMU, European Central Bank Working Papers, forthcoming.
- Dahl, C. M. (2002) An investigation of tests for linearity and the accuracy of likelihood based inference using random fields, *Econometrics Journal*, **5**, 263-284.
- Hamilton, J. D. (2001) A parametric approach to flexible nonlinear inference, *Econometrica*, **69**, 537-573.
- Hamilton, J. D. (2003) What is an oil shock?, *Journal of Econometrics*, **113**, 363-398.
- Huang, H.-C. and Lin, S.-C. (2006) A flexible nonlinear inference to Okun's relationship, *Applied Economics Letters*, **13**, 325-331.
- Kim, D. H., Osborn, D. R. and Sensier, M. (2005) Nonlinearity in the Fed's monetary policy rule, *Journal of Applied Econometrics*, **20**, 621-639.
- Taylor, A. M. and Taylor, M. P. (2004) The purchasing power parity debate, *Journal of Economic Perspectives*, **18**, 135-158.
- Taylor, M. P. (2006) Real exchange rates and purchasing power parity: mean-reversion in economic thought, *Applied Financial Economics*, **16**, 1-17.

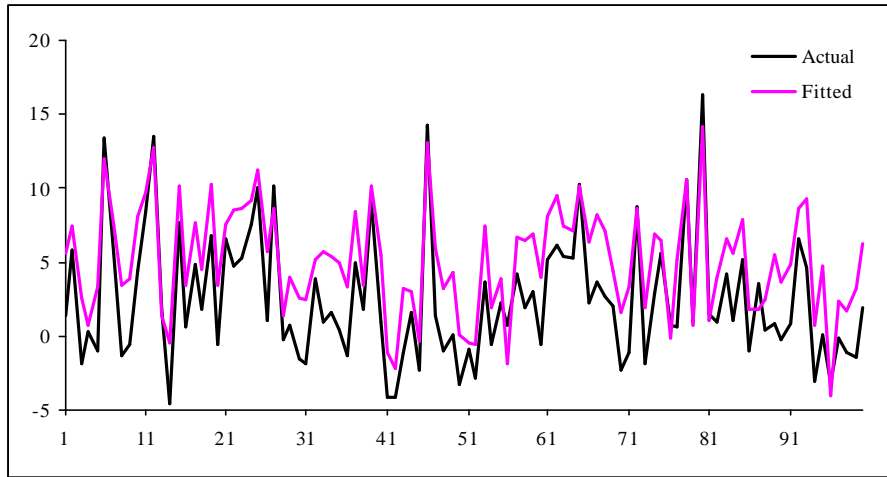


Figure 1: Plot of  $y_t$  and its linear approximation,  $\hat{l}_t$ .

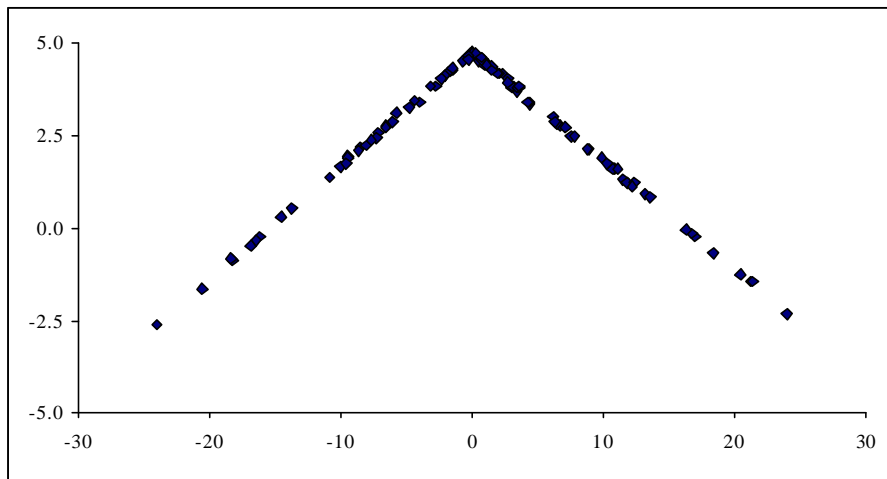


Figure 2: Plot of  $y_t - \hat{l}_t$  against  $x_t$ .

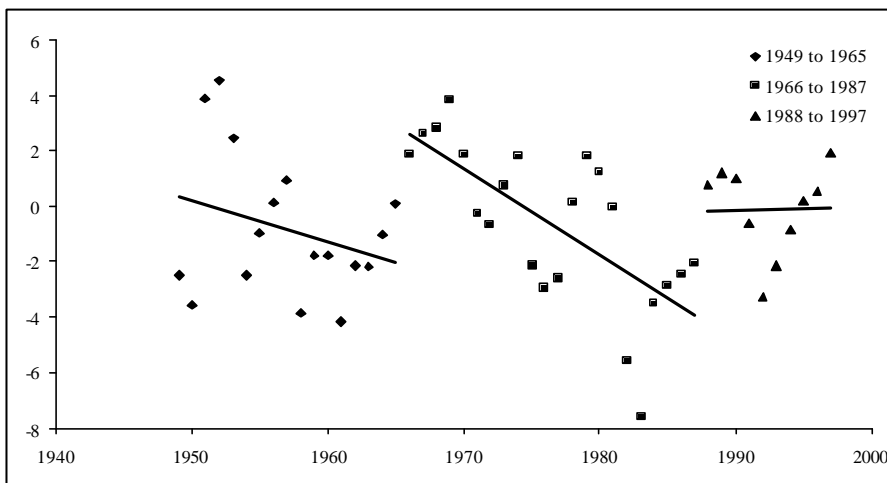


Figure 3: Plot of  $p_t - \hat{l}_t$  against  $t$ .

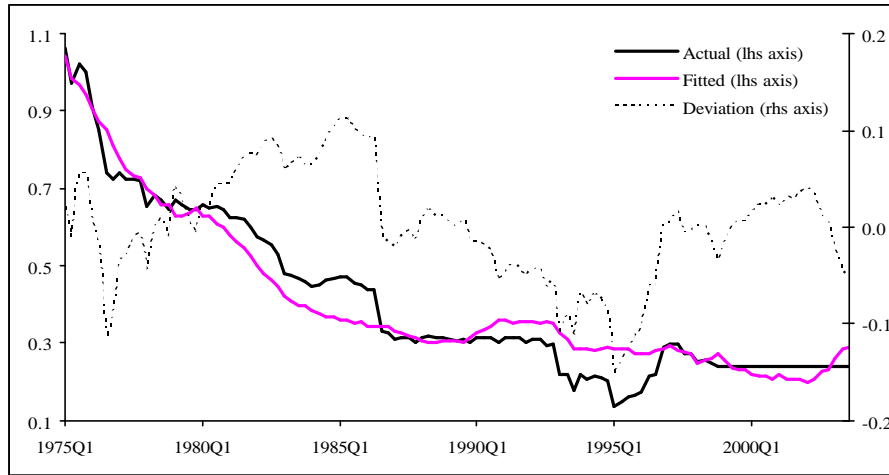


Figure 4: Plot of  $s_t$ ,  $\hat{l}_t$  and  $s_t - \hat{l}_t$  over time.