OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 72, 1 (2010) 0305-9049
doi: 10.1111/j.1468-0084.2009.00569.x

With or Without U? The Appropriate Test for a U-Shaped Relationship*

Jo Thori Lind and Halvor Mehlum

Department of Economics, University of Oslo, PB 1095 Blindern, 0317 Oslo, Norway (e-mail: j.t.lind@econ.uio.no; halvor.mehlum@econ.uio)

Abstract

Nonlinear relationships are common in economic theory, and such relationships are also frequently tested empirically. We argue that the usual test of nonlinear relationships is flawed and derive the appropriate test for a U-shaped relationship. Our test gives the exact necessary and sufficient conditions for the test of a U shape in finite samples in a large class of models.

I. Introduction

Economics is full of humps and U's. Famous examples of non-monotone relationships are the 'Laffer curve', the 'Kuznets curve (an inverted U between income and inequality, Kuznets, 1955), and the 'Environmental Kuznets curve' (inverted U between income and pollution, Selden and Song, 1994; Grossman and Krueger, 1995). In growth theory, poverty traps are generated if the growth rate of per capita capital stock first increases and then decreases with income (Nelson, 1956). In industrial organization it has been found that innovation is most intense at intermediate levels of competition (Aghion *et al.*, 2005). In political science, a finding is that countries with an intermediate level of democracy are more prone to war compared with both dictatorships and democracies (Hegre *et al.*, 2001). Finally, it has been argued that there is a hump-shaped relationship between union bargaining centralization and wage growth (Calmfors and Driffill, 1988), although few econometric studies of this exist, mostly as a result of the scarcity of data.

*While carrying out this research, the authors have been associated with the ESOP centre at the Department of Economics, University of Oslo. ESOP is supported by The Research Council of Norway. The authors are grateful for comments from two anonymous referees, Kalle Moene, and seminar participants at the University of Oslo and PRIO. They also thank Dustin Chambers for making his data available. A companion STATA module `utest` is available from: http://ideas.repec.org/c/boc/bocode/s456874.html
JEL Classification numbers: C12, C20.

For many of the examples shown, the existence of genuinely U-shaped (or hump-shaped) relationships have been the subject of intense debate. The debates take many forms and take issues with conceptual, empirical, and theoretical problems. There is one important common empirical question, however, that to the best of our knowledge has not been explicitly addressed anywhere in this diverse literature: namely, *given the estimates of a regression model, what is the test at some level* α *of the presence of a U shape?*

In most empirical work trying to identify U shapes, the researcher includes a nonlinear (usually quadratic) term in an otherwise standard regression model. If this term is significant and, in addition, the estimated extremum point is within the data range, it is common to conclude that there is a U-shaped relationship. We argue in this paper that this criterion is too weak. The problem arises when the true relationship is convex but monotone over relevant data values. A quadratic specification may then erroneously yield an extreme point and hence a U shape.

To test properly for the presence of a U shape on some interval of values, we need to test whether the relationship is decreasing at low values within this interval and increasing at high values within the interval. As the distribution of the estimated slope at any point is readily available, a test for the sign of the slope at any given point is straightforward. A test for a U shape gets more involved, however, as the null hypothesis is that the relationship is increasing at the left-hand side of the interval *and/or* is decreasing at the right-hand side. For this composite null hypothesis, standard testing methodology is no longer suitable. Fortunately, a general framework for such tests has been developed by Sasabuchi (1980).[1] We adopt his general framework to test for the presence of a U-shaped relationship. The extension to an inverse U shape is of course trivial.

Since 2001, there have been seven articles in the *American Economic Review* that use regression techniques to identify a U shape or an inverted U shape in their data. All of these revert to criteria that are intuitively sound, but potentially misleading. In this paper, we will present the appropriate test and then contrast it to the tests employed by the seven articles. In addition we provide a detailed example of the test in the case of a Kuznets curve.

II. A test of a U shape

To allow the regression to have a U shape, the standard approach has been to include a quadratic or an inverse term in a linear model. A more general formulation is:

$$y_i = \alpha + \beta x_i + \gamma f(x_i) + \xi' z_i + \varepsilon_i, \quad i = 1, \ldots, n. \tag{1}$$

Here x is the explanatory variable of main interest (e.g. income) whereas y is the variable to be explained (e.g. inequality), ε is an error term, and z is a vector of control variables. The known function f gives equation (1) a curvature and, depending on the

[1] Surprisingly, this contribution has so far not been applied in the econometrics literature with the exception of a few studies of economic inequality starting with Dardanoni and Forcina (1999).

parameters γ and β, equation (1) may be U-shaped or not. We assume that f is chosen so that the relationship has at most one extreme point. In that case the relationship is either hump-shaped, U-shaped, or monotone. In the following we focus on the case of testing for a U shape.[2]

Given equation (1) and the assumption of only one extreme point, the requirement for a U shape is that the slope of the curve is negative at the start and positive at the end of a reasonably chosen interval of x-values $[x_l, x_h]$. The natural choice of interval is in many contexts the observed data range [min(x), max(x)]. If we want to make sure that the inverse U shape is not only a marginal phenomenon the interval could also be in the interior of the domain of x. To assure at most one extreme point on $[x_l, x_h]$, as assumed before, we require f' to be monotone on this interval. A U shape is then implied by the conditions

$$\beta + \gamma f'(x_l) < 0 < \beta + \gamma f'(x_h). \tag{2}$$

If either of these inequalities are violated the curve is not U-shaped but inversely U-shaped or monotone.

To test whether the conditions in equation (2) are supported by the data, we need to test whether the combined null hypothesis

$$H_0 : \beta + \gamma f'(x_l) \geq 0 \text{ and/or } \beta + \gamma f'(x_h) \leq 0 \tag{3}$$

can be rejected in favour of the combined alternative hypothesis

$$H_1 : \beta + \gamma f'(x_l) < 0 \text{ and } \beta + \gamma f'(x_h) > 0. \tag{4}$$

Owing to the linearity of the specification (1) with respect to β and γ, the test of equation (3) vs. equation (4) is simply a test of linear restrictions on β and γ. The difficulty is that the test involves a set of inequality constraints. Hence the set of (β, γ) that satisfy H_1 is a sector in $\mathbb{R}^2$ contained between the two lines $\beta + \gamma f'(x_l) = 0$ and $\beta + \gamma f'(x_h) = 0$.

The test

Assuming that $\varepsilon_i \sim \text{NID}(0, \sigma^2)$, Sasabuchi (1980) shows that a test of H_0 in equation (3) based on the likelihood ratio principle takes the form

> Reject H_0 at the α level of confidence only if either,
> H_0^L or H_0^H or both can be rejected at the α level of confidence,

where H_0^L and H_0^H are the null hypotheses in the two standard one-sided tests

$$H_0^L : \beta + \gamma f'(x_l) \geq 0 \text{ vs. } H_1^L : \beta + \gamma f'(x_l) < 0,$$
$$H_0^H : \beta + \gamma f'(x_h) \leq 0 \text{ vs. } H_1^H : \beta + \gamma f'(x_h) > 0.$$

[2]The test for an inverse U is achieved by changing the sign on γ.

The rejection area is the convex cone

$$R_\alpha = \left\{ (\beta, \gamma) : \frac{\beta + \gamma f'(x_l)}{\sqrt{s_{11} + 2f'(x_l)s_{12} + f'(x_l)^2 s_{22}}} < -t_\alpha \right.$$
$$\left. \text{and} \quad \frac{\beta + \gamma f'(x_h)}{\sqrt{s_{11} + 2f'(x_h)s_{12} + f'(x_h)^2 s_{22}}} > t_\alpha \right\}, \tag{5}$$

where s_{11}, s_{22} and s_{12} are the estimated variances of β and γ and the covariance between them, whereas t_α is the α-level tail probability of the t-distribution with the appropriate degrees of freedom.

The test shown is also known as an intersection–union test as the null hypothesis is that the parameter vector is contained in a union of specified sets (see, e.g. Casella and Berger, 2002, ch. 8.2.3). Berger (1997) discusses the relationship between likelihood ratio tests and intersection–union tests and the conditions under which they coincide.

The two most common specifications of equation (1) is the quadratic form

$$y_i = \alpha + \beta x_i + \gamma x_i^2 + \xi' z_i + \varepsilon_i \tag{6}$$

and the inverse form

$$y_i = \alpha + \beta x_i + \gamma x_i^{-1} + \xi' z_i + \varepsilon_i. \tag{7}$$

In the first case, the presence of a U implies $\beta + 2\gamma x_l < 0$ and $\beta + 2\gamma x_h > 0$. A U shape in the second implies $\beta - \gamma x_l^{-2} < 0$ and $\beta - \gamma x_h^{-2} > 0$. In both cases, the test is easily carried out as two ordinary t-tests.[3]

The rejection area R_α can be manipulated so that it is expressed in terms of $f'(x_l)$ and $f'(x_h)$. Solving for $f'(x_l)$ and $f'(x_h)$ in R_α and inserting for $(\hat{\beta}, \hat{\gamma})$ yields the following two criteria for rejection:

$$f'(x_l) < \hat{\theta}_l \equiv \frac{s_{12}t_\alpha^2 - \hat{\beta}\hat{\gamma} - t_\alpha \sqrt{(s_{12}^2 - s_{22}s_{11})t_\alpha^2 + \hat{\gamma}^2 s_{11} + \hat{\beta}^2 s_{22} - 2s_{12}\hat{\beta}\hat{\gamma}}}{\hat{\gamma}^2 - s_{22}t_\alpha^2}, \tag{8}$$

$$f'(x_h) > \hat{\theta}_h \equiv \frac{s_{12}t_\alpha^2 - \hat{\beta}\hat{\gamma} + t_\alpha \sqrt{(s_{12}^2 - s_{22}s_{11})t_\alpha^2 + \hat{\gamma}^2 s_{11} + \hat{\beta}^2 s_{22} - 2s_{12}\hat{\beta}\hat{\gamma}}}{\hat{\gamma}^2 - s_{22}t_\alpha^2}. \tag{9}$$

Here $\hat{\theta}_l$ and $\hat{\theta}_h$ are the roots of the same quadratic equation. This way of formulating the criterion has a good parallel to the work of Fieller (1954). From his work, it is known how to construct an exact confidence interval for the ratio of two normally distributed estimates. Note first that the estimated extreme point of equation (1) is

[3]A routine to perform the test in these two cases in the software package Stata is available from: http://ideas.repec.org/c/boc/bocode/s456874.html.

$$f'(\hat{x}^{\min}) = -\frac{\hat{\beta}}{\hat{\gamma}}.$$

Following Fieller (1954), a $(1-2\alpha)$ confidence interval for $-\beta/\gamma$ is given by $[\hat{\theta}_l, \hat{\theta}_h]$ as defined in equations (8) and (9). Therefore, a $(1-2\alpha)$ confidence interval for $\hat{x}^{\min}$ is $[\tilde{x}_l, \tilde{x}_h]$ where $f'(\tilde{x}_i) = \hat{\theta}_i$. To perform the test of equation (3) vs. equation (4) at the α-level of significance is then equivalent to seeing whether the $(1-2\alpha)$ confidence interval for $\hat{x}^{\min}$ is inside the data range, $[\tilde{x}_l, \tilde{x}_h] \subset [x_l, x_h]$.

To find a confidence interval for $\hat{x}^{\min}$, one could also use the delta method. For finite samples, however, this may be severely biased.[4] Also when using the delta method the $(1-2\alpha)$ interval is the proper interval to use when testing for a U shape at the α-level.

The testing strategy is easily carried over to more involved estimations. First, it is readily extended to a more general formulation like $y_i = \alpha + \sum_{j=1}^{H} \beta_j f_j(x) + z_i'\gamma + \varepsilon_i$, where f_j is a set of known functions.[5] The appropriate test of the presence of an inverse U-shaped relationship between x and y is now

$$\begin{aligned} &H_0: \sum \beta_j f_j'(x_l) \le 0 \text{ and/or } \sum \beta_j f_j'(x_h) \ge 0 \\ \text{vs.} \\ &H_1: \sum \beta_j f_j'(x_l) > 0 \text{ and } \sum \beta_j f_j'(x_h) < 0. \end{aligned} \tag{10}$$

The test is also applicable to studies of U-shaped relationships in the full class of generalized linear models, encompassing most models of limited dependent variables. As the estimated parameters are asymptotically jointly normally distributed, the distribution of the test is asymptotic as explained before.

Performance of the test

To see the performance of the test, Table 1 shows the results of some Monte Carlo analyses. Two specifications are employed. First, data are generated with $x \sim U[0,1]$, $\varepsilon \sim N(0, 1/2)$ and $y = a + bx + x^2 + \varepsilon$, where b is chosen to achieve the desired minimum ranging from 0.6 to 1.6. We then regress y on x and x^2 and perform three tests: (i) whether the quadratic term is significant; (ii) whether the quadratic term is significant and the estimated extremum point is in $[0, 1]$; and (iii) the test described before. The table gives the fraction of cases where we reject the absence of a U-shaped

[4]This problem is clearly illustrated in Hirschberg and Lye (2005) who study different methods for constructing $(1-\alpha)$ intervals for the extremum of a quadratic equation and give practical recommendations consistent with our suggestions.

[5]The difficulty with this general specification is that although the relationship is decreasing at the left-hand side of the relevant interval and increasing at the right-hand side, it may not be a U-shaped relationship inside, but instead for instance an W-shaped relationship. It is then necessary to perform a joint test of the whole shape of the relationship, which is much more complicated. See Doveh, Shapiro and Feigin (2002) for some developments in this direction.

TABLE 1

Monte Carlo analysis of the performance of the test

		Number of observations									
		Correctly specified model					*Mis-specified model*				
Extreme point		10	50	100	500	1,000	10	50	100	500	1,000
0.6	(i)	0.11	0.18	0.32	0.91	1.00	0.12	0.26	0.45	0.98	1.00
	(ii)	0.11	0.18	0.32	0.91	1.00	0.12	0.26	0.45	0.98	1.00
	(iii)	0.10	0.18	0.30	0.82	0.97	0.13	0.30	0.49	0.98	1.00
0.8	(i)	0.11	0.18	0.31	0.91	1.00	0.12	0.26	0.45	0.98	1.00
	(ii)	0.11	0.18	0.31	0.89	0.97	0.12	0.26	0.45	0.98	1.00
	(iii)	0.09	0.11	0.14	0.36	0.57	0.11	0.20	0.31	0.83	0.98
1	(i)	0.11	0.19	0.31	0.91	1.00	0.12	0.26	0.44	0.98	1.00
	(ii)	0.11	0.19	0.31	0.51	0.50	0.12	0.26	0.44	0.93	0.98
	(iii)	0.07	0.06	0.05	0.05	0.05	0.09	0.13	0.17	0.45	0.68
1.2	(i)	0.11	0.19	0.32	0.91	1.00	0.12	0.25	0.45	0.98	1.00
	(ii)	0.11	0.18	0.26	0.10	0.03	0.12	0.25	0.45	0.81	0.90
	(iii)	0.05	0.02	0.01	0.00	0.00	0.08	0.09	0.11	0.23	0.36
1.4	(i)	0.11	0.18	0.31	0.91	1.00	0.13	0.25	0.45	0.98	1.00
	(ii)	0.11	0.16	0.13	0.00	0.00	0.13	0.25	0.44	0.71	0.78
	(iii)	0.04	0.01	0.00	0.00	0.00	0.08	0.07	0.08	0.13	0.19
1.6	(i)	0.12	0.18	0.32	0.91	1.00	0.13	0.25	0.45	0.99	1.00
	(ii)	0.11	0.11	0.05	0.00	0.00	0.13	0.25	0.44	0.63	0.68
	(iii)	0.04	0.00	0.00	0.00	0.00	0.08	0.06	0.07	0.09	0.12

Notes: Tests performed at the 5% level of significance. Results are based on 10,000 simulations. The tests studied are: (i) whether the quadratic term is significant; (ii) whether the quadratic term is significant and the estimated extremum point is in [0, 1]; and (iii) the test described previously.

relationship. To see how the different tests perform in a mis-specified model, the right-hand part of the table shows simulations from a true model $y = a + bx + 1/x + \varepsilon$, but where we still regress y on x and x^2. Now we have[6] $x \sim U[0.5; 1]$ and $\varepsilon \sim N(0, 1/2)$.

In the correctly specified model, test (i) has a huge rate of Type I errors, erroneously rejecting the absence of a U-shaped relationship. This indicates that this is an inappropriate test. When it is also required that the estimated extremum be in the relevant interval as in (ii), the test is stronger, particularly for large samples where this point estimate becomes precise. However, the proper test (iii) performs better in all cases. Notice, though, that it has somewhat lower test power than test (ii) when there actually is a U-shaped relationship. In the mis-specified model, both tests (i) and (ii) perform badly. There is some excess probability of committing Type I errors for our test as well, but the outcomes are much better than the commonly used procedures.

[6]The smaller interval for x was chosen as the asymptote $y \to +\infty$ when $x \to 0$ would have a too large impact on the estimates if x was allowed to be too small. In this case, all three procedures perform badly (results available upon request).

III. A comparison with applied work

We have read through a large number of applied econometrics works where U shapes are identified using equation (6). Most authors focus on the significance and sign of both $\hat{\beta}$ and $\hat{\gamma}$. How does the significance of these parameters relate to the test in equation (5)?

If the data range is the whole of $\mathbb{R}$ the significance of $\hat{\gamma}$ (with right sign) is necessary and sufficient for rejecting H_0 . If the data range is any subset of $\mathbb{R}$, the significance of $\hat{\gamma}$ is still necessary but not sufficient.

If the data range is exactly $\mathbb{R}^+$ or $\mathbb{R}^-$ the significance (with the right sign) of $\hat{\beta}$ and $\hat{\gamma}$ respectively is necessary and sufficient for rejecting H_0. If the data range is a subset of $\mathbb{R}^+$ (or of $\mathbb{R}^-$) the individual significance of both $\hat{\beta}$ and $\hat{\gamma}$ is necessary but not sufficient.

If the data range is neither $\mathbb{R}$, $\mathbb{R}^+$ or $\mathbb{R}^-$, the necessary and sufficient conditions are only found using equation (5). Obviously, a simple first check is whether the estimated minimum point ($\hat{x}^{\min} = -\hat{\beta}/(2\hat{\gamma})$) itself is within the date range.

Most works use the criterion that if both $\hat{\beta}$ and $\hat{\gamma}$ are significant and if the implied extreme point is within the data range, they have found a U. This is a sensible criterion but it is neither sufficient nor necessary. It is insufficient as the estimated extreme point may be too close, given the uncertainty, to an end point of the data range. It is not generally necessary as β may be zero if the data range extends to both sides of $x = 0$.

A small number of works use the delta method to calculate the standard deviation of the extreme point. This method, although sound, is only reliable with a sufficiently large number of observations.

We have in particular looked at articles in the *American Economic Review*. There are seven articles since 2001 that use regression techniques to identify a U shape. All use formulations similar to equation (6).

The most common procedure is the one followed by Abiad and Mody (2006), McKinnish (2004), and Kalemli-Ozcan, Sørensen and Yosha (2003). They find that both $\hat{\beta}$ and $\hat{\gamma}$ in the quadratic specification have the right sign and are individually significant. Based on this, they all conclude that they have found an inverted U shape without checking whether the implied extreme point is within any relevant range.

Sigman (2002) follows the same procedure, but she also calculates the extreme point and finds it to be within the data range. She writes, relating to $\hat{\beta}$ and $\hat{\gamma}$ (p. 1157): 'The GDP coefficients always appear to follow an inverted U shaped pattern, but are jointly statistically significant at 5 percent only in columns (3) and (4)'. Her testing for joint significance of $\hat{\beta}$ and $\hat{\gamma}$, however, do not add much when looking for a U shape. As we have argued before, the significance of $\hat{\gamma}$ alone is always a necessary condition in the test of a U shape.

Similarly, Aghion *et al.* (2005) use their estimates from the quadratic specification to calculate predicted value $\hat{y}$. They find that the estimated minimum is inside the x-range, but they do not explicitly assess the significance of the result.

Carr, Markusen and Maskus (2001) only include γx^2, and not βx, in their specification. As their x-range includes both negative and positive values, their specification rules out monotonically increasing or decreasing relationships between x and y. Hence their analysis does not allow for any meaningful test of a U shape against a monotone relationship.

Imbs and Wacziarg (2003) first use non-parametric techniques to find the relationship between per capita income x and sectoral concentration y and they find a U shape. They use this technique also to construct 95% intervals for the minimum point. They then turn to a parametric specification with x and x^2 in the right-hand side. Both coefficients have the right sign and are individually significant. The significance of the U shape is assessed by plotting x and $\hat{y}$ vs. the scatter plot of the data. The regression curve fits the data very well and based on the plot alone one can safely conclude that the U shape is significant at all conventional levels. But, the precise significance level is unclear and so is the precision of the estimated minimum point.

IV. Illustration

As a concrete illustration of our methodology, we use a recent study by Chambers (2007). The paper contains a standard Kuznets regression between gross domestic product (GDP) and inequality. The data are an unbalanced panel of 29 countries giving a total of 232 observations. He has information on the level of inequality measured by the Gini coefficient, log of purchasing power parity adjusted GDP per capita, and a set of control variables (see Chambers, 2007, for details).

We replicate and extend his analysis. Results from the regression analysis, test statistics from the appropriate U test, and the derived Fieller interval are reported in Table 2. The conventional method would simply test the significance of $\hat{\beta}$ and $\hat{\gamma}$, which

TABLE 2

Estimates of the Kuznets curve

Dependent variable: Gini index		
log per capita GDP (X_i)	$\hat{\beta}=$	32.27
		(11.63)***
log per capita GDP-squared (X_i^2)	$\hat{\gamma}=$	−1.88
		(0.65)***
Slope at X_l	$\hat{\beta}+2\hat{\gamma}X_l=$	8.15
		(3.65)**
Slope at X_h	$\hat{\beta}+2\hat{\gamma}X_h=$	−4.33
		(2.34)*
Appropriate U test		1.85
		[0.033]
Extremum point	$-\hat{\beta}/(2\hat{\gamma})=$	8.60
90% confidence interval, Fieller method		[7.44, 9.58]
90% confidence interval, Delta method		[7.73, 9.47]

Notes: Robust standard errors in parentheses and P-values in square brackets. ***, ** and * denote significances at the 1, 5, and 10% levels.
GDP, gross domestic product.

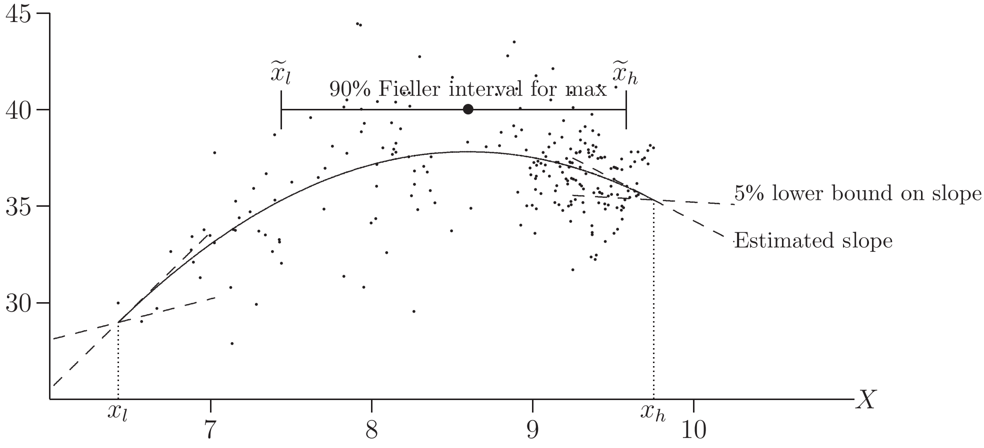

Figure 1. Illustrative Kuznets curve

in this case yield t-values of 2.77 and 2.89. With the appropriate U test, however, we see that the test for a positive slope at x_h yields a t-value of only 1.85, and hence a P-value of 0.032 with the required one-sided test.

Figure 1 illustrates the estimated relationship, the confidence interval for the maximum and the two lower bounds on the slopes in each endpoint.[7] We see that the turning point of the relationship is quite close to x_h, and that the slope of the curve at x_h is negative, but not very steep and only just significant at the 5% level. Hence there is a significant hump-shaped relationship over the range of the data, but the significance of this relationship is weaker than what would be detected by traditional approaches.

V. Conclusion

In this paper, we have provided an appropriate test of a U-shaped relationship in a regression model. In the applied econometrics literature a large number of articles try to identify non-monotone relationships using regression analysis. Any of these articles hardly use the adequate formal procedures when they test for the presence of a U shape. To the best of our knowledge none of them have used the simple test that we are suggesting. Most works, nevertheless, seem to be on fairly safe ground when they claim to have found a U shape. The reason is that the *de facto* common practice seems to be to check two necessary conditions, namely that the second derivative has the right sign and that the extremum point is within the data range. However, only the results from the former are usually reported. This criterion will be misleading, however, if the estimated extremum point is too close to the end point of the data range. Our test gives the exact necessary and sufficient conditions for the test of a U shape. In addition, the interval interpretation provides a confidence interval for the extremum point.

Final Manuscript Received: June 2009

[7]We have used average values for all the controls.

References

Abiad, A. and Mody, A. (2006). 'Financial reform: what shakes it? what shapes it?' *American Economic Review*, Vol. 95, pp. 66–88.

Aghion, P., Bloom, N., Blundell, R., Griffith, R. and Howitt, P. (2005). 'Competition and innovation: an inverted-U relationship', *Quarterly Journal of Economics*, Vol. 120, pp. 701–728.

Berger, R. L. (1997). 'Likelihood ratio tests and intersection–union test', Ch. 15 in Panchapakesan S. and Balakrishnan N. (eds), *Advances in Statistical Decision Theory and Application*, Birkhäuser, Boston, pp. 225–237.

Calmfors, L. and Driffill, J. (1988). 'Centralization of wage bargaining', *Economic Policy*, Vol. 6, pp. 14–61.

Carr, D. L., Markusen, J. R. and Maskus, K. E. (2001). 'Estimating the knowledge-capital model of the multinational enterprise', *American Economic Review*, Vol. 91, pp. 693–708.

Casella, G. and Berger, R. L. (2002). *Statistical Inference*, 2nd edn, Duxbury Press, Pacific Grove, CA.

Chambers, D. (2007). 'Trading places: does past growth impact inequality?' *Journal of Development Economics*, Vol. 82, pp. 257–266.

Dardanoni, V. and Forcina, A. (1999). 'Inference for Lorenz curve orderings', *Econometrics Journal*, Vol. 2, pp. 49–75.

Doveh, E., Shapiro, A. and Feigin, P. D. (2002). 'Testing of monotonicity in parametric regression models', *Journal of Statistical Planning and Inference*, Vol. 107, pp. 289–306.

Fieller, E. C. (1954). 'Some problems in interval estimation', *Journal of the Royal Statistical Society, Series B*, Vol. 16, pp. 175–185.

Grossman, G. M. and Krueger, A. B. (1995). 'Economic growth and the environment', *Quarterly Journal of Economics*, Vol. 110, pp. 353–377.

Hegre, H., Ellingsen, T., Gates, S. and Gleditsch, N. P. (2001). 'Toward a democratic civil peace? Democracy, political change, and civil war, 1816–1992', *American Political Science Review*, Vol. 95, pp. 17–33.

Hirschberg, J. and Lye, J. (2005). 'Inferences for the extremum of quadratic regression models', Research Paper 906, Department of Economics, University of Melbourne.

Imbs, J. and Wacziarg, R. (2003). 'Stages of diversification', *American Economic Review*, Vol. 93, pp. 63–86.

Kalemli-Ozcan, S., Sørensen, B. E. and Yosha, O. (2003). 'Risk sharing and industrial specialization: regional and international evidence', *American Economic Review*, Vol. 93, pp. 903–918.

Kuznets, S. (1955). 'Economic growth and income inequality', *American Economic Review*, Vol. 45, pp. 1–28.

McKinnish, T. G. (2004). 'Gender in policy and the labor market occupation, sex-integration, and divorce', *American Economic Review*, Vol. 94, pp. 322–325.

Nelson, R. R. (1956). 'A theory of the low-level equilibrium trap in underdeveloped economies', *American Economic Review*, Vol. 46, pp. 894–908.

Sasabuchi, S. (1980). 'A test of a multivariate normal mean with composite hypotheses determined by linear inequalities', *Biometrika*, Vol. 67, pp. 429–439.

Selden, T. M. and Song, D. (1994). 'Environmental quality and development: is there a Kuznets curve for air pollution emissions?' *Journal of Environmental Economics and Management*, Vol. 27, pp. 147–162.

Sigman, H. (2002). 'International spillovers and water quality in rivers: do countries free ride?' *American Economic Review*, Vol. 92, pp. 1152–1159.