



THE UNIVERSITY OF
MELBOURNE

Department of Economics

Working Paper Series

Inverse Test Confidence Intervals for Turning points: A Demonstration with Higher Order Polynomials

J. N. Lye and J. G. Hirschberg

Sept 2012

Research Paper Number 1160

ISSN: 0819 2642

ISBN: 978 0 7340 4511 9

Department of Economics
The University of Melbourne
Parkville VIC 3010
www.economics.unimelb.edu.au

Inverse Test Confidence Intervals for Turning-points: A Demonstration with Higher Order Polynomials.¹

J. N. Lye and J. G. Hirschberg²

September 2012

Abstract:

In this paper we demonstrate the construction of inverse test confidence intervals for the turning points in estimated nonlinear relationships by the use of the marginal or first derivative function. First, we outline the inverse test confidence interval approach. Then we examine the relationship between the traditional confidence intervals based on the Wald test for the turning-points for a cubic, a quartic and fractional polynomials estimated via regression analysis and the inverse test intervals. We show that the confidence interval plots of the marginal function can be used to estimate confidence intervals for the turning points that are equivalent to the inverse test. We also provide a method for the interpretation of the confidence intervals for the second derivative function to draw inferences for the characteristics of the turning-point.

This method is applied to the examination of the turning points found when estimating a quartic and a fractional polynomial from data used for the estimation of an Environmental Kuznets Curve. The *Stata* do files used to generate these examples are listed in the appendix along with the data.

Key words: Inverse Test Confidence Intervals, Likelihood Profile, Quartic, Fractional Polynomials, *Stata*, Marginal Effect Function, Fieller Method.

¹ We would like to thank the anonymous referee for a number of comments that proved very helpful in the preparation of this version of the paper. We also wish to thank participants at the Econometric Society Australasian Meeting held in July of 2012. This research is supported by research funds supplied by the Faculty of Business and Economics, University of Melbourne. The responsibility for any errors is entirely ours.

² Department of Economics, University of Melbourne, Melbourne, Vic 3010, Australia.

1. Introduction

Regression models in many applications employ a functional specification which allows for a nonlinear relationship between a regressor and the dependent variable. The most common of these is the quadratic specification with a wide application of more complex higher order polynomials and semi-parametric methods where part of the specification is linear and part is allowed to be flexible functions of x . Typically these models are of the form:

$$y_t = \beta(x_t) + \left(\alpha_0 + \sum_{j=1}^M \alpha_j z_{jt} \right) + \varepsilon_t$$

In this specification we assume that the nonlinear specification is in the variable x while the other regressors (z) are assumed to have a linear influence on the dependent variable.³

An important characteristic of these models is that they possess the property of a non-constant marginal effect of the x variable over the values of x . In this paper we demonstrate that the confidence intervals of the marginal effect function with respect to x can be used for inferences concerning the turning-points (sometimes referred to as the stationary points). We also show that these intervals are equivalent to a class of previously defined inferential methods. In addition we demonstrate how the marginal of the marginal effects function, or more directly the second order function can be used for inferences in high order models. In this discussion we follow the implications of the consideration of the derivative functions implied by these nonlinear specifications.

³ In this discussion we assume that there is only one variable to be entered in a nonlinear form. In most cases the extension to more than one can be accommodated in a similar manner to the techniques we discuss here.

Marginal effects or derivative functions of nonlinear regression specifications have a significant history in the interpretation of regression results. The interpretation of the marginal effects is central to many econometric applications and is extensively covered in texts such as Wooldridge (2010, chapter 2). Recently, the statistics package *Stata* has introduced a series of post-estimation procedures to compute marginal effects and functions based on the parameters estimated from a wide variety of models. The original version of the currently available *margins* routine was a user written routine as described in Bartus (2005).

An early exposition of the marginal effects for polynomial regressions can be found in Williams (1959, chapter 6), however his treatment has been rarely referenced and it is only recently that the estimation of high order polynomials has become widespread. In the case of the quadratic specification the location of the maximum or minimum is defined as $-1/2$ times the ratio of the parameter on the linear term divided by the parameter on the squared term. Thus to draw inferences concerning the location of the turning-point requires the construction of confidence intervals for the ratio of regression parameters. However, Dufour (1997) in his survey of “Impossibility Theorems in Econometrics” observes that the direct application of the Wald test to the ratio of parameters “deviates arbitrarily from any approximating distribution” because it is a function of the data to which it is applied. He then proposes that a valid interval can be found using the Fieller (1944) approach as the inversion of the appropriate likelihood ratio test and demonstrates that the interval is defined as the roots to a quadratic expression. In defining the general form of Fieller’s method for the construction of confidence intervals for ratios of linear functions of regression parameters Rao (1973) and Zerbe (1978) obtain equivalent expressions (equations 4b.2.16 and 2.5-2.7 respectively). Applications of the Fieller

technique for the inferences for varying income elasticities from Engle curve estimation can be found in Hirschberg et al (2008).

The inversion of a two sided test for a linear combination of the regression parameters is not only limited to the consideration of ratios of parameter values. In the examination of the existence of turning-points in nonlinear functions estimated by polynomial regressions, Williams (chapter 6, 1959) and Hellbronner (1979) demonstrated that the confidence interval of the location of the turning-points for a cubic regression could be constructed by inverting the test for the location of the value of the regressor where the first derivative function was equal to zero. Because the first derivative function of the cubic is a quadratic that implies two turning-points, the inversion test used to construct their confidence intervals, involves the solution of the four roots for a quartic to define the intervals of interest.⁴ In the particular application considered by Hellbronner, he limited his interest to only the interval that lay in the range of the x variable. Williams also employs the second derivative to locate the point of maximum rate of change and although he alludes to the application of inverse tests, he does not discuss the details of its implementation.

In this paper we demonstrate the uses for the inversion of the test for a single linear combination of parameters to draw inferences for a set of related nonlinear specifications. Vandaele (1981) has shown that these inverse t-tests are equivalent to the profile likelihood tests for these cases since these tests are equivalent to the corresponding likelihood ratio tests.

First we will outline the concept behind the inversion test. Then we discuss the application of the Fieller/likelihood profile test that involves the inversion of the t-test for the location of the turning-points (in this case an extremum) point defined by

⁴ Williams refers to the inverse test confidence intervals as fiducial limits.

the estimated quadratic regression. We then introduce the cubic and demonstrate that inferences concerning the location of turning-points can be found via the inversion tests. We then extend the cubic to consider the quartic specification that has recently been found in more applications in economics. This discussion is followed by the consideration of an example of the use of a quartic. The last section introduces a more general specification referred to as a fractional polynomial. This specification differs from a traditional polynomial in that all the higher order derivatives are not zero by assumption. We then demonstrate how this method can be applied to the same data examined using the quartic. The *Stata* routines for these analyses are given in the appendices along with a listing of the data used in the applications.

2. The inverse test confidence interval

The inverse test confidence interval is found by first defining a relationship for a critical value of a test statistic for a variable of interest and then solving for the values of the variable that define the limits of the interval. For example, a test of a linear combination of the mean vector of normally distributed random k dimensional vector $\mathbf{c}_{k \times 1} \sim N[\boldsymbol{\mu}_{k \times 1}, \boldsymbol{\Omega}_{k \times k}]$ defined as $\mathbf{R}_{1 \times k} \boldsymbol{\mu} = r$ we can propose the following 2-sided hypothesis test:

$$\begin{aligned} H_0 &: \mathbf{R}_{1 \times k} \boldsymbol{\mu} = r \\ H_1 &: \mathbf{R}_{1 \times k} \boldsymbol{\mu} \neq r \end{aligned}$$

And the usual confidence interval for the scalar r is given by:

$$CI = \mathbf{Rc} \pm z_{\alpha/2} \sqrt{\mathbf{R}\boldsymbol{\Omega}\mathbf{R}'} \quad (1)$$

Where $z_{\alpha/2}$ is the critical standard normal value for a $(1-\alpha)$ confidence interval. Under

H_0 we can construct the test statistic defined by:

$$T = \frac{\mathbf{Rc} - r}{\sqrt{\mathbf{R}\boldsymbol{\Omega}\mathbf{R}'}} \quad (2)$$

To establish if the H_0 can be rejected or not we need to determine if $T \geq |z_{\alpha/2}|$.

Alternatively, one could “invert the problem” by determining the critical values for r referred to as \tilde{r} , where the null hypothesis would be rejected at the $(1-\alpha)$ level of significance as defined by $z_{\alpha/2}$. Thus the implied value of the test statistic (\tilde{T}^2) is given by:

$$\tilde{T}^2 = \frac{(\mathbf{R}\mathbf{c} - \tilde{r})^2}{\mathbf{R}\mathbf{\Omega}\mathbf{R}'} = z_{\alpha/2}^2 \quad (3)$$

The inversion involves the solution to a quadratic in \tilde{r} defined as:

$$(\mathbf{R}\mathbf{c} - \tilde{r})^2 - z_{\alpha/2}^2 \mathbf{R}\mathbf{\Omega}\mathbf{R}' = 0 \quad (4)$$

The two solutions for \tilde{r} define the interval.⁵

$$\{\tilde{r}_U, \tilde{r}_L\} = \hat{r} \pm z_{\alpha/2} \sqrt{\mathbf{R}\mathbf{\Omega}\mathbf{R}'}, \text{ where } \hat{r} = \mathbf{R}\mathbf{c} \quad (5)$$

In this case the solution is trivial. To paraphrase Molière's *Bourgeois Gentilhomme* who discovers that he is already speaking prose, we can now observe that we have been using inverse confidence intervals all along. However, in a number of cases the solution is not equivalent to the traditional Wald test approach where we add and subtract a multiple of the standard error of the parameter estimate to the estimate.

Consider a related case where the multiplicative matrix \mathbf{R} is a function of an unknown parameter θ and r is given. In this case we wish to test the value(s) of θ for which $\mathbf{R}(\theta)\boldsymbol{\mu} = r$ using the two-sided test defined by:

$$\begin{aligned} H_0 &: \mathbf{R}(\theta)\boldsymbol{\mu} = r \\ H_1 &: \mathbf{R}(\theta)\boldsymbol{\mu} \neq r \end{aligned}$$

One way to proceed is to use estimates of $\boldsymbol{\mu}$ defined by \mathbf{c} and to solve

$\mathbf{R}(\hat{\theta})\mathbf{c} = r$ for estimate(s) of θ referred to as $\hat{\theta}$ when it is possible to solve

⁵ We will follow the convention where we use the circumflex (^) to denote the estimated location and the tilda (~) to denote the estimated interval limits.

$\mathbf{R}^{-1}(r, \mathbf{c}) = \hat{\theta}$. We could then construct a Wald test with an estimate of the variance of the $\hat{\theta}$ derived from the application of the delta method to find a linear approximation for $\mathbf{R}^{-1}(r, \mathbf{c})$.⁶

In this case the inverse test provides an alternative solution. Again specify the test statistic under the null (note that we expect up to twice the number of values for the value of θ to define the interval(s) for all the solutions for $\mathbf{R}^{-1}(r, \mathbf{c})$).

$$\tilde{T}^2 = \frac{(\mathbf{R}(\tilde{\theta})\mathbf{c} - r)^2}{\mathbf{R}(\tilde{\theta})\mathbf{\Omega}\mathbf{R}(\tilde{\theta})'} = z_{\alpha/2}^2 \quad (6)$$

Thus we can write this expression as a quadratic in $\mathbf{R}(\tilde{\theta})$ given by:

$$\mathbf{R}(\tilde{\theta})(\mathbf{c}\mathbf{c}' - z_{\alpha/2}^2\mathbf{\Omega})\mathbf{R}(\tilde{\theta})' - 2r\mathbf{c}'\mathbf{R}(\tilde{\theta})' + r^2 = 0, \quad (7)$$

which can be solved for the confidence interval(s). Since this is a quadratic form there are now potentially two solutions for every $\hat{\theta}$. Note that for any value of α , there is no guarantee that they will be real valued.

In this paper we investigate a particular set of cases where the null hypothesis of interest can be defined as $\mathbf{R}(\theta)\mathbf{c} = 0$ and where the restriction vector can be written as a polynomial function in θ and \mathbf{c} is the vector of estimated regression parameters. Typically we have the linear function vector defined as:

$$\mathbf{R}(\theta) = (R_0 \quad R_1\theta \quad R_2\theta^2 \quad \dots \quad R_j\theta^j \quad \dots \quad R_k\theta^k) \quad (8)$$

Thus in the cases dealt with here there are $2k$ potential confidence bounds defined by the roots to the polynomial defined by the solution to:

$$\mathbf{R}(\tilde{\theta})(\mathbf{c}\mathbf{c}' - z_{\alpha/2}^2\mathbf{\Omega})\mathbf{R}(\tilde{\theta})' = 0 \quad (9)$$

⁶ Note that the analytic solution may not be available as we find in our discussion below. As an alternative to the delta method one could use a simulation or empirical bootstrap constructed by draws from an estimate of $N[\boldsymbol{\mu}, \mathbf{\Omega}]$.

The graphical equivalent that we demonstrate is found from the plot of the confidence interval defined by:

$$CI(f(\theta)) = \mathbf{R}(\theta)\mathbf{c} \pm z_{\alpha/2} \sqrt{\mathbf{R}(\theta)\mathbf{\Omega}\mathbf{R}(\theta)'} \quad (10)$$

From which we can locate the values of θ where the upper and lower bounds cross the zero reference line to define the $\tilde{\theta}$.⁷

3. The Quadratic Specification

The quadratic functional form is widely used in regression to identify a U or inverted U-shape relationship between a regressor and the dependent variable. Many authors claim to have found such a relationship by focussing of the significance and signs of the estimated coefficients corresponding to the quadratic (Clark et al. 1996) and possibly also examining whether the position of the extremum falls within the range of the data (see Richmond and Kaufman 2006). Dufour's (1997) admonition notwithstanding, a number of authors continue to use the delta method to calculate a standard error for the extremum and then construct symmetric confidence intervals (e.g. Kearsley and Riddell 2010).⁸ Although there has been a considerable number of Monte Carlo studies that record the similarity of the confidence intervals obtained using the Fieller and the delta it has been shown that the coverage of the delta confidence interval is related to the characteristics of the particular sample used in estimation (Hirschberg and Lye 2010a). More recently, authors have used confidence intervals for the extremum based on the Fieller method as well as the Sasabuchi

⁷ The case where they do not cross the zero reference line will indicate the situation where no real root for the equivalent polynomial defined by (9).

⁸ We refer to the delta as the Wald test confidence interval derived by combination of the delta approximation for the variance of the ratio with the definition of the confidence interval by adding and subtracting a multiple of the estimated standard error.

(1980) intersection-union test (see Schnabel, C. and J. Wagner 2008; Huang et al 2012).⁹

The quadratic specification is typically written as:

$$y_t = \beta_1 x_t + \beta_2 x_t^2 + \left(\alpha_0 + \sum_{j=1}^M \alpha_j z_{jt} \right) + \varepsilon_t \quad (11)$$

Again we assume that there is a single regressor x entered as a quadratic function, M linearly related regressors denoted by z_j , and N observations. By setting the marginal effect or the first derivative of $E(y|x, z)$ with respect to x , $\frac{\partial E(y|x, z)}{\partial x} = \beta_1 + 2\beta_2 x_t$, equal to zero, we can derive the value of x that defines the extremum of the relationship between y and x given values for the other regressors, as $\theta = \frac{-\beta_1}{2\beta_2}$. θ is either the maximum (when $\beta_2 < 0$) or the minimum (when $\beta_2 > 0$). These conditions are established via the signs of the second order derivative of the function where $x = \theta$. In this case $\frac{\partial^2 E(y|\theta, z)}{\partial x^2} = 2\beta_2$, thus the sign of β_2 will establish the nature of the extremum.¹⁰ The usual estimate of the extremum is given by $\hat{\theta} = \frac{-\hat{\beta}_1}{2\hat{\beta}_2}$, where $\hat{\beta}_1$ and $\hat{\beta}_2$ are the OLS estimates of β_1 and β_2 in (11) respectively. The point defined by $\hat{\theta}$ also defines the value of x where the marginal effect of x on y changes sign.

3.1 The Delta method Applied to the Quadratic

The delta method for drawing inferences about a nonlinear function such as θ , is to estimate its variance by use of a first order Taylor series expansion to linearize

⁹ Lind and Mehlum (2010) observe that the Sasabuchi (1980) test at the α level of significance is equivalent to determining if the $(1 - 2\alpha)$ confidence interval for the extremum is within the range of the x variable.

¹⁰ This also justifies the use of the test of the null hypothesis that $\beta_2 = 0$ to establish whether an inflection point is present or not and to justify the consideration of this functional form.

the non-linear function. In this case the estimated variance of $\hat{\theta}$ is given by (see, for example, Rao 1973, pp 385-389):

$$\frac{1}{4} \begin{bmatrix} \left(\frac{-1}{\hat{\beta}_2}\right) & \left(\frac{\hat{\beta}_1}{\hat{\beta}_2^2}\right) \end{bmatrix} \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{12} \\ \hat{\sigma}_{12} & \hat{\sigma}_2^2 \end{bmatrix} \begin{bmatrix} \left(\frac{-1}{\hat{\beta}_2}\right) \\ \left(\frac{\hat{\beta}_1}{\hat{\beta}_2^2}\right) \end{bmatrix} = \frac{\hat{\sigma}_1^2 \hat{\beta}_2^2 - 2\hat{\beta}_1 \hat{\beta}_2 \hat{\sigma}_{12} + \hat{\beta}_1^2 \hat{\sigma}_2^2}{4\hat{\beta}_2^4} \quad (12)$$

where $\hat{\sigma}_i^2$ is the estimated variance of $\hat{\beta}_i$ and $\hat{\sigma}_{ij}$ is the estimated covariance between $\hat{\beta}_i$ and $\hat{\beta}_j$. A $100(1 - \alpha)\%$ confidence interval for θ is defined by:

$$\{\tilde{\theta}_U, \tilde{\theta}_L\}_{delta} = \hat{\theta} \pm t_{\alpha/2} \sqrt{\frac{\hat{\sigma}_1^2 \hat{\beta}_2^2 - 2\hat{\beta}_1 \hat{\beta}_2 \hat{\sigma}_{12} + \hat{\beta}_1^2 \hat{\sigma}_2^2}{4\hat{\beta}_2^4}} \quad (13)$$

where $t_{\alpha/2}$ is the value from the t distribution with an $(\alpha/2)\%$ level of significance and $N-k$ degrees of freedom. This is commonly referred to as a delta method confidence interval and can readily be computed using post-regression software routines available in most statistics packages.

3.2 The Inversion method Applied to the Quadratic

An alternative approach to construct a confidence interval for θ is found by inverting the t -test associated with the null hypothesis:

$$H_0 : \beta_1 + 2\beta_2\theta = 0$$

$$H_1 : \beta_1 + 2\beta_2\theta \neq 0$$

A $100(1 - \alpha)\%$ confidence interval for θ is obtained by solving¹¹

$$\tilde{T}^2 = \frac{(\hat{\beta}_1 + 2\hat{\beta}_2\tilde{\theta})^2}{(\hat{\sigma}_1^2 + 4\hat{\sigma}_{12}\tilde{\theta} + 4\hat{\sigma}_2^2\tilde{\theta}^2)} = t_{\alpha/2}^2 \quad (14)$$

¹¹ We will substitute t^2 for the F with one degree of freedom in the numerator.

The solution is found by solving $a\tilde{\theta}^2 + b\tilde{\theta} + c = 0$ where $a = 4\hat{\sigma}_2^2(\hat{t}_2^2 - t_{\alpha/2}^2)$,

$b = 4(\hat{\beta}_1\hat{\beta}_2 - t_{\alpha/2}^2\hat{\sigma}_{12})$ and $c = \hat{\sigma}_1^2(\hat{t}_1^2 - t_{\alpha/2}^2)$, where we define the t -ratios for the

parameter estimates as: $\hat{t}_1 = \hat{\beta}_1/\hat{\sigma}_1$ and $\hat{t}_2 = \hat{\beta}_2/\hat{\sigma}_2$. Then, the roots can be defined as:

$$\{\tilde{\theta}_U, \tilde{\theta}_L\}_{Fieller} = \frac{t_{\alpha/2}^2\hat{\sigma}_{12} - \hat{\beta}_1\hat{\beta}_2 \pm \sqrt{(\hat{\beta}_1\hat{\beta}_2 - t_{\alpha/2}^2\hat{\sigma}_{12})^2 - \hat{\sigma}_2^2\hat{\sigma}_1^2(\hat{t}_2^2 - t_{\alpha/2}^2)(\hat{t}_1^2 - t_{\alpha/2}^2)}}{2\hat{\sigma}_2^2(\hat{t}_2^2 - t_{\alpha/2}^2)} \quad (15)$$

If $\hat{t}_2^2 > t_{\alpha/2}^2$, then (15) has two real roots and a finite confidence interval is constructed.

This condition corresponds to being able to reject $H_0 : \beta_2 \neq 0$ when α is the level of significance (Buonaccorsi 1979). In addition to the finite interval case, the resulting confidence interval may be the complement of a finite interval when $(b^2 - 4ac > 0, a < 0)$ or of the whole real line when $b^2 - 4ac < 0, a < 0$. These conditions are discussed in Scheffé (1970) and Zerbe et al (1982). This approach results in a confidence interval that is equivalent to the Fieller method (Fieller 1944) which provides a general procedure for constructing confidence limits for the ratio of two parameters, estimated by normally distributed random variables.

Graphically, (as shown in Hirschberg and Lye 2010b) the $100(1 - \alpha)\%$ confidence interval for θ following the inversion method, is obtained by plotting an estimate of the first derivative as a function of x ($\hat{\beta}_1 + 2\hat{\beta}_2x$) with its corresponding $100(1 - \alpha)\%$ confidence interval defined as:

$$CI\left(\frac{\partial E(y|x,z)}{\partial x}\right) = (\hat{\beta}_1 + 2\hat{\beta}_2x) \pm t_{\alpha/2}\sqrt{(\hat{\sigma}_1^2 + 4x\hat{\sigma}_{12} + 4x^2\hat{\sigma}_2^2)} \quad (16)$$

An estimate of the extremum value, $x = \hat{\theta}$, is found by solving the equation

$\hat{\beta}_1 + 2\hat{\beta}_2\hat{\theta} = 0$. Similarly, the bounds that define a $100(1 - \alpha)\%$ confidence interval

on θ are found by solving for x in the relationship:

$$\left(\hat{\beta}_1 + 2\hat{\beta}_2 x\right) \pm t_{\frac{\alpha}{2}} \sqrt{\left(\hat{\sigma}_1^2 + 4x\hat{\sigma}_{12} + 4x^2\hat{\sigma}_2^2\right)} = 0 \quad (17)$$

which is equivalent to solving the roots of the equation:

$$\left(\hat{\beta}_1 + 2\hat{\beta}_2 x\right)^2 - t_{\frac{\alpha}{2}}^2 \left(\hat{\sigma}_1^2 + 4x\hat{\sigma}_{12} + 4x^2\hat{\sigma}_2^2\right) = 0. \quad (18)$$

By rearranging the terms in (18), this can be written as the quadratic equation,

$ax^2 + bx + c = 0$, where a , b and c are the terms defined above.

4. The Cubic Specification

Aside from the quadratic the next most commonly employed polynomial specification is the cubic. Many applications that use the cubic are accompanied by a plot of the estimated function along with descriptive comments on its shape (e.g. Bagliani et al. 2008). In a limited number of cases, the definition of the turning-points implied by the roots of the first derivative function and the determination of the signs of the second order derivative are used to determine the nature of the turning-points (Hellbronner 1979, Naghshpour 2005; Vitaliano and Held 1991). In some cases Wald tests are employed to find standard errors for the turning-points as in French and Zarkin (1995). Alternative techniques for the construction of confidence intervals have been applied in the application of the cubic; Plassmann and Khanna (2007) use a Bayesian approach for the establishment of turning-points.

The cubic polynomial specification is defined as:

$$y_t = \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + \left(\alpha_0 + \sum_{j=1}^M \alpha_j z_{jt} \right) + \varepsilon_t \quad (19)$$

In this case the first derivative function is given by a quadratic equation defined as:

$$\frac{\partial E(y/x, z)}{\partial x} = \beta_1 + 2\beta_2 x + 3\beta_3 x^2 \quad (20)$$

The solution of this equation when set to zero is given by the solution to a quadratic.

Thus there are two turning-points from the two roots (given $\beta_3 \neq 0$) defined as:

$$\theta_i = -\frac{1}{3\beta_3}(\beta_2 + \phi_i \gamma), \text{ where } \gamma = \sqrt{\beta_2^2 - 3\beta_1\beta_3}, \phi_1 = 1, \text{ and } \phi_2 = -1 \quad (21)$$

The second derivative of the cubic evaluated at the i th root of the quadratic in (21) defined by θ_i is given as $\frac{\partial^2 y}{\partial x^2} = 2\beta_2 + 6\beta_3\theta_i$. Both roots of (21) are real if $\gamma > 0$.

If this is the case the second derivative test indicates that θ_i is a relative (or local) minimum if $(2\beta_2 + 6\beta_3\theta_i > 0)$ and θ_i is a relative maximum when $(2\beta_2 + 6\beta_3\theta_i < 0)$.

4.1 The Delta method Applied to the Cubic

The approximate variance of $\hat{\theta}_i$ based on the delta method is

$$\widehat{V}(\theta_i) = d_1\hat{\sigma}_1^2 + 2d_1d_2\hat{\sigma}_{12} + 2d_1d_3\hat{\sigma}_{13} + d_2\hat{\sigma}_2^2 + 2d_2d_3\hat{\sigma}_{23} + d_3\hat{\sigma}_3^2 \quad (22)$$

where $[d_{i1} \quad d_{i2} \quad d_{i3}] = \left[\frac{\partial \theta_i}{\partial \beta_1} \quad \frac{\partial \theta_i}{\partial \beta_2} \quad \frac{\partial \theta_i}{\partial \beta_3} \right]$ evaluated at the estimates, thus $d_{i1} = \frac{\phi_i}{2\hat{\gamma}}$,

$$d_{i2} = \frac{-(\phi_i\hat{\beta}_2 + \hat{\gamma})}{3\beta_3\hat{\gamma}}, \text{ and } d_{i3} = \frac{2\hat{\beta}_2\hat{\gamma} - \phi_i(3\hat{\beta}_1\hat{\beta}_3 - 2\hat{\beta}_2^2)}{6\hat{\beta}_3^2\hat{\gamma}} \text{ for } i = 1, 2. \text{ A } 100(1 - \alpha)\% \text{ delta}$$

method based confidence interval for θ_i is given by: $\{\tilde{\theta}_{iU}, \tilde{\theta}_{iL}\}_{delta} = \hat{\theta}_i \pm t_{\alpha/2} \sqrt{\widehat{V}(\theta_i)}$.

A number of programs in computer packages can be used to calculate the delta variance estimates such as *nlcom* in *Stata*.

4.2 The Inversion Method Applied to the Cubic

A confidence interval for θ_i can be constructed by inverting the t -test associated with the null hypothesis

$$H_0 : \beta_1 + 2\beta_2\theta + 3\beta_3\theta^2 = 0$$

$$H_1 : \beta_1 + 2\beta_2\theta + 3\beta_3\theta^2 \neq 0$$

A $100(1 - \alpha)\%$ confidence interval for θ_i is obtained by solving for the values of θ at which the following expression holds.

$$T^2 = \frac{(\hat{\beta}_1 + 2\hat{\beta}_2\tilde{\theta} + 3\hat{\beta}_3\tilde{\theta}^2)^2}{(\hat{\sigma}_1^2 + 4\tilde{\theta}\hat{\sigma}_{12} + 4\tilde{\theta}^2\hat{\sigma}_2^2 + 6\tilde{\theta}^3\hat{\sigma}_{13} + 12\tilde{\theta}^4\hat{\sigma}_{23} + 9\tilde{\theta}^4\hat{\sigma}_3^2)} = t_{\alpha/2}^2 \quad (23)$$

By rearrangement we define a quartic equation whose roots are the limits of the implied confidence intervals for the turning-points of the original cubic function.

This quartic expression that is equivalent to (23) is given below.

$$9(\hat{\beta}_3^2 - t_{\alpha/2}^2\hat{\sigma}_3^2)\tilde{\theta}^4 + 12(\hat{\beta}_2\hat{\beta}_3 - t_{\alpha/2}^2\hat{\sigma}_{23})\tilde{\theta}^3 + 2\left(\left(3\hat{\beta}_1\hat{\beta}_3 + 2\hat{\beta}_2^2\right) - t_{\alpha/2}^2(3\hat{\sigma}_{13} + 2\hat{\sigma}_2^2)\right)\tilde{\theta}^2 + 4(\hat{\beta}_1\hat{\beta}_2 - t_{\alpha/2}^2\hat{\sigma}_{12})\tilde{\theta} + (\hat{\beta}_1^2 - t_{\alpha/2}^2\hat{\sigma}_1^2) = 0 \quad (24)$$

The corresponding four roots to this quartic define the confidence intervals for the turning points and can be found numerically using widely available software such as the *polyroots* function in *Stata* and other widely used software that implement the algorithm proposed by Jenkins and Traub (1970) for the numerical derivation of the roots for high order polynomials. The roots can then be associated with the corresponding confidence bounds for θ_i .

These can also be obtained graphically by plotting an estimate of the first derivative as a function of x ($\hat{\beta}_1 + 2\hat{\beta}_2x_t + 3\hat{\beta}_3x_t^2$) and a $100(1 - \alpha)$ confidence interval.

$$CI\left(\frac{\partial E(y/x_t, z)}{\partial x_t}\right) = \left(\hat{\beta}_1 + 2\hat{\beta}_2x_t + 3\hat{\beta}_3x_t^2\right) \pm t_{\alpha/2}\sqrt{\left(\begin{array}{l} \hat{\sigma}_1^2 + 4x_t\hat{\sigma}_{12} + 4x_t^2\hat{\sigma}_2^2 \\ +6x_t^2\hat{\sigma}_{13} + 12x_t^3\hat{\sigma}_{23} + 9x_t^4\hat{\sigma}_3^2 \end{array}\right)} \quad (25)$$

The estimated values of θ_i correspond to the values at which $\hat{\beta}_1 + 2\hat{\beta}_2x_t + 3\hat{\beta}_3x_t^2 = 0$ and the corresponding bounds are given by the values where the confidence interval intersects the zero reference line. Note that the CI is defined for all possible values of x_t however if the interval never cuts the zero reference line we can conclude that for this value of α that the interval is infinite. The possibility of an infinite bound is one of the characteristics of the inversion test and corresponds to the case where we obtain imaginary roots for (24).

In order to clearly identify the nature of the root as either a maximum or a minimum we invert the test that the second derivative is equal to zero as we have done for the confidence intervals of the turning-points. The function of the parameters defined by the second derivative function is given as $\left(\frac{\partial^2 y}{\partial x^2} = 2\beta_2 + 6\beta_3\theta_i\right)$ thus we can write the test as:¹²

$$T^2 = \frac{(2\hat{\beta}_2 + 6\hat{\beta}_3\bar{\theta})^2}{(36\bar{\theta}^2\hat{\sigma}_3^2 + 24\hat{\sigma}_{23}\bar{\theta} + 4\hat{\sigma}_2^2)} = t_{\alpha/2}^2 \quad (26)$$

We can solve for the roots of the quadratic defined below in (27) to establish the interval over which the turning-points can be identified as a maximum or a minimum.

$$36(\hat{\beta}_3^2 - t_{\alpha/2}^2\hat{\sigma}_3^2)\bar{\theta}^2 + 24(\hat{\beta}_2\hat{\beta}_3 - t_{\alpha/2}^2\hat{\sigma}_{23})\bar{\theta} + 4(\hat{\beta}_2^2 - t_{\alpha/2}^2\hat{\sigma}_2^2) = 0 \quad (27)$$

In analogous fashion to the intervals found for the first derivative function, we also find that this interval is equivalent to plotting the second derivative function with the corresponding $100(1 - \alpha)$ confidence interval evaluated at different values of x is given by:

$$CI\left(\frac{\partial^2 E(y/x, z)}{\partial x^2}\right) = (2\hat{\beta}_2 + 6\hat{\beta}_3x_t) \pm t_{\alpha/2}\sqrt{36x_t^2\hat{\sigma}_3^2 + 24\hat{\sigma}_{23}x_t + 4\hat{\sigma}_2^2} \quad (28)$$

Thus, the plot of the confidence interval for the second derivative function can be used to define the characteristics of the roots of the first derivative function.

5. The Quartic Specification

Following the trend in the applied economics literature to estimate higher order polynomials a number of studies report the estimation of a quartic or a fourth-order polynomial (Coate and Fry 2012; Kumar and Viswanathan 2007; Wong 2011). As in the application of lower order polynomials, the most common approach to describing the estimated function is to plot the estimated functional form. In this section we demonstrate that the inversion test confidence interval for the turning-

¹² Here we use $\bar{\theta}$ to denote confidence bounds for the 2nd derivative function.

points can be found from the confidence interval of the first derivative function. We follow the approach outlined by Williams (Section 6.14, 1959).

For the quartic or fourth order polynomial the regression is specified as:

$$y_t = \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + \beta_4 x_t^4 + \left(\alpha_0 + \sum_{j=1}^M \alpha_j z_{jt} \right) + \varepsilon_t \quad (29)$$

where the first derivative function is given by a cubic equation defined as:

$$\frac{\partial y}{\partial x} = \beta_1 + 2\beta_2 x + 3\beta_3 x^2 + 4\beta_4 x^3 \quad (30)$$

The turning-points can be found from the solution of this equation when set to zero:

$$\theta^3 + a_1 \theta^2 + a_2 \theta + a_3 = 0 \quad (31)$$

Where $a_1 = \frac{3\beta_3}{4\beta_4}$, $a_2 = \frac{\beta_2}{2\beta_4}$, and $a_3 = \frac{\beta_1}{4\beta_4}$ and define the intermediate terms:

$Q = \frac{a_1^2 - 3a_2}{9}$ and $R = \frac{2a_1^3 - 9a_1 a_2 + 27a_3}{54}$. If $Q^3 - R^2 \geq 0$ then there are 3 real roots.

Defining $\lambda = \arccos\left(\frac{R}{\sqrt{Q^3}}\right)$ then the roots are given by the equations:

$$\theta_1 = -2\sqrt{Q} \cos\left(\frac{\lambda}{3}\right) - \frac{a_1}{3} \quad (32)$$

$$\theta_2 = -2\sqrt{Q} \cos\left(\frac{\lambda + 2\pi}{3}\right) - \frac{a_1}{3} \quad (33)$$

$$\theta_3 = -2\sqrt{Q} \cos\left(\frac{\lambda + 4\pi}{3}\right) - \frac{a_1}{3}. \quad (34)$$

If $R^2 - Q^3 > 0$ then there is only 1 real root given by the equation:

$$\theta_1 = -\text{sgn}(R) \left[\left(\sqrt{R^2 - Q^3} + |R| \right)^{\frac{1}{3}} + Q \left(\sqrt{R^2 - Q^3} + |R| \right)^{\frac{1}{3}} \right] - \frac{a_1}{3}. \quad (35)$$

For the quartic the second order derivative given the value of the roots as defined by the quadratic equation:

$$\left[\frac{\partial^2 y}{\partial x^2} \mid \theta_i \right] = 2\beta_2 + 6\beta_3 \theta_i + 12\beta_4 \theta_i^2 \quad (36)$$

where θ_i represents the root being examined as estimated by $\hat{\theta}_i$ based on the estimates of the parameters of the model.

To determine whether a particular root is a relative maximum or minimum point we would test if the value is < 0 , or > 0 respectively. Note that although the sign of the second derivative function is sufficient to determine if a turning-point is a relative maximum or minimum, the second derivative being equal to zero is only a necessary condition for the location of an inflection point.

5.1 The Delta method Applied to the Quartic

The expressions defined by (32) – (35) for the roots of the quartic though they are complex can be used to derive delta standard errors when a program is available to obtain the approximate standard errors for non-linear functions of regression parameters. An example is shown in the *Stata* routine provided in Appendix A.1 for the application in Section 5.3. Note that due to the complexity of these expressions it may be necessary to derive the estimated variance covariance matrix for the intermediate expression such as the a_i defined above.

The delta confidence interval of the second derivative function can also be determined by estimating the variance of the roots to (36) evaluated at the turning-points defined by $\hat{\theta}_i$ derived in (32)-(35) above. An example of this is described in Section 5.3.

5.2 The Inversion Method Applied to the Quartic

Alternatively we can construct the intervals from the inversion of the test using the tests for the first derivatives and the second derivatives. Again the first derivative function for the quartic is defined as:

$$\frac{dy_t}{dx_t} = \beta_1 + 2\beta_2 x_t + 3\beta_3 x_t^2 + 4\beta_4 x_t^3 \quad (37)$$

A confidence interval for the intervals θ^i can be constructed by inverting the F-test associated with the null hypothesis that the first derivative is equal to zero versus the two-sided alternative:

$$\begin{aligned} H_0 : \beta_1 + 2\beta_2\theta + 3\beta_3\theta^2 + 4\beta_4\theta^3 &= 0 \\ H_1 : \beta_1 + 2\beta_2\theta + 3\beta_3\theta^2 + 4\beta_4\theta^3 &\neq 0. \end{aligned}$$

A $100(1 - \alpha)\%$ confidence interval for θ_i is obtained by solving for the solutions to

$$T^2 = \frac{(\hat{\beta}_1 + 2\hat{\beta}_2\tilde{\theta} + 3\hat{\beta}_3\tilde{\theta}^2 + 4\hat{\beta}_4\tilde{\theta}^3)^2}{(\hat{\sigma}_1^2 + 4\tilde{\theta}\hat{\sigma}_{12} + 2(2\hat{\sigma}_2^2 + 3\hat{\sigma}_{13})\tilde{\theta}^2 + 4(3\hat{\sigma}_{23} + 2\hat{\sigma}_{14})\tilde{\theta}^3 + (9\hat{\sigma}_3^2 + 16\hat{\sigma}_{24})\tilde{\theta}^4 + 24\hat{\sigma}_{34}\tilde{\theta}^5 + 16\hat{\sigma}_4^2\tilde{\theta}^6)} = t_{\alpha/2}^2 \quad (38)$$

The solution can be found by finding the roots of the 6th order polynomial defined by:

$$\begin{aligned} &16(\hat{\beta}_4^2 - t^2\hat{\sigma}_4^2)\tilde{\theta}^6 + 24(\hat{\beta}_3\hat{\beta}_4 - t^2\hat{\sigma}_{34})\tilde{\theta}^5 + (9\hat{\beta}_3^2 + 16\hat{\beta}_2\hat{\beta}_4 - t^2(9\hat{\sigma}_3^2 + 16\hat{\sigma}_{24}))\tilde{\theta}^4 \\ &+ (12\hat{\beta}_2\hat{\beta}_3 + 8\hat{\beta}_1\hat{\beta}_4 - t^2(12\hat{\sigma}_{23} + 8\hat{\sigma}_{14}))\tilde{\theta}^3 + 2((3\hat{\beta}_1\hat{\beta}_3 + 2\hat{\beta}_2^2) - t^2(3\hat{\sigma}_{13} + 2\hat{\sigma}_2^2))\tilde{\theta}^2 \\ &+ 4(\hat{\beta}_1\hat{\beta}_2 - t^2\hat{\sigma}_{12})\tilde{\theta} + (\hat{\beta}_1^2 - t^2\hat{\sigma}_1^2) = 0 \end{aligned} \quad (39)$$

Note that although a general analytic solution for these roots is not available we can obtain numerical solutions as discussed above.

An estimate of the first derivative as a function of x can be plotted with a $1 - \alpha$ confidence interval as:

$$CI\left(\frac{\partial Y_t}{\partial x_t}\right) = \left(\hat{\beta}_1 + 2\hat{\beta}_2x_t + 3\hat{\beta}_3x_t^2 + 4\hat{\beta}_4x_t^3\right) \pm t_{\alpha/2}^2 \sqrt{\begin{pmatrix} \hat{\sigma}_1^2 + 4x_t\hat{\sigma}_{12} + 2(2\hat{\sigma}_2^2 + 3\hat{\sigma}_{13})x_t^2 \\ +4(3\hat{\sigma}_{23} + 2\hat{\sigma}_{14})x_t^3 + (9\hat{\sigma}_3^2 + 16\hat{\sigma}_{24})x_t^4 \\ +24\hat{\sigma}_{34}x_t^5 + 16\hat{\sigma}_4^2x_t^6 \end{pmatrix}} \quad (40)$$

and the equivalent bounds for the first derivative function can be located where the confidence interval includes zero.

In this case the second order derivative function can also be used as a test of which solutions are relative extremum points. Thus we are interested in the region in which the null hypothesis can be rejected.

$$H_0 : 2\beta_2 + 6\beta_3\theta + 12\beta_4\theta^2 = 0$$

$$H_1 : 2\beta_2 + 6\beta_3\theta + 12\beta_4\theta^2 \neq 0$$

Using the test for the second order derivative function at the estimated root we find that the solution becomes the roots of a quartic.

$$T^2 = \frac{(2\hat{\beta}_2 + 6\hat{\beta}_3\tilde{\theta} + 12\hat{\beta}_4\tilde{\theta}^2)^2}{(4\hat{\sigma}_{22} + 24\hat{\sigma}_{23}\tilde{\theta} + (48\hat{\sigma}_{24} + 36\hat{\sigma}_{33})\tilde{\theta}^2 + 144\hat{\sigma}_{34}\tilde{\theta}^3 + 144\hat{\sigma}_{44}\tilde{\theta}^4)} = t_{\alpha/2}^2 \quad (41)$$

For the limits of the second derivative function where we are unable to reject the null hypothesis we can solve for the solutions defined by $\tilde{\theta}$ of the quartic defined in (42):

$$\begin{aligned} & 144[\hat{\beta}_4^2 - t_{\alpha/2}^2\hat{\sigma}_{44}] \tilde{\theta}^4 + 144[\hat{\beta}_3\hat{\beta}_4 - t_{\alpha/2}^2\hat{\sigma}_{34}] \tilde{\theta}^3 \\ & + [36(\hat{\beta}_3^2 - t_{\alpha/2}^2\hat{\sigma}_{33}) + 48(\hat{\beta}_2\hat{\beta}_4 - t_{\alpha/2}^2\hat{\sigma}_{24})] \tilde{\theta}^2 \\ & + 24[\hat{\beta}_2\hat{\beta}_3 - t_{\alpha/2}^2\hat{\sigma}_{23}] \tilde{\theta} + 4[\hat{\beta}_2^2 - t_{\alpha/2}^2\hat{\sigma}_{22}] = 0 \end{aligned} \quad (42)$$

The four roots of (42) define the limits of the regions over which we cannot reject the hypothesis that the sign of the second derivative is neither positive nor negative.

Again the alternative to the solution to this polynomial is to inspect the plot of the (100- α)% confidence interval of second derivative function defined by:

$$CI\left(\frac{\partial^2 Y_t}{\partial x_t^2}\right) = 2\hat{\beta}_2 + 6\hat{\beta}_3x_t + 12\hat{\beta}_4x_t^2 \pm t_{\alpha/2}^2 \sqrt{\begin{pmatrix} 4\hat{\sigma}_{22} + 24\hat{\sigma}_{23}x_t \\ + (48\hat{\sigma}_{24} + 36\hat{\sigma}_{33})x_t^2 \\ + 144(\hat{\sigma}_{34}x_t^3 + \hat{\sigma}_{44}x_t^4) \end{pmatrix}} \quad (43)$$

By establishing if the CI for the second derivative function at the location of estimated roots of the first derivative function (37) defined by the values of $\hat{\theta}_i$ is greater or less than zero we can infer the sign of the second order derivative at the turning-point.

5.3 Example of a Quartic¹³

Utt et al. (2001) estimate an Environmental Kuznets Curve (EKC) for the USA using annual observations on CO₂ per capita emissions from 1929 to 1996. Using a quartic polynomial they estimate an EKC with two peaks at about US\$8,000 and US\$28,000 and a trough at about US\$20,000 (as measured in constant 1990 dollars).¹⁴ In this case the regression is specified with only one independent variable which is included as a quartic of the form:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 x_t^3 + \beta_4 x_t^4 + \varepsilon_t \quad (44)$$

Where y is the CO₂ emissions per capita as reported in Utt et al (2001) and x is the real GDP per capita (in constant US\$2000) from Winer et al. (2008). Although these data are time series in nature and they exhibit the characteristics of potential unit-roots we appeal to the analysis by Liu et al (2009) in which they demonstrate that the properties of the asymptotic inferential methods work well when such a high order polynomial is estimated in the levels. Consequently, for this example we will apply OLS to this sample and employ the asymptotic standard errors.

<i>Coefficient</i>	<i>Estimate</i>	<i>Std Error.</i>	<i>t-Stat</i>
$\hat{\beta}_0$	2.98	0.10	28.42
$\hat{\beta}_1$	8.49	0.69	12.34
$\hat{\beta}_2$	-9.12	1.12	-8.16
$\hat{\beta}_3$	3.76	0.63	5.95
$\hat{\beta}_4$	-0.53	0.11	-4.62
<i>Adjusted R²</i>	0.86	<i>N</i>	68

Table 1. The parameter estimates for the quartic function applied to the Utt et al data series as defined in Appendix B.

Table 1 presents the resulting parameter estimates and their standard errors. In Figure 1 we have plotted the scatter plot of the observations along with the predicted

¹³ The *Stata* routine to generate the results in this section is given in Appendix A.1. Note that the figures have been edited to add comments.

¹⁴ The data used in this analysis are listed in Appendix B.

function. Note that we have included additional observations with x values from \$30,000 to \$35,000 that are outside the range of the observations.¹⁵

Using the first derivative approach we estimate three turning-points. Two maxima at US\$8076 and US\$27,045 and a minimum at \$US18,476. Listed in Table 2 are the estimated standard errors for these non-linear functions of the data as estimated by the use of the delta method applied to the non-linear equations defined in (32) to (34). Also listed in Table 2 are the Confidence intervals obtained using both the standard errors estimated via the delta method and using the inversion test approach based on the solution to the sixth order polynomial as defined in (39). The latter estimates were obtained numerically through the application of the *Stata* routine *polyroots*.

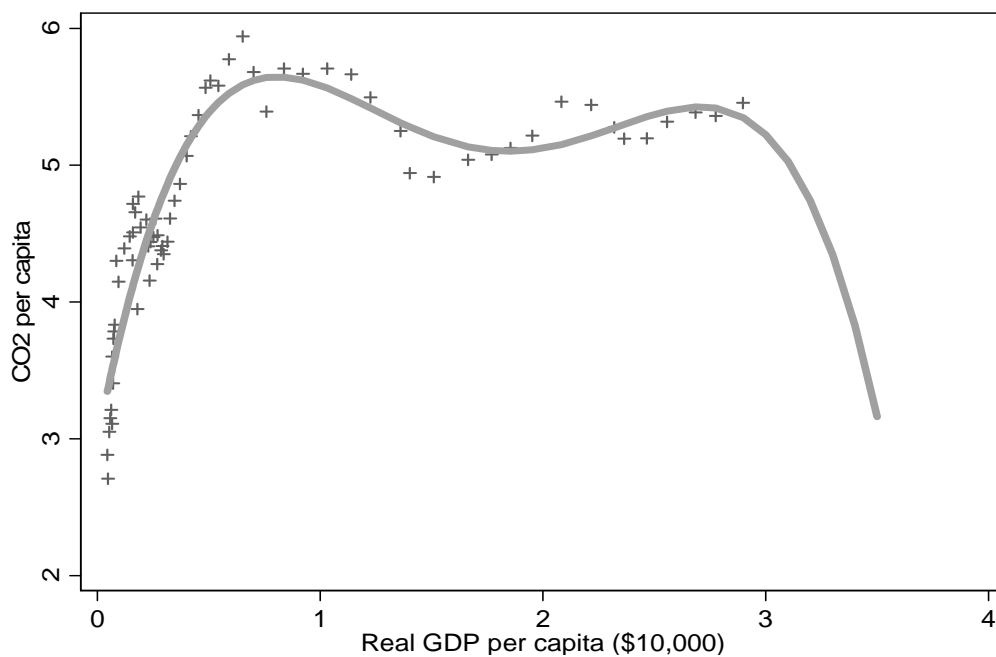


Figure1. The scatter plot of the observations and the quartic function fit by the regression.

¹⁵ Here we exploit a common feature in many regression software packages to predict out-of-sample when the dependent variable is missing but the regressors are available.

<i>Turning-point</i>	<i>Estimate</i>	<i>Std. Err.</i>	<i>95% Confidence Interval.</i>			
			<i>Delta</i>		<i>Inverse test</i>	
$\hat{\theta}_1$	0.81	0.04	0.72	0.89	0.73	0.91
$\hat{\theta}_2$	1.85	0.11	1.62	2.07	1.66	2.10
$\hat{\theta}_3$	2.70	0.18	2.35	3.06	2.46	3.41

Table 2 The estimated turning-points and the standard error estimated via the application of the delta method with the confidence intervals based on the two methods. All values are in \$10,000.

From Table 2 we note that the most of the bounds are similar except for the upper bound on the last turning-point that appears to be far out of the sample range for the x variable. The maximum value for real GDP per capita in the sample is \$29,000 which is well within the interval over which we are unable to reject the null hypothesis that the first derivative function is zero thus implying a flat relationship up to \$30,600 implied by the delta CI or \$34,100 found via the inversion test.

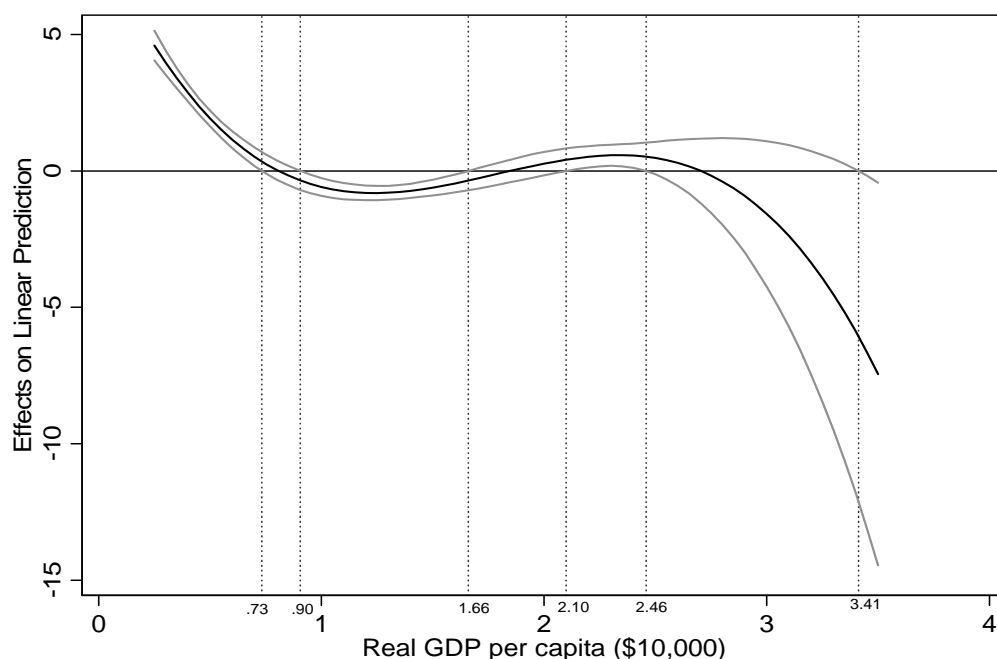


Figure 2 The first derivative function with a 95% confidence interval and the confidence intervals found using the inverse test.

Alternatively, the inverse test confidence interval from the CI of the first derivative function defined by (40) can be found from Figure 2.¹⁶ We have included the location of the roots obtained from the numeric solution to demonstrate that they coincide with the values found by locating the points where the 95% CI of the 1st derivative function cut the zero reference line.

In the case that the polynomial defined by (39) does not have 6 real roots we would find that the equivalent figure to Figure 2 would have CIs that would not cross the zero line for these imaginary roots. Thus it may be possible for the equivalent plot to display open ended intervals for some roots in which case the upper or lower limit may be ∞ or $-\infty$ respectively.

In order to determine the nature of these turning-points we now can investigate the 2nd derivative function of these estimated values. The second order derivative function for the quartic is given by (36) and we can now evaluate this function at the turning-points defined by the roots of the cubic implied by the first derivative function defined in (31). The delta method would involve the substitution of the estimates for the roots ($\hat{\theta}_i$) into the second derivative function to determine the sign. The *Stata* routine in the appendix has been set to evaluate the second derivatives at these points. The resulting intervals for each turning-point are given in Table 3. From this table we can infer that the first turning-point has a negative second derivative indicating a local maximum, the second turning-point has a positive second derivative indicating a local minimum and the third has a positive second order derivative which implies another local maximum. In addition, we also find that the estimated standard errors for the first two turning-points indicate that each of these values of $\frac{\partial^2 y}{\partial x^2} | \hat{\theta}_i$ are significantly

¹⁶ The first derivative plot can be provided by the *marginplot* routine in *Stata* as shown in the program provided in the appendix. Alternatively one can use the *nlcom* routine with a little more programming this method is also shown in the *Stata* routine in the appendix.

different from zero. However for the third turning-point we are unable to reject the null hypothesis that the second derivative is equal to zero for this point which is necessary but not sufficient to indicate that the last turning-point (located at \$27,045) is not a relative extrema.

<i>Turning-point</i>	$\widehat{\frac{\partial^2 y}{\partial x^2}} \hat{\theta}_i$	<i>Standard error</i>	<i>95% Confidence Interval (\$10,000)</i>			
			<i>Delta</i>		<i>Inverse</i>	
$\hat{\theta}_1 = 0.81$	-4.15	0.31	-4.76	-3.54	-4.75	-3.55
$\hat{\theta}_2 = 1.85$	1.87	0.33	1.22	2.53	1.24	2.51
$\hat{\theta}_3 = 2.70$	-3.42	2.10	-7.61	0.77	-7.53	0.69

Table 3. The 95% CI for the second order derivative evaluated at each of the estimated turning-points with the corresponding delta and Inverse confidence intervals.

Alternately, we can use the inverse test to compute the confidence interval for the implied turning-points by computing the CI for the second derivative function evaluated at each turning-point. The plot of the 95% confidence interval for the second derivative function can be used to obtain equivalent intervals for Table 3. Figure 3 plots the first derivative function as the solid line and the 95% CI for the second derivative function as defined in (36) as the dotted lines. Note that the turning-points have been located by the vertical lines. The intervals for the second derivative as shown in Figure 3, can be found by locating the values along the reference lines defined at the turning-points of the 1st derivative function, that are cut by the CI of the 2nd derivative functions. These intervals are listed in Table 3. Figure 3 provides an indication of the regions for which the sign of the second derivative is positive, negative and indeterminate. Thus one can locate the CI for the 2nd derivative at those values of the regressor that coincide with the roots of the quartic defined in (42).

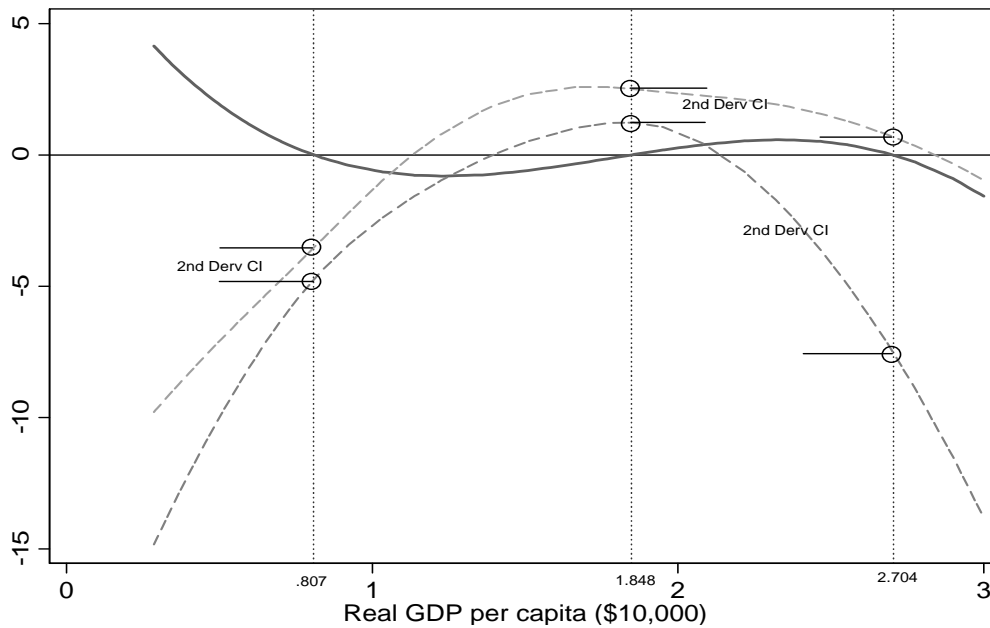


Figure 3 1st Derivative Function with 2nd Derivative 95% CI. The CI for the 2nd derivative function can be read at each turning-point.

In sum the location and classification of the turning-points can be established by use of the appropriate plot of the first and second derivatives of the estimated function along with the corresponding confidence intervals. Thus it is possible to implement the inversion test by the use of the appropriate plots of the 1st and 2nd derivatives and their associated CIs.

6. Fractional Polynomials

In order to demonstrate the general nature of the inverse test intervals we consider an extension of standard polynomials referred to as “fractional polynomials” proposed by Royston and Altman (1994). In the economics literature these specifications are similar to the application of the minflex Laurent expansion employed by Barnett (1983) to generalize the use of a second order Taylor series for demand system estimation. Fractional polynomials involve the use of the positive and negative integer powers as well as the log, square-root and inverse of the square root of the variable of interest. A primary application of this approach has been where regression models cannot be fit using the usual least squares approach (e.g. limited

dependent variables and other maximum likelihood models). Due to the more complex nature of the estimation needed for such methods as smoothing splines (as discussed in Ruppert et al 2003) and other “true” non-parametric techniques, fractional polynomials have been used where linear specifications can be accommodated by existing software such as in the application of the Cox proportional hazards model (see Kristiansen et al 2008).

A typical specification for a 3rd order fractional polynomial as identified by the powers $\{-3, -2, -1, -1/2, 0, 1/2, 1, 2, 3\}$ would be defined as:

$$y_t = \beta_0 + x_t^{-3}\beta_1 + x_t^{-2}\beta_2 + x_t^{-1}\beta_3 + x_t^{-1/2}\beta_4 + \beta_5 \ln x_t + x_t^{1/2}\beta_6 + x_t\beta_7 + x_t^2\beta_8 + x_t^3\beta_9 + \varepsilon_t \quad (45)$$

As in the cases for the regular polynomials we can define a vector of regressors that are functions of the covariate of interest (x).¹⁷

$$y_t = X_t B + \varepsilon_t$$

where

$$X_t = (1 \quad x_t^{-3} \quad x_t^{-2} \quad x_t^{-1} \quad x_t^{-1/2} \quad \ln(x_t) \quad x_t^{1/2} \quad x_t \quad x_t^2 \quad x_t^3) \quad (46)$$

And $B = (\beta_0 \quad \beta_1 \quad \dots \quad \beta_9)'$. We can define $w_t^2 = x_t$ so that the vector of regressors can be written in terms of integer powers:

$$X_t = (1 \quad w_t^{-6} \quad w_t^{-4} \quad w_t^{-2} \quad w_t^{-1} \quad 2\ln(w) \quad w_t \quad w_t^2 \quad w_t^4 \quad w_t^6) \quad (47)$$

To determine the locations of the extrema for this function we can evaluate the first order and second order derivatives using the vectors for the first and second derivatives defined as:

$$\frac{\partial X_t}{\partial w_t} = \begin{bmatrix} 0 & -6w_t^{-7} & -4w_t^{-5} & -2w_t^{-3} & -w_t^{-2} & 2w_t^{-1} & 1 & 2w_t & 4w_t^3 & 6w_t^5 \end{bmatrix} \quad (48)$$

¹⁷ Here we drop all other covariates to simplify our expressions. They would not influence the remainder of this section as long as they were not interacted with the x variable.

$$\frac{\partial^2 X_t}{\partial w_t^2} = \begin{bmatrix} 0 & 42w_t^{-8} & 20w_t^{-6} & 6w_t^{-4} & 2w_t^{-3} & -2w_t^{-2} & 0 & 2 & 12w_t^2 & 30w_t^4 \end{bmatrix} \quad (49)$$

Thus the estimated first derivative function is defined as $\left[\frac{\partial X_t}{\partial w_t} \right] \hat{\mathbf{B}}$ can be written as a 12th order polynomial defined as:

$$\left[\frac{\partial X_t}{\partial w_t} \right] \hat{\mathbf{B}} = w_t^{-7} \begin{pmatrix} 6\hat{\beta}_9 w_t^{12} + 4\hat{\beta}_8 w_t^{10} + 2\hat{\beta}_7 w_t^8 + \hat{\beta}_6 w_t^7 + 2\hat{\beta}_5 w_t^6 \\ -\hat{\beta}_4 w_t^5 - 2\hat{\beta}_3 w_t^4 - 4\hat{\beta}_2 w_t^2 - 6\hat{\beta}_1 \end{pmatrix} \quad (50)$$

Thus, (50) can be used to solve for the locations of the extrema in terms of the

estimated parameters $\hat{\beta}_i$ as $\hat{\theta}$ when $\left[\frac{\partial X_t}{\partial w_t} \right] \hat{\mathbf{B}} = 0$ thus we have:

$$\left(6\hat{\beta}_9 \hat{\theta}^{12} + 4\hat{\beta}_8 \hat{\theta}^{10} + 2\hat{\beta}_7 \hat{\theta}^8 + \hat{\beta}_6 \hat{\theta}^7 + 2\hat{\beta}_5 \hat{\theta}^6 - \hat{\beta}_4 \hat{\theta}^5 - 2\hat{\beta}_3 \hat{\theta}^4 - 4\hat{\beta}_2 \hat{\theta}^2 - 6\hat{\beta}_1 \right) = 0 \quad (51)$$

However, the roots of a 12th order polynomial have no analytic solution, thus the application of the traditional delta method would require the use of numeric derivatives for the approximation of the variance of the real valued roots. Or alternatively, one could derive confidence intervals for the roots via the use of a parametric bootstrap where one draws values of the coefficients from a multivariate normal pseudo-random number generator defined by using the estimated parameters and covariance.¹⁸

The test that we invert to locate the confidence intervals for the extrema is the test that the value of the first derivative to be equal to zero is defined as:

$$\frac{\left[\frac{\partial X_t}{\partial w_t} \right] \hat{\mathbf{B}}}{\sqrt{\left[\frac{\partial X_t}{\partial w_t} \right] \widehat{\text{cov}}(\hat{\mathbf{B}}) \left[\frac{\partial X_t}{\partial w_t} \right]'}} = t_{\alpha/2} \quad (52)$$

The value of the covariate (w) that satisfies this expression can be found by the solution of the polynomial in w defined by:

$$\left[\frac{\partial X_t}{\partial w_t} \right] \left(\hat{\mathbf{B}} \hat{\mathbf{B}}' - t_{\alpha/2}^2 \widehat{\text{cov}}(\hat{\mathbf{B}}) \right) \left[\frac{\partial X_t}{\partial w_t} \right]' = 0 \quad (53)$$

¹⁸ This technique may prove problematic when roots are only derived as real for a subset of simulations.

In this case it would be necessary to solve the 24th order polynomial. We list part of this 24 order polynomial in (54) to demonstrate that such applications can lead to highly complex expressions that would rarely be used for computation.¹⁹

$$\begin{aligned}
& \left(36\hat{\beta}_9^2 - 36t_{\alpha/2}^2 \hat{\sigma}_{99}\right) \tilde{\theta}^{24} \\
& + \left(48\hat{\beta}_8 \hat{\beta}_9 - 48t_{\alpha/2}^2 \hat{\sigma}_{89}\right) \tilde{\theta}^{22} \\
& + \left(16\hat{\beta}_8^2 + 24\hat{\beta}_7 \hat{\beta}_9 - 24t_{\alpha/2}^2 \hat{\sigma}_{79} - 16t_{\alpha/2}^2 \hat{\sigma}_{88}\right) \tilde{\theta}^{20} \\
& + \left(12\hat{\beta}_6 \hat{\beta}_9 - 12t_{\alpha/2}^2 \hat{\sigma}_{69}\right) \tilde{\theta}^{19} \\
& \quad \vdots \\
& \quad \vdots \\
& + \left(16\hat{\beta}_2^2 + 24\hat{\beta}_1 \hat{\beta}_3 - 24t_{\alpha/2}^2 \hat{\sigma}_{13} - 16t_{\alpha/2}^2 \hat{\sigma}_{22}\right) \tilde{\theta}^4 \\
& + \left(48\hat{\beta}_1 \hat{\beta}_2 - 48t_{\alpha/2}^2 \hat{\sigma}_{12}\right) \tilde{\theta}^2 \\
& + \left(36\hat{\beta}_1^2 - 36t_{\alpha/2}^2 \hat{\sigma}_{11}\right)
\end{aligned} \tag{54}$$

Again the roots of this polynomial could be determined by using a numeric solution as in Section 4.4. Furthermore, it may be difficult to interpret the implications for imaginary roots.

The alternative graphic solution can be found by defining the confidence intervals from the plots of the estimated 1st derivative function for any nonlinear specification. We only require the estimation of a linear combination of the estimated parameters to define the first derivative function as $\left[\frac{\partial X_t}{\partial w_t}\right] \hat{\mathbf{B}}$ and the confidence interval is given by:

$$\text{CI}\left(\frac{\partial Y_t}{\partial x_t}\right) = \left[\frac{\partial X_t}{\partial w_t}\right] \hat{\mathbf{B}} \pm t_{\frac{\alpha}{2}} \sqrt{\left[\frac{\partial X_t}{\partial w_t}\right] \widehat{\text{cov}}(\hat{\mathbf{B}}) \left[\frac{\partial X_t}{\partial w_t}\right]'} \tag{55}$$

This interval can be used to locate the confidence bounds for the turning points of the polynomial by locating the points where the CI of the first derivative cut the zero reference line.

¹⁹ This expression was obtained using the *MuMath* computer algebra routine.

The second order derivative function can also be plotted in order to establish the intervals for the classification of the turning-points. In this case we would use the confidence intervals for the second derivative model defined by:

$$CI\left(\frac{\partial^2 Y_t}{\partial x_t^2}\right) = \left[\frac{\partial^2 X_t}{\partial w_t^2}\right] \hat{B} \pm t_{\alpha/2} \sqrt{\left[\frac{\partial^2 X_t}{\partial w_t^2}\right] \widehat{\text{cov}}(\hat{B}) \left[\frac{\partial^2 X_t}{\partial w_t^2}\right]'} \quad (56)$$

In the same manner as we have discussed in the case of the application to the quartic described in Section 5.3.

6.1 An application of the Fractional Polynomial²⁰

Following in analogous fashion to the application in Section 5.3 for the quartic we reanalyse the Utt et al. (2001) data using a 3rd order fractional polynomial as defined in (45). The resulting estimated function with the data for estimation is plotted in Figure 4.

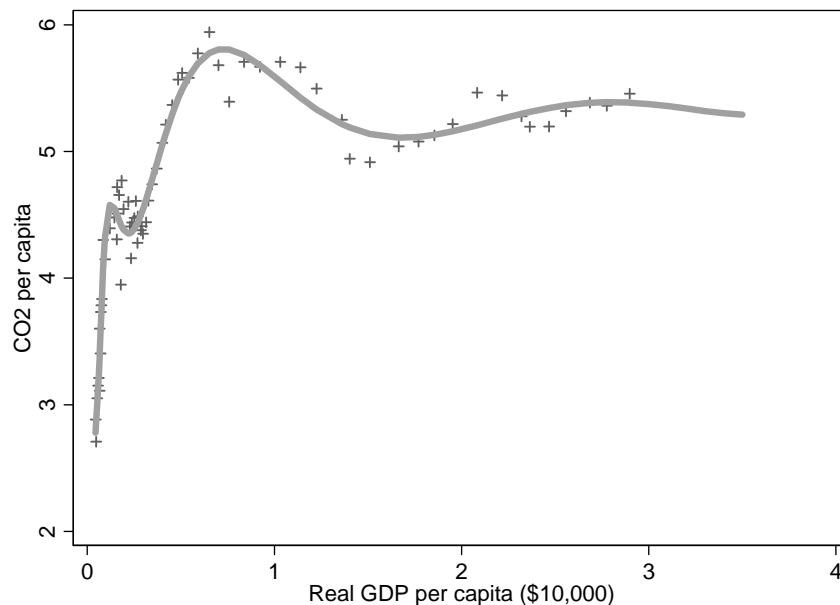


Figure 4 The scatter plot and fitted function based on 3rd order fractional polynomial.

By comparing Figures 1 and 4 one can note the two primary differences in the estimated functions. First, the fractional polynomial estimates a local maximum and a minimum between .2 and .4 that was missed by the quartic. Secondly, the fractional

²⁰ The *Stata* routine to generate the results in this section is given in Appendix A.2.

estimate does not exhibit the pronounced fall in the extrapolated values past the maximum value of the data that was present in the estimated quartic function. The estimated parameters for the 3rd order fractional polynomial are listed in Table 4. Note that not all are significantly different from zero but the fit is better than the quartic.

<i>Coefficient</i>	<i>Estimate</i>	<i>Std Error.</i>	<i>t-Stat</i>
$\hat{\beta}_0$	1721.771	651.417	2.64
$\hat{\beta}_1$	-0.014	0.005	-2.87
$\hat{\beta}_2$	1.598	0.503	3.18
$\hat{\beta}_3$	-166.668	50.490	-3.30
$\hat{\beta}_4$	1750.692	541.094	3.24
$\hat{\beta}_5$	1952.256	635.487	3.07
$\hat{\beta}_6$	-4217.859	1496.489	-2.82
$\hat{\beta}_7$	967.629	385.969	2.51
$\hat{\beta}_8$	-54.342	29.278	-1.86
$\hat{\beta}_9$	2.785	2.129	1.31
<i>Adjusted R²</i>	0.95	<i>N</i>	68

Table 4. Parameter estimates for the 3rd order fractional polynomial regression.

Due to the complications in the estimation of the turning-points and the confidence intervals based on the delta method and the inversion test, we will only use plots of the first and second derivatives and their corresponding confidence intervals to establish the nature of the turning-points in this case.

Figure 5 is a plot of the first derivative function with the corresponding 95% confidence interval.²¹ From this plot we can locate the 5 points where the first derivative function is equal to zero which correspond to the first maximum at \$1,200 then a minimum at approximately \$2,200 followed by a maximum at \$7,300, a minimum at \$16,700 and a maximum at \$28,000. From this plot we note that the confidence intervals for the first derivative indicate that once the level of GDP per capita exceeds \$22,200 the first derivative is never significantly different from zero, thus we can conclude that for values of GDP per capita over 22,200 there is no further

²¹ The *Stata* program used to create these results is listed in appendix A.2

variation in CO2 with respect to GDP per capita. The second order function can be used to establish the nature of the turning-points identified by the first derivative.

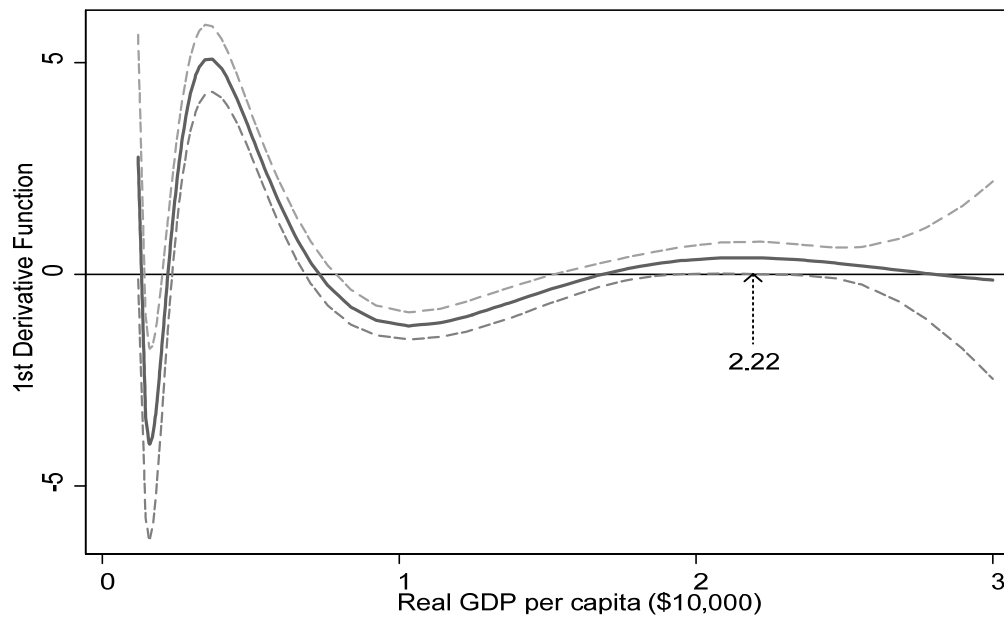


Figure 5 The first derivative function of the 3rd order fractional polynomial function with the corresponding 95% CI.

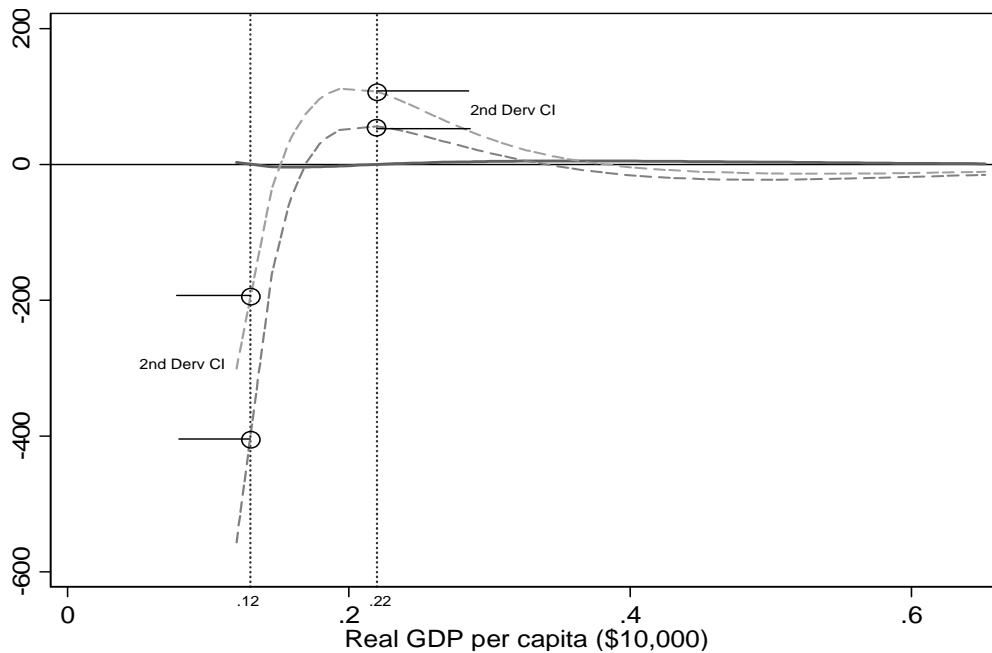


Figure 6 The first derivative function and the second derivative confidence interval for GDP per capita from 0 to \$7,000.

In order to more accurately determine the location of the turning points and the confidence intervals we have plotted the first derivative function and the confidence interval for the second derivative function in two figures defined for two

different ranges of the regressor. Figure 6 shows the second derivative confidence interval and the first derivative function for values of the per capita income under \$7,000. Figure 7 has the same plot from \$7,000 to \$30,000. From Figure 6 we find that the second derivative test indicates that the first turning-point at \$1,200 is a relative maximum, the second turning-point at \$2,200 is a relative minimum.

From Figure 7 we note that the third turning-point at \$7,000 is a maximum and the fourth turning-point at \$16,700 is a minimum. However, we can also read from Figure 7 that the confidence interval for the second derivative function evaluated at the position of the last turning-point at \$28,000 includes zero thus we cannot reject the null hypothesis that the turning-point is neither a maximum nor a minimum.

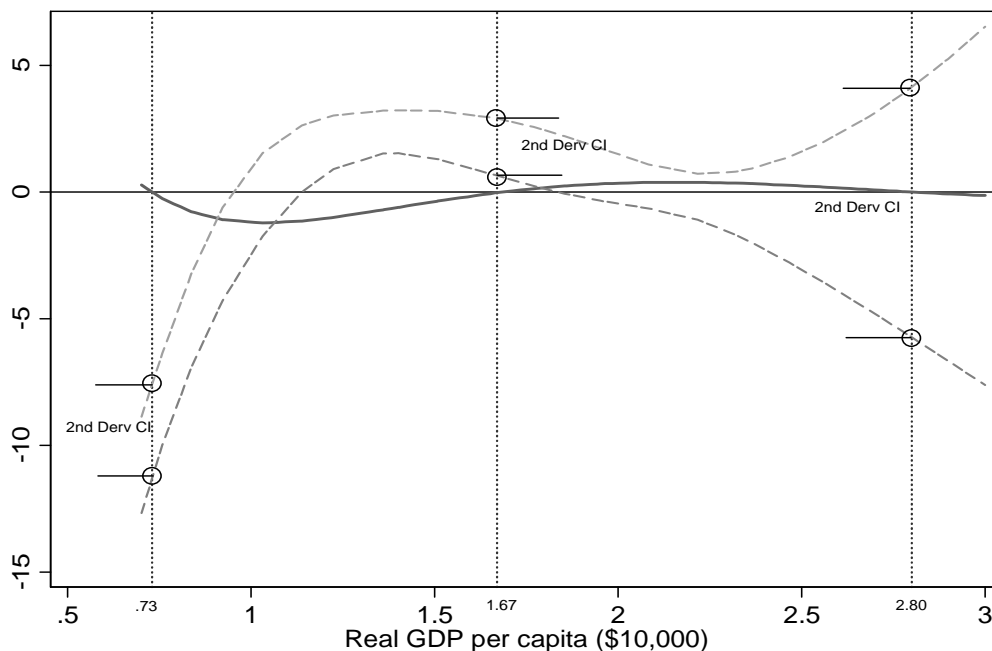


Figure 7 The first derivative function and the second derivative confidence interval for GDP per capita from \$7,000 to \$30,000.

7. Conclusions

In this paper we have presented a method for the determination of the location and properties of turning-points estimated for non linear models specified in regressions. By the construction of the inverse test we can locate the limits for

confidence intervals of high order polynomials as well as for fractional polynomials as the roots to high order polynomials. We then demonstrate that the results for the first derivative function can be extended to the second derivative function which can provide tests for the determination of the nature of the turning-points identified by the first derivative. Alternatively, we have shown that graphic methods can be used to provide equivalent results and that they can be much more easily employed especially for models that may be specified that do not have an analytical solution found from higher order polynomials.

Future research in this area could be directed to the consideration of the more complex results obtained when multiple regressors are allowed to interact in an extension of the general response surface model. In addition to the fractional polynomials it would also be possible to use the methods proposed here to evaluate the turning-points for other polynomial related specifications such as the “Fourier Flexible Form” specification as proposed by Gallant (1981) and the class of Chebyshev polynomials (see Smyth 2005 for an example). These specifications are similar to the fractional polynomial in that they also allow for the definition of higher order derivative functions that are linear in the regression parameters thus they can be plotted in a similar manner to the fractional polynomial examined here.

References

- Bagliani, M., G. Bravo and S. Dalmazzone, 2008, "A consumption-based approach to environmental Kuznets curves using the ecological footprint indicator", *Ecological Economics*, 65, 650-661.
- Barnett, W. A., 1983, "New Indices of Money Supply and the Flexible Laurent Demand System", *Journal of Business and Economic Statistics*, 1, 7-23.
- Bartus, T., 2005, "Estimation of marginal effects using margeff", *The Stata Journal*, 5, 309-329.
- Buonaccorsi, J. P. (1979), "On Fieller's Theorem and the General Linear Model", *The American Statistician*, 33, 162
- Clark, A., A. Oswald and P. Warr, 1996, "Is job satisfaction U-shaped in age?", *Journal of Occupation and Organizational Psychology*, 69, 57-81.
- Coate, B. and T. Fry, 2012, "Better off Dead? Prices realised for Australian Paintings sold at Auction", *ACEI working paper*, AWP-02-2012.
- Dufour, J., 1997, "Some Impossibility Theorems in Econometrics With Applications to Structural and Dynamic Models", *Econometrica*, 65, 1365-1387.
- Fieller, E. C., 1944, "A Fundamental Formula in the Statistics of Biological Assay and Some Applications", *Quarterly Journal of Pharmacy and Pharmacology*, 17, 117-123.
- French, M. and G. Zarkin, 1995, "Is moderate alcohol use related to wages? Evidence from four worksites", *Journal of Health Economics*, 14, 319-344.
- Gallant, A. R., 1981, "On the bias in flexible functional forms and an essentially unbiased form The Fourier flexible form", *Journal of Econometrics*, 15, 211-245.
- Hellbronner, E., 1979, "Position and Confidence Limits of an Extremum", *Journal of Chemical Education*, 56, 240-243.
- Hirschberg, J. G. and J. N. Lye, 2010a, "A Geometric Comparison of the Delta and Fieller Confidence Intervals", *The American Statistician*, 64, 234-241.
- Hirschberg, J. G. and J. N. Lye, 2010b, "Two geometric representations of confidence intervals for ratios of linear combinations of regression parameters: An application to the NAIRU", *Economics Letters*, 108, 73-76.
- Hirschberg, J. G., J. N. Lye and D. J. Slottje, 2008, "Inferential Methods for Elasticity Estimates", *Journal of Econometrics*, 147, 299-315.
- Huang, H., Y. Lin and C. Yeh, 2012, "An appropriate test of the Kuznets hypothesis", *Applied Economics Letters*, 19, 47-51.

- Jenkins, M. A. and J. F. Traub, 1970, "A Three-Stage Algorithm for Real Polynomials Using Quadratic Iteration", *SIAM Journal on Numerical Analysis*, 7, 545-566.
- Kearsley, A. and M. Riddell, 2010, "A further inquiry into the Pollution Haven Hypothesis and the Environmental Kuznets Curve", *Ecological Economics*, 69, 905-191.
- Kristiansen, L., M. Grønbaek, U. Becker, and J. S. Tolstrup, 2008, "Risk of Pancreatitis According to Drinking Habits: A Population-based Cohort Study", *American Journal of Epidemiology*, 168, 932-937.
- Kumar, K. and B. Viswanathan, 2007, "Changing structure of income indoor air pollution relationships in India", *Energy Policy*, 35, 5496-5504.
- Lind, J. and H. Mehlum, 2010, "With or without U? The appropriate test for a U-shaped relationship", *Oxford Bulletin of Economics and Statistics*, 72, 109-118.
- Liu, G., T. Skjerpen, and K. Telle, 2009, "Unit roots, polynomial transformations and the environmental Kuznets curve", *Applied Economics Letters*, 16, 285-288.
- Naghshpour, S., 2005, "The cyclical nature of family income distribution in the United States: an empirical note", *Journal of Economics and Finance*, 29, 138-143.
- Plassmann, F. and N. Khanna, 2007, "Assessing the precision of turning-point estimates in polynomial regression functions", *Econometric Reviews*, 26, 503-528.
- Rao, C. R., 1973, *Linear Statistical Inference and Its Applications*, 2nd ed., John Wiley & Sons, New York, NY.
- Richmond, A. and R. Kaufmann, 2006, "Is there a turning-point in the relationship between income and energy use and/or carbon emissions?", *Ecological Economics*, 56, 176-189
- Royston, P. and D.G. Altman, 1994, "Regression using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling", *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 43, 429-467.
- Ruppert, D., M. P. Wand and R. J. Carroll, 2003, *Semiparametric Regression*, Cambridge University Press, UK.
- Sasabuchi, S., 1980, "A test of a multivariate normal mean with composite hypotheses determined by linear inequalities", *Biometrika*, 67, 429-439.
- Scheffé, H. (1970), "Multiple Testing versus Multiple Estimation. Improper Confidence Sets. Estimation of Directions and Ratios", *The Annals of Mathematical Statistics*, 41, 1-29.

- Schnabel, C. and J. Wagner, 2008, "Union Membership and Age: The Inverted U-shape Hypothesis under Test", Institute for the Study of Labor (IZA) Discussion Paper No. 3842.
- Smyth, G. K. (2005), "Polynomial Approximation", *Encyclopaedia of Biostatistics*, 2nd ed, John Wiley & Sons, Ltd. (on line edition)
- Utt, J. A., W. W. Hunter, R. E. McCormick, 2001, "On the Relationship between Net Carbon Emissions and Income", mimeo, Washington State University.
- Vandaele, Walter, 1981, "Wald, Likelihood Ratio and Lagrange Multiplier Tests as an F Test", *Economics Letters*, 8, 361-365.
- Vitaliano, D. and J. Held, 1991, "Road accident external effects: an empirical assessment", *Applied Economics*, 23, 373-378.
- Williams, E. J., 1959, *Regression Analysis*, John Wiley & Sons, New York, NY.
- Winer, S. L., M. W. Tofias, B. Grofman and J. H. Aldrich, 2008, "Trending economic factors and the structure of Congress in the growth of government, 1930-2002", *Public Choice*, 135, Pages 415-448
- Wong, M., 2011, "Estimating the Impact of the Ethnic Housing Quotas in Singapore", downloaded <http://maisy.wharton.upenn.edu/research> 8th March 2012.
- Wooldridge, J. M., 2010, *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Mass.
- Zerbe, G. O., 1978, "On Fieller's Theorem and the General Linear Model", *The American Statistician*, 32, 103-105.
- Zerbe, G.O., Laska, E., Meisner, M., and Kushner, A.B. 1982. "On Multivariate Confidence Regions and Simultaneous Confidence Limits for Ratios", *Communications in Statistics, Theory and Methods* 11, 2401-2425.

Appendix A.1 Stata program to generate the results in Section 5.3

```

/*
    Read the Carbon data from Utt etal and run the quartic regression
    Save the estimated parameters, the predicted values and plot the function and the data.
    Use the mono colour scheme for all graphs.
*/
*/
set scheme s1mono
use "C:\data\old_v\cubic\carbon_stock\utt.dta" , clear

label variable co2 "CO2 per capita"
label variable g "Real GDP per capita ($10,000)"

regress co2 g c.g#c.g c.g#c.g#c.g c.g#c.g#c.g#c.g
estimates store eq_1
predict p_co2
label variable p_co2 "Predicted CO2"

twoway (scatter co2 g, msymbol(+)) (line p_co2 g, sort lwidth(1)), ylabel(,nogrid) ///
ytitle("CO2 per capita") legend(off) name(Functionfrac, replace)

/*
    Define the roots of the quartic and apply the delta method to estimate confidence intervals
    Using nlcom to compute the non-linear function of the parameters to find the roots of the quartic
    and the delta standard errors with the confidence interval. By using sequential calls to nlcom
    with the post option we can build up the complex functions for these roots.
*/
*/
quietly nlcom      (a1: (3*_b[c.g#c.g#c.g])/(4*_b[c.g#c.g#c.g#c.g])) ///
                  (a2: (_b[c.g#c.g])/(2*_b[c.g#c.g#c.g#c.g])) ///
                  (a3: (_b[g])/(4*_b[c.g#c.g#c.g#c.g])) ///
                  (b2: _b[c.g#c.g]) (b3: _b[c.g#c.g#c.g]) (b4: _b[c.g#c.g#c.g#c.g]), post

quietly nlcom      (a1: (_b[a1])) ///
                  (q1: ( (_b[a1]^2-3*_b[a2])/9 )) ///
                  (q2: (2*_b[a1]^3 - 9*_b[a1]*_b[a2] + 27*_b[a3])/54 )) ///
                  (b2: _b[b2]) (b3: _b[b3]) (b4: _b[b4]), post

nlcom      (x1: -2*sqrt(_b[q1])*cos( acos((_b[q2])/sqrt((_b[q1]^3))) /3) - (_b[a1]/3)) ///
          (x2: -2*sqrt(_b[q1])*cos( (acos((_b[q2])/sqrt((_b[q1]^3)))+ 2*_pi)/3) - (_b[a1]/3)) ///
          (x3: -2*sqrt(_b[q1])*cos( (acos((_b[q2])/sqrt((_b[q1]^3)))+ 4*_pi)/3) - (_b[a1]/3)) ///
          (t1a: 2*_b[b2] + 6*_b[b3]*.8076 + 12*_b[b4]*.8076^2 ) ///
          (t2b: 2*_b[b2] + 6*_b[b3]*2.7045 + 12*_b[b4]*2.7045^2 ) ///
          (t3c: 2*_b[b2] + 6*_b[b3]*1.8476 + 12*_b[b4]*1.8476^2 )

/*
    Retrieve the original quartic regression parameter and covariance estimates
*/
*/
estimates restore eq_1

/*
    Store the parameters from the quartic as b1-b4, and the elements of the covariance matrix as Vij
*/
*/
scalar b1=_b[g]
scalar b2=_b[c.g#c.g]
scalar b3=_b[c.g#c.g#c.g]
scalar b4=_b[c.g#c.g#c.g#c.g]

matrix V = get(VCE)
scalar v34 = V[3,4]
scalar v24 = V[2,4]
scalar v14 = V[1,4]
scalar v23 = V[2,3]
scalar v13 = V[1,3]
scalar v12 = V[1,2]
scalar v11 = V[1,1]
scalar v22 = V[2,2]
scalar v33 = V[3,3]
scalar v44 = V[4,4]

/*
    Define the 6th order polynomial with roots that are the equivalent 95% inverse CIs for
    each inflexion point. Set t_s to the square of the critical value for two tailed t-test.
*/
*/
scalar t_ = 1.96
scalar t_s = t_^2

scalar c0 = b1^2 - t_s*v11
scalar c1 = 4*b1*b2 - 4*t_s*v12
scalar c2 = 6*b1*b3 + 4*b2^2 - 6*t_s*v13 - 4*t_s*v22
scalar c3 = 12*b2*b3 + 8*b1*b4 - 12*t_s*v23 - 8*t_s*v14
scalar c4 = 9*b3^2 + 16*b2*b4 - 9*t_s*v33 - 16*t_s*v24

```

```

scalar c5 = 24*b3*b4 - 24*t_s*v34
scalar c6 = 16*b4^2 - 16*t_s*v44

/*
Use the mata program polyroots to compute the roots of this polynomial.
st_numscalar inserts the scalar values to the polyroot program
*/
mata polyroots((st_numscalar(("c0")),st_numscalar(("c1")),st_numscalar(("c2")), ///
st_numscalar(("c3")), st_numscalar(("c4")),st_numscalar(("c5")),st_numscalar(("c6"))))

/*
Use the margins program to estimate the first derivative function and the estimated 95%CI then
plot with marginsplot along with the indicators for the solution found from the polyroots command
to show they coincide with the Fieller type bounds.
*/
quiet margins, dydx(g) at(g=(0.25(0.05)3.5))

marginsplot, recast(line) recastci(rline) ciopts(color(*.7)) yline(0, lstyle(foreground)) ///
name(FirstDerv1, replace) plotregion(fcolor(gs16)) ylabel(,nogrid) ///
xline(3.41144302 2.45864884 2.09915614 1.65867625 .904947489 .732782011 , lpattern(".") lwidth( .2 ))

/*
Plot out the 1st and 2nd derivative functions to find the regions for inflexion points as opposed to
maxima and minima using predictnl
*/
predictnl d1 = _b[c.g] + 2 *_b[c.g#c.g] * g + 3 *_b[c.g#c.g#c.g] * g^2 ///
+ 4 *_b[c.g#c.g#c.g#c.g] * g^3 , se(d1_se)
generate up1 = d1 + t_ * d1_se
generate lw1 = d1 - t_ * d1_se

line d1 up1 lw1 g, lwidth( .4 .3 .3) lpattern( "1" "-#-" "-#-" ) sort ///
xline(0.8076194 2.704491 1.847637, lpattern(".") lwidth( .2 ) ) ///
name(d1, replace)yline(0, lstyle(foreground)) ///
ylabel(,nogrid) ytitle("1st Derivative Function") legend(off)

predictnl d2 = 2 *_b[c.g#c.g] + 6 *_b[c.g#c.g#c.g] * g + 12 *_b[c.g#c.g#c.g#c.g] * g^2 , se(d2_se)
generate up2 = d2 + t_ * d2_se
generate lw2 = d2 - t_ * d2_se

line d1 up2 lw2 g if lw2 > -15 , lwidth( .4 .3 .3) lpattern( "1" "-#-" "-#-" ) sort ///
xline(0.8076194 2.704491 1.847637, lpattern(".") lwidth( .2 ) ) ///
name(d1_d2, replace)yline(0, lstyle(foreground)) ///
ylabel(,nogrid) title("1st Derivative Function with 2nd Derivative 95% CI") legend(off)

twoway (scatter co2 g) (line p_co2 g, sort), name(Function2, replace) ///
xline(3.41144302 2.45864884 2.09915614 1.65867625 .904947489 .732782011, lpattern(".") lwidth( .2 )) ///
ylabel(,nogrid) title("Function with turning-point intervals") legend(off)

line p_co2 d1 d2 up2 lw2 g, ///
name(all, replace) lstyle(p4 p1 p2 p2) yline(0,lstyle(foreground)) ///
xline(3.41144302 2.45864884 2.09915614 1.65867625 .904947489 .732782011, lpattern(".") lwidth( .2 )) ///
ylabel(,nogrid) title("Function with turning-point intervals") legend(off)

```

Appendix A.2 Stata program to generate the results in Section 6.2

```

/*
    Read the Carbon data from Utt etal and run a fractional polynomial
    with powers -3 -2 -1 -1/2 0 1/2 1 2 3
    Save the estimated parameters, the predicted values and plot the function and the data
    Use the mono-colour scheme for all graphs.
*/
use "C:\data\old_v\cubic\carbon_stock\utt.dta" , clear
label variable co2 "CO2 per capita"
label variable g "Real GDP per capita ($10,000)"
set scheme smono

generate ig3 = 1/g^3
generate ig2 = 1/g^2
generate ig1 = 1/g
generate igs = 1/sqrt(g)
generate g0 = log(g)
generate gs = sqrt(g)
generate g1 = g
generate g2 = g^2
generate g3 = g^3

regress co2 ig3 ig2 ig1 igs g0 gs g1 g2 g3
estimates store eq_1
predict p_co2
label variable p_co2 "Predicted CO2"
twoway (scatter co2 g, msymbol(+)) (line p_co2 g, sort lwidth(1)), ylabel(,nogrid) ///
ytitle("CO2 per capita") legend(off) name(Functionfrac, replace)

/*
    Plot out the 1st and 2nd derivative functions to find the regions for inflexion points as opposed to
    maxima and minima
*/
generate t_ = 1.96
predictnl d1 = -2*_b[ig2] / g^3 - _b[ig1] / g^2 + _b[g0] / g + _b[g1] + 2*_b[g2] * g ///
              - .5 * _b[igs]*ig1*igs + .5 * _b[gs]* igs ///
              - 3 * _b[ig3] / g^4 + 3*_b[g3] * g^2 , se(d1_se)

generate up1 = d1 + t_ * d1_se
generate lw1 = d1 - t_ * d1_se

line d1 up1 lw1 g if ((g > .095 )*( g < 3.1)), lwidth( .4 .3 .3) lpattern( "1" "-#" "-#-" ) sort , ///
ylabel(,nogrid) ytitle("1st Derivative Function") legend(off) ///
name(d1frac, replace)yline(0, lstyle(foreground))

predictnl d2 = 6*_b[ig2] / g^4 + 2*_b[ig1] / g^3 - _b[g0] / g^2 + 2*_b[g2] ///
              + .75 * _b[igs]*ig2*igs - .25 * _b[gs]*ig1* igs ///
              + 12 * _b[ig3] / g^5 + 6 * _b[g3] * g , se(d2_se)

generate up2 = d2 + t_ * d2_se
generate lw2 = d2 - t_ * d2_se

line d1 up2 lw2 g if ((g < .7 )*( g > .12)), lwidth( .4 .3 .3) lpattern( "1" "-#" "-#-" ) sort ///
xline(.73 1.67 2.8 , lpattern(".") lwidth( .3 ) ) ///
name(d1_d2frac1, replace)yline(0, lstyle(foreground)) ///
ylabel(,nogrid) title("1st Derivative Function with 2nd Derivative 95% CI") legend(off)

line d1 up2 lw2 g if ((g < .7 )*( g > .12)), lwidth( .4 .3 .3) lpattern( "1" "-#" "-#-" ) sort ///
xline(.13 .22 , lpattern(".") lwidth( .3 ) ) ///
name(d1_d2frac2, replace)yline(0, lstyle(foreground)) ///
ylabel(,nogrid) title("1st Derivative Function with 2nd Derivative 95% CI") legend(off)

```


Appendix B Data used for example from Utt et al (2001).

co2	g	co2	g
4.3	0.085003	5.068	0.400714
3.786	0.074033	5.211	0.418914
3.212	0.06162	5.368	0.453311
2.709	0.046979	5.567	0.485656
2.882	0.044872	5.619	0.506366
3.05	0.05218	5.582	0.542679
3.151	0.057553	5.774	0.58988
3.601	0.065376	5.943	0.652405
3.732	0.071262	5.681	0.701269
3.109	0.066247	5.392	0.758539
3.404	0.070367	5.707	0.836963
3.835	0.076747	5.669	0.921925
4.149	0.094976	5.707	1.030728
4.393	0.12005	5.664	1.138708
4.479	0.14524	5.496	1.224937
4.716	0.158819	5.25	1.360127
4.51	0.159439	4.943	1.4017
4.305	0.157226	4.915	1.509263
4.656	0.169435	5.041	1.663832
4.771	0.18359	5.078	1.769473
3.949	0.17917	5.127	1.854223
4.545	0.193692	5.216	1.951673
4.603	0.219915	5.464	2.082665
4.408	0.228283	5.442	2.216931
4.439	0.237771	5.278	2.319561
4.156	0.234249	5.195	2.364967
4.476	0.250976	5.197	2.46678
4.609	0.260075	5.318	2.557764
4.488	0.269218	5.386	2.684405
4.277	0.268288	5.361	2.774956
4.378	0.286005	5.455	2.898218
4.409	0.291215		2.9
4.35	0.296448		3
4.442	0.313843		3.1
4.611	0.326307		3.2
4.74	0.345757		3.3
4.863	0.370008		3.4
			3.5