Journal of Econometrics 187 (2015) 376–384

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

# Testing linearity using power transforms of regressors[☆]

Yae In Baek[a], Jin Seo Cho[b,*], Peter C.B. Phillips[c,d,e,f]

[a] *Department of Economics, University of California, San Diego, 9500 Gilman Dr. La Jolla, CA, 92093, USA*
[b] *School of Economics, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul, 120-749, Republic of Korea*
[c] *Yale University, USA*
[d] *University of Auckland, New Zealand*
[e] *Singapore Management University, Singapore*
[f] *University of Southampton, United Kingdom*

ARTICLE INFO

*Article history:*
Received 15 April 2014
Received in revised form
18 March 2015
Accepted 18 March 2015
Available online 1 April 2015

*JEL classification:*
C12
C18
C46
C52

*Keywords:*
Box–Cox transform
Gaussian stochastic process
Neglected nonlinearity
Power transformation
Quasi-likelihood ratio test
Trend exponent
Trifold identification problem

ABSTRACT

We develop a method of testing linearity using power transforms of regressors, allowing for stationary processes and time trends. The linear model is a simplifying hypothesis that derives from the power transform model in three different ways, each producing its own identification problem. We call this modeling difficulty the *trifold identification problem* and show that it may be overcome using a test based on the quasi-likelihood ratio (QLR) statistic. More specifically, the QLR statistic may be approximated under each identification problem and the separate null approximations may be combined to produce a composite approximation that embodies the linear model hypothesis. The limit theory for the QLR test statistic depends on a Gaussian stochastic process. In the important special case of a linear time trend regressor and martingale difference errors asymptotic critical values of the test are provided. Test power is analyzed and an empirical application to crop-yield distributions is provided. The paper also considers generalizations of the Box–Cox transformation, which are associated with the QLR test statistic.



## 1. Introduction

Linear models are a natural starting point in empirical work. They also relate in a fundamental way to underlying Gaussian assumptions and the use of wide sense conditional expectations. Testing linearity is therefore a familiar practice in applications whenever there is concern over specification and Gaussianity. Such tests fall within the framework of general model specification tests.

Power transformations are especially popular as alternatives to linearity. Tukey (1957, 1977) provides several rationales for the use of power transformations, and Box and Cox (1964) further developed their use in nonlinear modeling. The Box–Cox transformation, in particular, successfully implements the so-called Tukey 'ladder of power' option. In time series applications, some studies (notably, Wu (1981) and Phillips (2007)) considered power transforms of a time trend, providing limit theories that are useful in estimation and inference concerning the relevant parameters.

Power transformations can be used to form tests that deliver consistent power against arbitrary alternatives to linearity. As Stinchcombe and White (1998) showed, any non-polynomial analytic function can be used to construct generically comprehensively revealing (GCR) tests, in the sense that linear projection errors are not necessarily orthogonal to any power transform when the linear model is misspecified. This property motivates use

☆ The co-editor, Jianqing Fan, Associated editor, and two anonymous referees provided helpful comments for which we are grateful. We have also benefited from discussions with Shun-ichiro Bessho, Seonghoon Cho, Stan Hurn, Isao Ishida, Tae-Hwan Kim, Shandre Thangavelu, Byungsam Yoo, Valentin Zelenyuk, and other participants at the NZESG meeting (Auckland, 2013) and the 8th Joint Economics Symposium of Four Leading East Asian Universities (Shanghai, 2014). Baek and Cho acknowledge research support from the second stage BK21 project and research grant (NRF-2010-332-B00025) of the National Research Foundation of Korea, respectively, and Phillips acknowledges support from a Kelly Fellowship and the National Science Foundation of USA under Grant No. SES 12-58258.

* Corresponding author.
*E-mail address:* jinseocho@yonsei.ac.kr (J.S. Cho).

http://dx.doi.org/10.1016/j.jeconom.2015.03.041
0304-4076/

of power transforms for constructing tests with omnibus power. In spite of this apparently useful property, testing linearity using power transforms is largely undeveloped in the literature, mainly because of the identification problem that arises under the null of linearity. As detailed below, the linear model hypothesis can be deduced from a power transformation in three different ways, each of which involves its own identification problem, a feature that we call the *trifold identification problem*. To our knowledge, this problem has never before been addressed in the literature.

Our primary goal in the present paper is to resolve this complex trifold problem. Our focus is pragmatic and involves constructing mechanisms needed in using power transformations. We focus on models involving power transforms of a strictly stationary (SS) variable or a time trend. While this excludes some possibilities, such as nonlinear transforms of nonstationary variates (e.g. Park and Phillips (1999), and Shi and Phillips (2012)), the range of potential applications is large and includes both microeconometric and time series data.

This paper restricts attention to a particular statistic, the quasi-likelihood ratio (QLR) statistic. As we demonstrate, the QLR statistic may produce a composite form that embodies the linear model hypothesis. An additional benefit from focusing on the QLR test is its relationship to the Box–Cox transformation. The score of the test turns out to be related to an augmented form of Box–Cox transform. Our approach to developing a null approximation of the QLR test extends the methodology of Cho and Ishida (2012), who studied how to test the effects of omitted power transformations. We advance that work and compare our null approximation with the QLR tests that are popular in the artificial neural network (ANN) literature where there is at most a twofold identification problem. Our approach also exploits the properties of time-trend power transforms and regressions studied recently in Phillips (2007). Time trend regressors and their power transforms have very different properties from those of stationary regressors in view of the asymptotic degeneracy of the signal matrix.

The paper is organized as follows. Section 2 examines power transformations of a stationary process and tests linearity. The null approximation and the power properties of the QLR test are developed. Section 3 extends the discussion and asymptotic results to power transforms of a time-trend regressor. Simulations and empirical applications are contained in Sections 4 and 5, respectively. Concluding remarks are given in Section 6. All proofs are collected in an Appendix to the paper which is available as an online supplement (Baek et al., 2014).

## 2. Testing for neglected power transforms of a stationary regressor

We seek to model the conditional mean $\mathbb{E}[Y_t|\mathbf{W}_t]$ of a dependent variable $Y_t$ given a collection of explanatory variables $\mathbf{W}_t$. We define the class of (parameter dependent) conditional mean functions as $m_t(\boldsymbol{\omega}) := \alpha + \mathbf{W}_t'\boldsymbol{\delta} + \beta X_t^{\gamma} = \mathbb{E}[Y_t|\mathbf{W}_t]$, where the parameter vector $\boldsymbol{\omega} := (\alpha, \boldsymbol{\delta}', \beta, \gamma)' \in \boldsymbol{\Omega} \subset \mathbb{R}^{k+4}$, with $\boldsymbol{\delta} \in \mathbb{R}^{k+1}$ for some $k \in \mathbb{N}$. In this specification, the variables $(Y_t, \mathbf{W}_t)$ comprise a strictly stationary and absolutely regular mixing process, the variable $X_t$ is positively valued, and $\boldsymbol{\Omega}$ is the parameter space of $\boldsymbol{\omega}$. In addition to appearing nonlinearly as $X_t^{\gamma}$, the variable $X_t$ also enters linearly in $m_t(\boldsymbol{\omega})$ so that $X_t$ is the first element of $\mathbf{W}_t$. Then $\mathbf{W}_t = (X_t, \mathbf{D}_t')'$ for some $\mathbf{D}_t \in \mathbb{R}^k$. Similarly, we partition the parameter vector $\boldsymbol{\delta} := (\xi, \boldsymbol{\eta}')'$, so that $\mathbf{W}_t\boldsymbol{\delta} = \xi X_t + \mathbf{D}_t'\boldsymbol{\eta}$. In Section 3, $X_t$ is a linear time trend and so the conditional mean function includes both a linear and nonlinear (power function) trend.

Our interest is primarily in testing the effective form of $X_t$ in the conditional mean $\mathbb{E}[Y_t|\mathbf{W}_t]$. We consider the following explicit hypotheses. Given that $\mathbb{E}[Y_t|\mathbf{W}_t]$ is linear with respect to the components $(1, \mathbf{W}_t)$, we focus on the null hypothesis $\mathcal{H}_0$ : $\exists(\alpha_*, \boldsymbol{\delta}_*)$, $\mathbb{E}[Y_t|\mathbf{W}_t] = \alpha_* + \mathbf{W}_t'\boldsymbol{\delta}_*$ w.p. 1 and the alternative hypothesis $\mathcal{H}_1$ : $\forall(\alpha, \boldsymbol{\delta})$, $\mathbb{E}[Y_t|\mathbf{W}_t] = \alpha + \mathbf{W}_t'\boldsymbol{\delta}$ w.p. $< 1$, which implies that nonlinear elements of $X_t$ appear in the conditional mean that cannot be embodied in $\mathcal{H}_0$. The affix '$*$' is used to parameterize $\mathbb{E}[Y_t|\mathbf{W}_t]$, so that for some $\alpha_o$ and $\beta_o$, $(\alpha_*, \beta_*, \gamma_*) \in \{(\alpha, \beta, \gamma) : \alpha + \beta X_t^{\gamma} = \alpha_o \text{ or } \alpha + \beta X_t^{\gamma} = \beta_o X_t\}$ under $\mathcal{H}_0$.

Testing the linear model hypothesis using a maintained model with a nonlinear component is common practice in the literature. Such tests may be regarded as a variant of the Bierens (1990) test. Similarly, Stinchcombe and White's (1998) GCR tests are constructed to test for a nonlinear component. A power transform representation is particularly popular for the nonlinear component. For example, Tukey (1957, 1977) introduced power transform flexible nonlinear models, and Box and Cox (1964) found that their transformation accords with Tukey's (1957) 'ladder of power' and it has been widely applied in empirical work (e.g. Sakia (1992)). The GCR property is delivered by non-polynomial analytic functions that can approximate arbitrary functions by Taylor expansion, so that for some $\gamma_*$, $\mathbb{E}[V_t X_t^{\gamma_*}] \neq 0$ in a misspecified linear model, where $V_t$ denotes the linear projection error. This property motivates the construction of power transforms to test linearity. The literature already has related variations of power transforms such as those used in Ramsey's (1969) test which have prefixed power exponents. The general power transforms used here do not fix power exponents, and this flexibility is used to gain powers in testing, as detailed below.

Notwithstanding considerable interest in power transforms, $\mathcal{H}_0$ has not been formally examined in the literature mainly because testing $\mathcal{H}_0$ cannot be conducted in a standard way. There are three different identification problems that arise under $\mathcal{H}_0$. If $\beta_* = 0$, $\gamma_*$ is not identified and Davies' (1977,1987) identification problem arises. On the other hand, if $\gamma_* = 0$, $\alpha_* + \beta_*$ is identified, but neither $\alpha_*$ nor $\beta_*$ is separately identified. Furthermore, if $\gamma_* = 1$ and $\boldsymbol{\delta}_*$ is conformably partitioned as $(\xi_*, \boldsymbol{\eta}_*)'$, $\xi_* + \beta_*$ is identified although neither $\xi_*$ nor $\beta_*$ is identified. Thus, *three* different identification problems arise under the linear model hypothesis. We denote these three hypotheses as $\mathcal{H}_0'$ : $\beta_* = 0$; $\mathcal{H}_0''$ : $\gamma_* = 0$; and $\mathcal{H}_0'''$ : $\gamma_* = 1$ and call this construct the *trifold identification problem*.

The current literature approaches the trifold identification problem only in a limited way. Hansen (1996), for instance, provided a testing methodology that employs the weighted bootstrap to treat $\mathcal{H}_0'$. Alternatively, the power coefficient might be fixed as in Ramsey (1969), so that the identification problems under $\mathcal{H}_0''$ and $\mathcal{H}_0'''$ are avoided. Accordingly, the main goal of the current study is to provide a tractable test that is able to handle the trifold identification problem within a unified framework without losing power.

Some related identification problems have appeared in the literature. Cho et al. (2011, 2014) test for neglected nonlinearity using ANN models and find that two different identification problems arise under the null of linearity. They show how this twofold identification problem may be addressed using the QLR test. Cho and Ishida (2012) similarly test for effects of power transforms using the same QLR statistic but their focus of interest differs from ours and their model has only a twofold identification problem. None of this work considers nonlinear trend effects.

The approach taken in the current work is to extend the analysis of Cho et al. (2011, 2014) and Cho and Ishida (2012). The maximum order involved in the null approximation used in Cho et al. (2011) is the fourth order, whereas that used in Cho et al. (2014) is the sixth order. They observe that the maximum order depends on the activation function used in constructing the test. On the other hand, Cho and Ishida (2012) use a second-order approximation, as is common in econometric practice. The present paper examines how these approximations are modified by the trifold identification problem.