Center for the Study of Finance and Insurance
Osaka University

Discussion Paper Series 2011-06

Testing for the Effects of Omitted Power Transformation

Jin Seo Cho and Isao Ishida

Center for the Study of Finance and Insurance
Osaka University

Testing for the Effects of Omitted Power Transformation

JIN SEO CHO
School of Economics
Yonsei University
Email: jinseocho@yonsei.ac.kr

ISAO ISHIDA
CSFI
Osaka University
i-ishida@sigmath.es.osaka-u.ac.jp

First version: September 2006
This version: November 2010

Abstract

This study considers testing the hypothesis that the coefficient of power-transformed conditioning variable is zero or not. The hypothesis is not standard because there is an unidentified parameter under the null hypothesis, and also because the null can be tested by imposing zero power coefficient. That is, the so-called twofold identification problem arises. We examine the asymptotic null distribution of the quasi-likelihood ratio (QLR) test and show that it can be obtained by employing a conventional quadratic expansion, which is a different consequence from Cho, Ishida, and White (*Neural Computation*, 2010), in which a fourth-order Taylor expansion is the minimum order of model expansion. This difference is caused from that power-transformed conditioning variable is different from the other conditioning variables.

Key Words: Power transformation, twofold identification, quasi-likelihood ratio test.

Subject Class: C12, C13.

1 Introduction

We consider the following simple model using power transformation:

$$Y_t = \alpha + \mathbf{W}_t'\boldsymbol{\delta} + \beta X_t^{\gamma} + U_t, \tag{1}$$

where $(Y_t, X_t, \mathbf{W}_t')' \in \mathbb{R}^{2+k}$ ($k \in \{0\} \cup \mathbb{N}$) are identically and independently distributed (IID); X_t is positively valued; and $\mathbf{Z}'\mathbf{Z} = \sum_{t=1}^{n} \mathbf{Z}_t\mathbf{Z}_t'$ is nonsingular, where $\mathbf{Z}_t := (1, \mathbf{W}_t')'$, and n is the sample size. We also suppose that $\mathbf{Z}_t$ does not contain a constant and X_t.

Econometric models using power transformations are popular and also often easily applied due to their flexibility. For example, Turkey [6] considers its properties and advocates use of power transformation when linear model is misspecified. As another example, Box and Cox [1] reparameterize Eq. (1) by letting $\alpha = -\lambda/\gamma$ and $\beta = \lambda/\gamma$ for some unknown λ to secure the continuity of power transformation (see Turkey [6]) and estimate λ and γ separately. This yields the so-called Box-Cox transformation. That is, Eq. (1) is now transformed into

$$Y_t = \mathbf{W}_t'\boldsymbol{\delta} + \lambda \left(\frac{X_t^{\gamma} - 1}{\gamma} \right) + U_t. \tag{2}$$

Thus, estimating and inferring Eq. (2) is equivalent to Eq. (1) by the invariance principle. Many other interesting power transformations are obtained by adjusting the coefficients in Eq. (1).

Our interest is in testing the effect of X_t to $E[Y_t|\mathbf{Z}_t]$, and the hypotheses can be given as

$$\mathcal{H}_0 : \exists(\alpha_*, \boldsymbol{\delta}_*), E[Y_t|\mathbf{Z}_t] = \alpha_* + \mathbf{W}_t'\boldsymbol{\delta}_* \text{ w.p. } 1 \text{ ; vs. } \mathcal{H}_1 : \forall(\alpha, \boldsymbol{\delta}), E[Y_t|\mathbf{Z}_t] = \alpha + \mathbf{W}_t'\boldsymbol{\delta} \text{ w.p. } < 1. \tag{3}$$

Here, subscript '$*$' is used to parameterize $E[Y_t|\mathbf{Z}_t]$, and $(\beta_*, \gamma_*) \in \{(\beta, \gamma) : \beta X_t^{\gamma} = \text{constant}\}$ under $\mathcal{H}_0$.

These hypotheses are motivated by its abundant applicabilities. First, a researcher may suffer from the loss of power of t-statistic by incorrectly specifying a linear model for $E[Y_t|\mathbf{Z}_t]$. We instead consider the power transformation as an alternative. Second, the researcher may think of X_t as an omitted variable with unknown functional form. As a simple example, $\mathbf{W}_t$ and X_t can be thought of as a set of dummy variables and a continuous random variable, respectively. When the functional form of $E[Y_t|\mathbf{Z}_t]$ is not necessarily linear with respect to X_t, a natural extension is supposing a flexible function for X_t. Here, the power function serves for this purpose.

Our test statistic is the quasi-likelihood ratio (QLR) test considered by Cho, Ishida, and White [2]:

$$QLR_n := n(1 - \hat{\sigma}_{n,A}^2/\hat{\sigma}_{n,0}^2),$$

where

$$\hat{\sigma}^2_{n,A} := \inf_{\alpha,\beta,\gamma,\boldsymbol{\delta}} \frac{1}{n}\sum_{t=1}^{n}(Y_t - \alpha - \mathbf{W}_t'\boldsymbol{\delta} - \beta X_t^{\gamma})^2, \quad \text{and} \quad \hat{\sigma}^2_{n,0} := \inf_{\alpha,\boldsymbol{\delta}} \frac{1}{n}\sum_{t=1}^{n}(Y_t - \alpha - \mathbf{W}_t'\boldsymbol{\delta})^2.$$

As Cho, Ishida, and White [2] point out, testing the hypotheses in (3) is not standard because it does have identification problem two different ways under $\mathcal{H}_0$. If $\beta_* = 0$, then γ_* is not identified, so that Davies' [3, 4] identification problem arises. On the other hand, if $\gamma_* = 0$, then $\alpha_* + \beta_*$ is identified, but each of α_* and β_* is not identified. Cho, Ishida, and White [2] call this *twofold identification* and show that the QLR test may be successfully used for the twofold identification. In particular, a quartic expansion was necessary for obtaining the asymptotic null distribution in their context.

The goal of this study is twofold. The first goal is in examining the QLR test and resolving the twofold identification problem under our new context. The QLR test is not always successful in resolving the twofold identification as pointed out by Cho, Ishida, and White [2], and our model is different from theirs in that $\mathbf{W}_t$ does not contain X_t. This aspect lets us examine this model separately. The second goal stems from our desire to search for a model pursued by the literature. More specifically, Hansen [5] provides regularity conditions for handling Davies's [3, 4] identification problem, and Box-Cox transformation is treated as a special case of his analysis. Nevertheless, Hansen [5] does not assume the presence of twofold identification, and this lets his analysis be restrictive in analyzing general power transformations: the space for γ has to be restricted not to contain zero and avoid a quartic expansion. We show that the model of this study can be indeed analyzed using a conventional quadratic expansion, although it has a twofold identification. This mainly follows from that the flexible function we examine is a power function.

The plan of this note is similar to that of Cho, Ishida, and White [2]. In Section 2, we consider the QLR statistic under $\beta_* = 0$ and introduce another statistic which is asymptotically equivalent to the QLR statistic under $\beta_* = 0$. In Section 3, we consider $\gamma_* = 0$ and proceed as in Section 2. In Section 4, we show how the QLR statistics in Section 2 and 3 are stochastically associated. Finally, concluding remarks are given in Section 5.

2 QLR Statistic under $\beta_* = 0$

For notational simplicity, we let the quasi-likelihood (QL) and concentrated QL (CQL) be

$$L_n(\alpha, \beta, \gamma, \boldsymbol{\delta}) := -\sum_{t=1}^{n}(Y_t - \alpha - \beta X_t^{\gamma} - \mathbf{W}_t'\boldsymbol{\delta})^2,$$

$$L_n(\beta;\gamma) := L_n(\hat{\alpha}_n(\beta;\gamma), \beta, \gamma, \hat{\boldsymbol{\delta}}_n(\beta;\gamma))$$

respectively, where $(\hat{\alpha}_n(\beta;\gamma), \hat{\boldsymbol{\delta}}_n(\beta;\gamma)')' := \arg\max_{\alpha,\boldsymbol{\delta}} L_n(\alpha,\beta,\gamma,\boldsymbol{\delta})$. We note that γ_* is not identified if $\beta_* = 0$, so that we fix γ for the moment and minimize the QL with respect to identified parameters $(\alpha_*, \boldsymbol{\delta}_*)$. The CQL is denoted as $L_n(\beta;\gamma)$, and its specific form is

$$L_n(\beta;\gamma) = -\mathbf{U}'\mathbf{M}\mathbf{U} + 2\beta\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U} - \beta^2\mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma)$$

using $\beta_* = 0$, where $\mathbf{U} := (U_1, U_2, \ldots, U_n)'$, $\mathbf{M} := \mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$, and $\mathbf{X}(\gamma) := (X_1^\gamma, X_2^\gamma, \ldots, X_n^\gamma)'$.

The CQL is conventionally approximated by a second-order Taylor expansion around $\beta_* = 0$ (e.g. Hansen [5]). That is,

$$L_n(\beta;\gamma) = L_n(0;\gamma) + L_n^{(1)}(0;\gamma)\beta + \frac{1}{2}L_n^{(2)}(0;\gamma)\beta^2,$$

where $L_n^{(k)}(0;\gamma) := \frac{\partial^k}{\partial\beta^k}L_n(\beta;\gamma)|_{\beta=0}$ for $k = 0, 1, 2, \ldots$. Thus, if we maximize the CQL with respect to β for each γ,

$$\sup_\beta\{L_n(\beta;\gamma) - L_n(0;\gamma)\} = \sup_\beta\left\{L_n^{(1)}(0;\gamma)\beta + \frac{1}{2}L_n^{(2)}(0;\gamma)\beta^2\right\} = -\frac{\{L_n^{(1)}(0;\gamma)\}^2}{2L_n^{(2)}(0;\gamma)} = \frac{\{\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\}^2}{\mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma)}. \tag{4}$$

We consider this maximization process to examine the following QLR statistic:

$$QLR_n^{(1)} := \sup_\gamma \sup_\beta n\left\{1 - \frac{L_n(\beta;\gamma)}{L_n(0;\gamma)}\right\} = \frac{\{n^{-1/2}\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\}^2}{\hat{\sigma}_{n,0}^2\{n^{-1}\mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma)\}}$$

using the fact that $\hat{\sigma}_{n,0}^2 = -n^{-1}L_n(0;\gamma)$. The statistic $QLR_n^{(1)}$ is defined to accommodate the fact that γ is not identified under $\beta_* = 0$. We let n tend to infinity before maximizing the CQL with respect to γ.

It is not hard to provide regularity conditions for obtaining the weak convergence of $QLR_n^{(1)}$ under $\mathcal{H}_0$. Under suitable regularity conditions as considered by Hansen [5], $n^{-1/2}\mathbf{X}(\,\cdot\,)'\mathbf{M}\mathbf{U}$ is tight, and the uniform law of large numbers holds for $n^{-1}\mathbf{X}(\,\cdot\,)'\mathbf{M}\mathbf{X}(\,\cdot\,)$. Further, $\hat{\sigma}_{n,0}^2$ converges to a constant in probability. This implies that $QLR_n^{(1)}$ weakly converges to $\sup_\gamma \mathcal{G}(\gamma)^2$ under the null $\beta_* = 0$, where $\mathcal{G}$ is a Gaussian stochastic process with

$$E[\mathcal{G}(\gamma)\mathcal{G}(\gamma')] = \frac{\rho(\gamma,\gamma')}{\sqrt{\kappa(\gamma,\gamma)}\sqrt{\kappa(\gamma',\gamma')}},$$

where

$$\rho(\gamma,\gamma') := E[U_t^2 X_t^{\gamma+\gamma'}] - E[X_t^{\gamma}\mathbf{Z}_t]E[\mathbf{Z}_t\mathbf{Z}_t']^{-1}E[U_t^2\mathbf{Z}_t X_t^{\gamma'}] \\ - E[X_t^{\gamma'}\mathbf{Z}_t]E[\mathbf{Z}_t\mathbf{Z}_t']^{-1}E[U_t^2\mathbf{Z}_t X_t^{\gamma}] + E[X_t^{\gamma}\mathbf{Z}_t']E[\mathbf{Z}_t\mathbf{Z}_t']^{-1}E[U_t^2\mathbf{Z}_t\mathbf{Z}_t']E[\mathbf{Z}_t\mathbf{Z}_t']^{-1}E[\mathbf{Z}_t X_t^{\gamma'}],$$

and $\kappa(\gamma,\gamma) := \sigma_*^2\{E[X_t^{2\gamma}] - E[X_t^{\gamma}\mathbf{Z}_t']E[\mathbf{Z}_t\mathbf{Z}_t']^{-1}E[\mathbf{Z}_t X_t^{\gamma}]\}$.

3 QLR Statistic under $\gamma_* = 0$

We now iterate this procedure to obtain the asymptotic null distribution under $\gamma_* = 0$. If $\gamma_* = 0$, $\alpha_* + \beta_*$ is identified, although each parameter is not separately identified. For resolving this, we proceed in two ways: first, we fix β to identify α_* and obtain the asymptotic null distribution. Alternatively, we fix α and identify β_*.

3.1 When β_* Is Not Identified

We first fix β and obtain the CQL with respect to identified parameters $(\alpha_*, \boldsymbol{\delta}_*)$ as before. We let

$$L_n(\gamma;\beta) := L_n(\hat{\alpha}_n(\gamma;\beta), \beta, \gamma, \hat{\boldsymbol{\delta}}_n(\gamma;\beta)),$$

where $(\hat{\alpha}_n(\gamma;\beta), \hat{\boldsymbol{\delta}}_n(\gamma;\beta)')' := \arg\max_{\alpha,\boldsymbol{\delta}} L_n(\alpha,\beta,\gamma,\boldsymbol{\delta})$. By this, the CQL is explicitly stated as

$$L_n(\gamma;\beta) = -\mathbf{U}'\mathbf{M}\mathbf{U} + 2\beta\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U} - \beta^2\mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma).$$

A second-order Taylor expansion is now applied. That is, if $L_n(\gamma;\beta)$ is approximated at around $\gamma_* = 0$,

$$L_n(\gamma;\beta) = L_n(0;\beta) + L_n^{(1)}(0;\beta)\gamma + \frac{1}{2}L_n^{(2)}(0;\beta)\gamma^2 + o_p(1),$$

so that

$$\sup_{\gamma}\{L_n(\gamma;\beta) - L_n(0;\beta)\} = \sup_{\gamma}\left\{L_n^{(1)}(0;\beta)\gamma + \frac{1}{2}L_n^{(2)}(0;\beta)\gamma^2 + o_p(1)\right\} \\ = -\frac{\{L_n^{(1)}(0;\beta)\}^2}{2L_n^{(2)}(0;\beta)} + o_p(1) = \frac{\{\beta\mathbf{L}_1'\mathbf{M}\mathbf{U}\}^2}{\beta^2\mathbf{L}_1'\mathbf{M}\mathbf{L}_1 - \beta\mathbf{L}_2\mathbf{M}\mathbf{U}} + o_p(1), \qquad (5)$$

where $L_n^{(k)}(0;\beta) := \frac{\partial^k}{\partial\gamma^k}L_n(\gamma;\beta)|_{\gamma=0}$ for $k = 0, 1, 2, \ldots$, $\mathbf{L}_1 := (\ln X_1, \ln X_2, \ldots, \ln X_n)'$, and $\mathbf{L}_2 :=$

$[(\ln X_1)^2, (\ln X_2)^2, \ldots, (\ln X_n)^2]'$. Here, we could apply a second-order Taylor expansion from the fact that $L_n^{(1)}(0;\beta)$ is not necessarily equal zero, whereas Cho, Ishida, and White [2] have to rely on a fourth-order Taylor expansion as it is zero under their context. This makes testing neglected nonlinearity by a power function be different from what Cho, Ishida, and White [2] examine, given that X_t is not contained in $\mathbf{W}_t$. We also note that the right-hand side (RHS) of (5) is free of β, provided that $\mathbf{L}_2\mathbf{MU} = o_p(n)$, which easily holds under mild regularities, as $E[(\ln X_t)^2 U_t] = 0$ and $E[\mathbf{Z}_t U_t] = \mathbf{0}$. We thus obtain that

$$\sup_{\gamma}\{L_n(\gamma;\beta) - L_n(0;\beta)\} = \frac{\{\mathbf{L}_1'\mathbf{MU}\}^2}{\mathbf{L}_1'\mathbf{ML}_1} + o_p(1),$$

and the asymptotic distribution of the following QLR statistic:

$$QLR_n^{(2)} := \sup_{\beta}\sup_{\gamma} n\left\{1 - \frac{L_n(\gamma;\beta)}{L_n(0;\beta)}\right\}$$

is simply identical to that of

$$\frac{\{n^{-1/2}\mathbf{L}_1'\mathbf{MU}\}^2}{\hat{\sigma}_{n,0}^2\{n^{-1}\mathbf{L}_1'\mathbf{ML}_1\}}. \tag{6}$$

Here, we used the fact that $L_n(0;\beta) = -n\hat{\sigma}_{n,0}^2$. We can apply the central limit theorem and law of large numbers to the numerator and denominator, respectively, so that $QLR_n^{(2)}$ weakly converges to a noncentral chi-square random variable. That is, a Gaussian process does not have to be introduced for this limit, although β_* is not identified.

3.2 When α_* Is Not Identified

When $\gamma_* = 0$, we can identify the model in another way. That is, we can fix α and identify $(\beta_*, \boldsymbol{\delta}_*)$. For examining this, we first let $(\hat{\beta}_n(\gamma;\alpha), \hat{\boldsymbol{\delta}}_n(\gamma;\alpha)')' := \arg\max_{\beta,\boldsymbol{\delta}} L_n(\alpha,\beta,\gamma,\boldsymbol{\delta})$, which is

$$\begin{bmatrix} \hat{\beta}_n(\gamma;\alpha) \\ \hat{\boldsymbol{\delta}}_n(\gamma;\alpha) \end{bmatrix} = \begin{bmatrix} \sum X_t^{2\gamma} & \sum X_t^{\gamma}\mathbf{W}_t' \\ \sum \mathbf{W}_t X_t^{\gamma} & \sum \mathbf{W}_t\mathbf{W}_t' \end{bmatrix}^{-1} \begin{bmatrix} \sum (Y_t - \alpha)X_t^{\gamma} \\ \sum (Y_t-\alpha)\mathbf{W}_t \end{bmatrix},$$

and we now obtain the QL as

$$\begin{aligned} L_n(\gamma;\alpha) &:= L_n(\alpha, \hat{\beta}_n(\gamma;\alpha), \gamma, \hat{\boldsymbol{\delta}}_n(\gamma;\alpha)) \\ &= \sum (Y_t-\alpha)^2 - \begin{bmatrix} \sum (Y_t - \alpha)X_t^{\gamma} \\ \sum (Y_t-\alpha)\mathbf{W}_t \end{bmatrix}' \begin{bmatrix} \sum X_t^{2\gamma} & \sum X_t^{\gamma}\mathbf{W}_t' \\ \sum \mathbf{W}_t X_t^{\gamma} & \sum \mathbf{W}_t\mathbf{W}_t' \end{bmatrix}^{-1} \begin{bmatrix} \sum (Y_t - \alpha)X_t^{\gamma} \\ \sum (Y_t-\alpha)\mathbf{W}_t \end{bmatrix}. \end{aligned}$$

We can also approximate this with respect to $\gamma_* = 0$. For this, we obtain the first two derivatives:

$$L_n^{(1)}(0;\alpha) = 2(\alpha_* - \alpha)\underbrace{\mathbf{L}_1'\mathbf{M}\mathbf{U}}_{O_p(\sqrt{n})} + 2\underbrace{\mathbf{U}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}}_{O_p(1)} - \underbrace{\mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z})(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}}_{O_p(1)}$$

under mild regularities, where $L_n^{(j)}(0;\alpha) := \frac{\partial^j}{\partial\gamma^j}L_n(\gamma;\alpha)|_{\gamma=0}$ and $\mathbf{K}_j := [\mathbf{L}_j \vdots \mathbf{0}_{n\times k}]$ for $j = 1, 2, \ldots$, so that $L_n^{(1)}(0;\alpha) = 2(\alpha_* - \alpha)\mathbf{L}_1'\mathbf{M}\mathbf{U} + o_p(\sqrt{n})$. Here, $L_n^{(1)}(0;\alpha)$ is not necessarily equal to zero. Thus, a second-order Taylor expansion is enough. Also,

$$\begin{aligned}
L_n^{(2)}(0;\alpha) = &- 2(\alpha_* - \alpha)^2\mathbf{L}_1'\mathbf{M}\mathbf{L}_1 \\
&+ 2(\alpha_* - \alpha)[\mathbf{L}_1'\mathbf{M}\mathbf{U} + 2\mathbf{L}_1'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} - 2\mathbf{L}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{U}] \\
&+ 2\mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}[\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z} - \mathbf{K}_1'\mathbf{K}_1 - \mathbf{Z}'\mathbf{K}_2](\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} \\
&+ 2[\mathbf{U}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{U} + \mathbf{U}'\mathbf{K}_2(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} - 2\mathbf{U}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\{\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z}\}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U}]
\end{aligned}$$

after some algebra. It is not hard to find sufficient conditions for having

- $\mathbf{L}_1'\mathbf{M}\mathbf{L}_1 = O_p(n)$;
- $\mathbf{L}_1'\mathbf{M}\mathbf{U} + 2\mathbf{L}_1'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} - 2\mathbf{L}_1'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{U} = O_p(\sqrt{n})$;
- $\mathbf{U}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\{\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z} - \mathbf{K}_1'\mathbf{K}_1 - \mathbf{Z}'\mathbf{K}_2\}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} = O_p(1)$; and
- $\mathbf{U}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{K}_1'\mathbf{U} + \mathbf{U}'\mathbf{K}_2(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} - 2\mathbf{U}'\mathbf{K}_1(\mathbf{Z}'\mathbf{Z})^{-1}\{\mathbf{Z}'\mathbf{K}_1 + \mathbf{K}_1'\mathbf{Z}\}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{U} = O_p(1)$,

either. Thus, we may conclude that $L_n^{(2)}(0;\alpha) = -2(\alpha_* - \alpha)^2\mathbf{L}_1'\mathbf{M}\mathbf{L}_1 + o_p(n)$ under mild conditions. Finally, it now follows that

$$\sup_\gamma\{L_n(\gamma;\alpha) - L_n(0;\alpha)\} = -\frac{\{L_n^{(1)}(0;\alpha)\}^2}{2L_n^{(2)}(0;\alpha)} + o_p(1) = \frac{\{\mathbf{L}_1'\mathbf{M}\mathbf{U}\}^2}{\mathbf{L}_1'\mathbf{M}\mathbf{L}_1} + o_p(1).$$

The unidentified parameter α is canceled off as in the previous subsection, and the weak limit of the QLR test defined as

$$QLR_n^{(3)} := \sup_\alpha\sup_\gamma n\left\{1 - \frac{L_n(\gamma;\alpha)}{L_n(0;\alpha)}\right\} = \frac{\{n^{-1/2}\mathbf{L}_1'\mathbf{M}\mathbf{U}\}^2}{\hat{\sigma}_{n,0}^2\{n^{-1}\mathbf{L}_1'\mathbf{M}\mathbf{L}_1\}} + o_p(1)$$

is asymptotically equivalent to (6) under $\gamma_* = 0$. Here, we also used that $L_n(0;\alpha) = -n\hat{\sigma}_{n,0}^2$. From this, $QLR_n^{(2)}$ is asymptotically equivalent to $QLR_n^{(3)}$.

4 The Relationship between the Three QLR Statistics

The separate weak limits we obtained in the previous sections are not independent. We can derive their stochastic interrelationship by letting γ converge to zero from the QLR test in Section 2. For examining this, we let $N_n(\gamma)$ and $D_n(\gamma)$ be defined as

$$N_n(\gamma) := \{\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U})\}^2 \quad \text{and} \quad D_n(\gamma) := \mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma),$$

respectively. Note that theses are the numerator and denominator of the RHS in (4). Also,

$$\operatorname*{plim}_{\gamma\to 0} N_n(\gamma) = 0 \quad \text{and} \quad \operatorname*{plim}_{\gamma\to 0} D_n(\gamma) = 0$$

because $\mathrm{plim}_{\gamma\to 0}\mathbf{X}(\gamma) = \boldsymbol{\iota}$ and $\mathbf{M}$ is an idempotent matrix constructed by $\mathbf{Z}_t$. We cannot simply apply L'Hôspital's rule using this. We therefore obtain

$$\operatorname*{plim}_{\gamma\to 0} \frac{d}{d\gamma} N_n(\gamma) = \operatorname*{plim}_{\gamma\to 0} 2\{\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\}\left\{\frac{d}{d\gamma}\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\right\} = 0;$$

$$\operatorname*{plim}_{\gamma\to 0} \frac{d}{d\gamma} D_n(\gamma) = \operatorname*{plim}_{\gamma\to 0} 2\left\{\frac{d}{d\gamma}\mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma)\right\} = 0,$$

and this shows that it is necessary to apply L'Hôspital's rule one more time. That is, from that

$$\operatorname*{plim}_{\gamma\to 0} \frac{d^2}{d\gamma^2} N_n(\gamma) = \operatorname*{plim}_{\gamma\to 0} 2\left\{\frac{d^2}{d\gamma^2}\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\right\}^2 + 2\{\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\}\left\{\frac{d^2}{d\gamma^2}\mathbf{X}(\gamma)'\mathbf{M}\mathbf{U}\right\} = 2\{\mathbf{L}_1\mathbf{M}\mathbf{U}\}^2;$$

$$\operatorname*{plim}_{\gamma\to 0} \frac{d^2}{d\gamma^2} D_n(\gamma) = \operatorname*{plim}_{\gamma\to 0} 2\left\{\frac{d^2}{d\gamma^2}\mathbf{X}(\gamma)'\mathbf{M}\mathbf{X}(\gamma) + \frac{d}{d\gamma}\mathbf{X}(\gamma)'\mathbf{M}\frac{d}{d\gamma}\mathbf{X}(\gamma)\right\} = 2\mathbf{L}_1\mathbf{M}\mathbf{L}_1,$$

it now follows that

$$\operatorname*{plim}_{\gamma\to 0} \frac{N_n(\gamma)}{\hat{\sigma}^2_{n,0} D_n(\gamma)} = \operatorname*{plim}_{\gamma\to 0} \frac{N_n^{(1)}(\gamma)}{\hat{\sigma}^2_{n,0} D_n^{(1)}(\gamma)} = \operatorname*{plim}_{\gamma\to 0} \frac{N_n^{(2)}(\gamma)}{\hat{\sigma}^2_{n,0} D_n^{(2)}(\gamma)} = \frac{\{\mathbf{L}_1\mathbf{M}\mathbf{U}\}^2}{\hat{\sigma}^2_{n,0}\mathbf{L}_1\mathbf{M}\mathbf{L}_1},$$

where $N_n^{(k)}(\gamma) := \frac{\partial^k}{\partial\gamma^k} N_n(\gamma)$ and $D_n^{(k)}(\gamma) := \frac{\partial^k}{\partial\gamma^k} D_n(\gamma)$, and $k = 1, 2$. Here, we exploited the simple fact that $\mathrm{plim}_{\gamma\to 0}\frac{d}{d\gamma}\mathbf{X}(\gamma) = \mathbf{L}_1$ and $\mathrm{plim}_{\gamma\to 0}\frac{d^2}{d\gamma^2}\mathbf{X}(\gamma) = \mathbf{L}_2$. We note that these probability limits are now asymptotically equivalent to $QLR_n^{(2)}$ or $QLR_n^{(3)}$, respectively. Also, neither $QLR_n^{(2)}$ nor $QLR_n^{(3)}$ can be greater than $QLR_n^{(1)}$. Thus,

$$QLR_n = \max[QLR_n^{(1)}, QLR_n^{(2)}, QLR_n^{(3)}] = QLR_n^{(1)} \Rightarrow \sup_{\gamma} \mathcal{G}(\gamma)^2, \quad (7)$$

which yields the asymptotic null distribution of the QLR test, and we achieve our first goal by this.

The implication of (7) is that we could obtain the previous result without a quartic approximation. The asymptotic null distribution in Cho, Ishida, and White [2] had to exploit L'Hôspital's rule four times successively, whereas we here use L'Hôspital's rule only twice. Further, Section 3 doesn't have to use quartic expansions to obtain the asymptotic distributions, which is also different from Cho, Ishida, and White [2]. This follows mainly from the fact that X_t is powered and is not contained in $\mathbf{W}_t$. In case X_t were one of the variables in $\mathbf{W}_t$, the analysis using the quartic approximation in Cho, Ishida, and White [2] should have been applied. From this, we can conclude that that our model can be analyzed in the framework of a quadratic approximation, achieving our second goal.

5 Conclusion

We examine testing for the effect of omitted power transformation using the QLR test in the framework of Cho, Ishida, and White [2]. The specification in Cho, Ishida, and White [2] had to exploit a *quartic* expansion, whereas we didn't have to exploit this for our model: a *quadratic* expansion is sufficient for our case, although it has a twofold identification. The main source of this is that the conditioning variables we employ for linearity are different from the conditioning variable for power transformation. Otherwise, a quartic approximation should have been exploited.

References

[1] BOX, G. AND COX, D. (1964): "An Analysis of Transformations," *Journal of the Royal Statistical Society*, *Series B*, 26, 211-252.

[2] CHO, J., ISHIDA, I., AND WHITE, H. (2010): "Revisting Tests for Neglected Nonlinearity Using Aritificial Neural Network," *Neural Computation*, forthcoming.

[3] DAVIES, R. (1977): "Hypothesis Testing When a Nuisance Parameter is Present only under the Alternative," *Biometrika*, 64, 247-254.

[4] DAVIES, R. (1987): "Hypothesis Testing When a Nuisance Parameter is Present only under the Alternative," *Biometrika*, 74, 33-43.

[5] HANSEN, B. (1996): "Inference When a Nuisance Parameter is Not Identified under the Null Hypothesis," *Econometrica*, 64, 413-430.

[6] TURKEY, J. (1957): "The Comparative Anatomy of Transformations," *Annals of Mathematical Statistics*, 28, 602-632.