Econometric Theory, **15**, 1999, 99–113. Printed in the United States of America.

ANALYTICAL POWER COMPARISONS OF NESTED AND NONNESTED TESTS FOR LINEAR AND LOGLINEAR REGRESSION MODELS

MASAHITO KOBAYASHI
Yokohama National University

MICHAEL MCALEER
University of Western Australia

This paper compares several tests for linear and loglinear regression models where both the dependent and independent variables are transformed. It is shown that the Lagrange multiplier test proposed by Godfrey and Wickens (1981, *Review of Economic Studies* 48, 487–496) in the framework of the Box–Cox regression model has the highest asymptotic power of the compared tests. The extended projection test of MacKinnon, White, and Davidson (1983, *Journal of Econometrics* 11, 53–70), the test of Bera and McAleer (1983, paper presented to the SSRC Econometric Study Group Conference on Model Specification and Testing, Warwick; 1989, *Sankhya B* 51, 212–224), and the test of Andrews (1971, *Biometrika* 58, 249–254) are shown to have asymptotically equivalent powers and to have lower powers than the nonnested test of Cox (1961, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, 105–123, Berkeley: University of California Press; 1962, *Journal of the Royal Statistical Society B* 24, 406–424).

1. INTRODUCTION

A basic problem in econometric modeling is whether the variables should be expressed in logarithms or in levels. For example, in estimating the Cobb–Douglas production function, the logarithms of the variables are appropriate, whereas the levels of the variables may be more suitable for estimating the ag-

We thank a co-editor and a referee for helpful suggestions and seminar participants at CORE, the Catholic University of Louvain, Erasmus University Rotterdam, Hiroshima University, Kyoto University, Southern Methodist University, Texas A&M University, Tilburg University, and the University of Western Australia for helpful comments and suggestions. Michael McAleer acknowledges the financial support of the Australian Research Council. Much of the paper was written while Michael McAleer was visiting the Tinbergen Institute at Erasmus University Rotterdam and CentER at Tilburg University. An earlier version of the paper was presented at the Seventh World Congress of the Econometric Society, Tokyo, August 1995. Address correspondence to: Masahito Kobayashi, Faculty of Economics, Yokohama National University, Hodogaya-ku, Yokohama 240, Japan; e-mail: mkoba@econ.ynu.ac.jp.

© 1999 Cambridge University Press 0266-4666/99 $9.50

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

gregate consumption function: constant elasticities are assumed in the former case and constant slopes in the latter. These alternative functional form hypotheses can be tested against each other by the use of nonnested tests, such as those of Cox (1961, 1962), which is based on the modified likelihood ratio, the extended projection (PE) test of MacKinnon, White, and Davidson (1983), and the test of Bera and McAleer (1983, 1989) (the B&M test), which are based on variable addition (for a useful recent survey of tests for nonnested hypotheses, see Gourieroux and Monfort, 1994). Kobayashi (1994) compared the asymptotic local powers of these nonnested tests when the logarithmic transformation is applied only to the dependent variable.

Linear and loglinear models can also be tested using conventional likelihood-based hypothesis tests using a more general model, rather than the nonnested tests considered in Kobayashi (1994). Godfrey and Wickens (1981) proposed the Lagrange multiplier (LM) test for linear and loglinear models, noting that both models are special cases of the Box–Cox regression model, where the dependent and explanatory variables are expressed in a parametric family of power transformations. Davidson and MacKinnon (1984, 1985) proposed a variant of the LM test, whose size is more reliable in finite samples (see also Pesaran and Pesaran, 1995). The purpose of the paper is to extend the analytical results of Kobayashi (1994), so that the explanatory variables are also transformed in the loglinear model, and the dependent variable. In addition, the LM test and the Andrews test, both of which are based on the Box–Cox regression model, are included in the comparison. It is shown that the LM test has the highest asymptotic power of the compared tests, including the nonnested Cox test, which was found to have the highest power in Kobayashi (1994). This finding is novel and interesting, because the LM test is not a nonnested test. The PE, B&M, and Andrews tests, all of which are based on variable addition, are shown to have asymptotically equivalent powers and to have lower powers than the Cox test.

2. MODEL SPECIFICATIONS

The two models to be tested are the loglinear model

$$\log y_t = \alpha_1 + \beta \log X_t + v_t, \qquad v_t \sim \text{NID}(0, \sigma^2)$$

and the linear model

$$y_t = \gamma_1 + \delta_1 X_t + u_t, \qquad u_t \sim \text{NID}(0, \omega^2),$$

where $t = 1, 2, \ldots, n$. To make the two models converge, it is assumed that σ and ω converge to 0 as the sample size n increases, namely, $\sigma = O(n^{-1/2})$ and $\omega = O(n^{-1/2})$, and that the standard deviation of the original variable X_t relative to its mean, which is assumed to be positive, is small, namely, $(X_t - \bar{X})/\bar{X} = O(\sigma)$. Denoting the new regressor as $x_t = (X_t - \bar{X})/\bar{X}$, and $\alpha = \alpha_1 + \beta \log \bar{X}$, $\delta = \delta_1 \bar{X}$,

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

and $\gamma = \gamma_1 + \delta_1 \bar{X}$, the loglinear and linear models to be considered in this paper are formally defined in Assumption 1.

Assumption 1. The loglinear model is defined as

$$\log y_t = \alpha + \beta q_t + v_t, \tag{1}$$

where $v_t \sim \text{NID}(0,\sigma^2)$, $\sum_{t=1}^{n} x_t = 0$, $q_t = \log(1 + x_t)$, $x_t = O(\sigma)$, and $\sigma = O(n^{-1/2})$. The linear model is defined as

$$y_t = \gamma + \delta x_t + u_t, \tag{2}$$

where $\gamma > 0$, $u_t \sim \text{NID}(0,\omega^2)$, and $\omega = O(n^{-1/2})$.

It should be noted here that, from the assumptions $\gamma > 0$, $x = O(\sigma)$, and $u = O_p(\sigma)$, it follows that $y_t > 0$ when n is sufficiently large and hence $\log y_t$ is well defined.

The assumption of the small variation of the regressor variable x_t and the small disturbance variances σ^2 and ω^2 is necessary to derive well-defined asymptotic power functions; under this assumption, the loglinear model converges to the linear model as the sample size n increases. If this assumption is not satisfied, the models are distinct and the powers of the tests converge to 1; the models are clearly distinguished by the nonnormality of the errors if the variances are large and by the nonlinear effect of x_t if the variation of x_t is large.

The asymptotic analysis based upon Assumption 1 is called small-σ expansion. As an example of its usage in a related context, see Bickel and Doksum (1981), who examined estimation issues associated with the Box–Cox model, a flexible nonlinear family of power transformations.

In considering the LM tests in the framework of the Box–Cox regression model, an additional assumption is necessary on the intercept terms. The details are given in Section 7.

3. THREE VARIABLE ADDITION TESTS

The PE, B&M, and Andrews tests are considered under the nomenclature of variable addition tests. In these procedures, an artificial variable is attached to the null model, and then the significance of the additional variable is tested by using an exact or asymptotic t-test. If the additional variable is not significant, the null model is accepted. For a heuristic derivation of these tests, see Godfrey, McAleer, and McKenzie (1988).

The PE test for the loglinear model is defined by the t-ratio for testing $\pi = 0$ in the linear regression

$$\log y = \alpha + \beta \log(1 + x) + \pi(\exp(\hat{z}) - \hat{\gamma} - \hat{\delta}\hat{x}) + v,$$

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

where $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$, and $\hat{\delta}$ are the ordinary least squares (OLS) estimators in (1) and (2) and $\hat{z} = \hat{\alpha} + \hat{\beta}\log(1 + x)$ is the fitted value in (1). In this case, the PE test coincides asymptotically with the t-test for the inclusion of q^2 in addition to a constant and q, because a second-order expansion yields

$$\exp(\hat{z}) = \exp(\hat{\alpha} + \hat{\beta}\log(1 + x))$$
$$= \exp(\hat{\alpha})[1 + \hat{\beta}\log(1 + x) + \hat{\beta}^2\{\log(1 + x)\}^2/2] + O_p(\sigma^3).$$

Thus, the PE test statistic can be expressed as

$$\text{PE} = \sum_{t=1}^{n} \lambda_t \log y_t/(\hat{\sigma}^2 \sum \lambda_t^2)^{1/2}, \tag{3}$$

where $\hat{\sigma}^2$ is the ML estimate of σ^2, $w_t = \log(1 + x_t) - \sum \log(1 + x_t)/n$, and $\lambda_t = w_t^2 - w_t \sum w_t^3/\sum w_t^2 - \sum w_t^2/n$, which is orthogonal to 1 and w.

The next proposition is obtained after some algebra given in the Appendix.

PROPOSITION 1. *Under Assumption 1, the distribution of the PE test statistic under the alternative is asymptotic normal with variance 1 and mean*

$$n^{1/2}\delta(1 - \delta/\gamma)H^{1/2}/(2\omega),$$

where

$$m_{xxxx} = \sum x^4/n, \qquad m_{xxx} = \sum x^3/n, \qquad m_{xx} = \sum x^2/n,$$

$$H = m_{xxxx} - m_{xx}^2 - m_{xxx}^2/m_{xx}.$$

The B&M test (Bera and McAleer, 1989) is the t-test for $\pi = 0$ in

$$\log y_t = \alpha + \beta\log(1 + x_t) + \pi\hat{\eta}_t + \varepsilon_t,$$

where ε_t is a disturbance and $\hat{\eta}_t$ is the OLS residual obtained by regressing $\exp(\hat{z})$ on 1 and x_t. The Andrews test (Andrews, 1971) is the exact t-test for $\pi = 0$ in

$$\log y_t = \alpha + \beta\log(1 + x_t) + \pi(\hat{z}_t^2 - \hat{\beta}q_t^2) + \varepsilon_t.$$

The B&M and Andrews test statistics follow exact t-distributions under the null hypothesis of a loglinear model.

Substituting $\hat{z}_t = \hat{\alpha} + \hat{\beta}q_t$ in the preceding equations and expanding them using a Taylor series, it can easily be seen that these three test statistics are asymptotically expressed as (3), which is the t-statistic for the presence of w^2 in addition to 1 and $\log(1 + x)$. It then follows that the power functions of these tests are asymptotically identical with that of the test given in Proposition 1.

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

4. THE COX TEST

The Cox test is an extension of the log-likelihood ratio procedure to test nonnested hypotheses. Denoting the log-likelihood ratio by T, the Cox test statistic is defined as $(T/n - \hat{\mu}_N)/\sigma_N$, where the expected value of T/n is denoted by μ_N, the ML estimator of μ_N by $\hat{\mu}_N$, and the variance of $T/n - \hat{\mu}_N$ by σ_N^2, evaluated under the null hypothesis. The log-likelihood ratio of T for (1) and (2) is

$$T = (n/2)\log(\hat{\omega}^2) - (n/2)\log(\hat{\sigma}^2) - \sum \log y,$$

because the log-likelihood for (1) is

$$(-n/2)\log(2\pi\sigma^2) - \sum (\log y - \alpha - \beta w)^2/(2\sigma^2) - \sum \log y.$$

After some lengthy algebra given in the Appendix, the expected value and variance of T, and hence the Cox test statistic for (1) against (2), can be derived in the following proposition.

PROPOSITION 2. *The Cox test statistic is defined as* $(T/n - \mu_N)/\sigma_N$, *where*

$$\mu_N = \tfrac{1}{8}\beta^2(1-\beta)^2 H/\sigma^2 + \beta^2 m_{xx} + \tfrac{3}{4}\sigma^2,$$

$$\sigma_N^2 = [\tfrac{1}{4}\beta^2(1-\beta)^2 H/\sigma^2 + 2\beta^2 m_{xx} + \tfrac{3}{2}\sigma^2]/n,$$

and the ML estimators of β *and* σ^2 *are substituted in* μ_N *and* σ_N^2.

This version of the Cox test statistic is different from the original form of the Cox test statistic given, for example, in Pesaran and Pesaran (1993). However, these two statistics coincide asymptotically when the null and alternative models converge, as given in Assumption 1, so that this form is sufficient to show the local properties of the Cox test statistic.

The distribution of T under the alternative hypothesis and then the power of the Cox test statistic are given as follows, with the algebraic details given in the Appendix.

PROPOSITION 3. *Under the alternative hypothesis of the linear model, the distribution of the Cox test statistic is asymptotic normal with mean*

$$-n^{1/2}[\tfrac{1}{4}\delta^2(1-\delta/\gamma)^2 H/\omega^2 + 2\delta^2 m_{xx}/\gamma^2 + \tfrac{3}{2}\omega^2/\gamma^2]^{1/2}$$

and variance 1.

5. THE LAGRANGE MULTIPLIER TEST

The LM tests for the linear and loglinear models suggested by Godfrey and Wickens (1981) are the tests of $\lambda = 0$ and $\lambda = 1$ in the Box–Cox regression model

$$y_t(\lambda) = \alpha + \beta x_t(\lambda) + v_t. \qquad \textbf{(4)}$$

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

The variables $y(\lambda)$ and $x(\lambda)$ are defined by

$$y(\lambda) = \begin{cases} (y^\lambda - 1)/\lambda & (\lambda \neq 0) \\ \log y & (\lambda = 0) \end{cases} \quad \text{and} \quad x(\lambda) = \begin{cases} (x^\lambda - 1)/\lambda & (\lambda \neq 0) \\ \log x & (\lambda = 0) \end{cases},$$

which are continuous in $y > 0$ and $x > 0$, because $(y^\lambda - 1)/\lambda$ and $(x^\lambda - 1)/\lambda$ converge to $\log y$ and $\log x$ if $y - 1$ and $x - 1$ are sufficiently small.

The definition of the linear and loglinear models used in this section is slightly different and is defined in the next assumption.

Assumption 2. The loglinear and linear models are defined by

$$\log y = \alpha + \beta \log(1 + x) + v, \tag{5}$$

$$y = 1 + \phi + \delta x + u, \tag{6}$$

where the notation is the same as that employed in (1) and (2), except that $\alpha = O(n^{-1/2})$ and $\phi = O(n^{-1/2})$.

The assumption regarding the intercept terms in Assumption 2 agrees with the original specification of Box and Cox (1964) that the variables should be standardized so that the mean of the logarithms of the variables is zero. This standardization ensures that $y(\lambda)$ and $x(\lambda)$ converge to $\log y$ and $\log(1 + x)$ rapidly as λ converges to 0.

The LM test statistic for this problem is the first-order derivative of the log-likelihood evaluated under the null hypothesis $\lambda = 0$, namely,

$$(1/n) \sum_{t=1}^{n} \{\log y_t - (\log y_t - \hat{\alpha} - \hat{\beta} q_t)[(\log y_t)^2 - \hat{\beta}(\log(1 + x_t))^2]/(2\hat{\sigma}^2)\},$$

divided by its estimated standard deviation, where the reference distribution is the standardized normal distribution.

After some algebra given in the Appendix, the next proposition is obtained.

PROPOSITION 4. *Under the alternative hypothesis, the* LM *test statistic follows the asymptotic normal distribution with mean*

$$n^{1/2}[\tfrac{3}{2}\omega^2 + \delta^2(1 - \delta)^2 H/(4\omega^2) + 2(\phi^2 - \delta^2 m_{xx})]^{1/2}$$

and variance 1.

6. POWER COMPARISONS

An examination of Propositions 1, 3, and 4 enables a comparison of powers through the means (or noncentrality parameters) of the tests under the alternative hypothesis. The means of the test statistics under the alternative hypothesis are summarized as follows:

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

(i) PE, B&M, and Andrews tests:

$$n^{1/2}\delta(1-\delta/\gamma)H^{1/2}/(2\omega)$$

(ii) Cox test:

$$-n^{1/2}[\tfrac{3}{2}\omega^2/\gamma^2+\delta^2(1-\delta/\gamma)^2H/(4\omega^2)+2\delta^2 m_{xx}/\gamma^2]^{1/2}$$

(iii) LM test:

$$n^{1/2}[\tfrac{3}{2}\omega^2+\delta^2(1-d)^2H/(4\omega^2)+2(\phi^2+\delta^2 m_{xx})]^{1/2}.$$

It can be seen that the Cox test has higher power than the PE, B&M, and Andrews tests, because

$$\begin{aligned}&[\tfrac{3}{2}\omega^2/\gamma^2+\delta^2(1-\delta/\gamma)^2H/(4\omega^2)+2\delta^2 m_{xx}/\gamma^2]^{1/2}\\&\quad\le[\delta^2(1-\delta/\gamma)^2H/(4\omega^2)]^{1/2}\\&\quad=|\delta(1-\delta/\gamma)|H^{1/2}/(2\omega).\end{aligned}$$

Under the assumption $\gamma=1+\phi=1+O(n^{-1/2})$ in Assumption 2, it follows that the LM test has higher power than the Cox test, because

$$\begin{aligned}&\tfrac{3}{2}\omega^2+\delta^2(1-\delta)^2H/(4\omega^2)+2(\delta^2 m_{xx}+\phi^2)\\&\quad\le\tfrac{3}{2}\omega^2+\delta^2(1-\delta)^2H/(4\omega^2)+2\delta^2 m_{xx}.\end{aligned}$$

This leads to the main results of the paper.

PROPOSITION 5. *The* PE *test, the* B&M *test, and the Andrews test have asymptotically equivalent power and have lower power than the Cox test. If* $\gamma=1+O(n^{-1/2})$, *the* LM *test has the highest power of the compared tests.*

An intuitive explanation of why the Cox test is more powerful than the variable addition tests is given in the next proposition.

PROPOSITION 6. *Assume that y is expanded as*

$$y=\exp(\alpha)+\exp(\alpha)\beta x+\exp(\alpha)v-\tfrac{1}{2}\exp(\alpha)\beta(1-\beta)x^2+O_p(n^{-3/2})$$

under the null, and that

$$\log y=\log\gamma+(\delta/\gamma)w+u/\gamma+\tfrac{1}{2}(\delta/\gamma)(1-\delta/\gamma)w^2+O_p(n^{-3/2})$$

under the alternative. Then the test statistic $(\tilde{T}/n-\tilde{\mu}_0)/\tilde{\sigma}_0$ *has the same power function as that of the variable addition tests, where* $\tilde{T}$ *is defined by* $\log(\hat{w}^2)-\log(\hat{\sigma}^2)$ *and* $\tilde{\mu}_0$ *and* $\tilde{\sigma}_0^2$ *are defined by the asymptotic mean and variance of* $\tilde{T}$ *under the null.*

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

The last term on the right-hand side of the expansion of y in Proposition 6 expresses the nonlinear effect of x on y. On the right-hand side of the correct expansion of y, namely,

$$y = \exp(\alpha) + \exp(\alpha)\beta x + \exp(\alpha)v - \tfrac{1}{2}\exp(\alpha)\beta(1-\beta)x^2 + \exp(\alpha)\beta xv + \exp(\alpha)v^2/2 + O_p(n^{-3/2}),$$

the fifth, sixth, and seventh terms correspond to heteroskedasticity, nonnormality, and order of magnitude $n^{-3/2}$, respectively. These terms are excluded from the expression for y in Proposition 5. Similar terms are also excluded from the correct expansion

$$\log y = \log \gamma + (\delta/\gamma)w + u/\gamma + \tfrac{1}{2}(\delta/\gamma)(1-\delta/\gamma)w^2 - (\delta/\gamma^2)xu - \tfrac{1}{2}u^2/\gamma^2 + O_p(n^{-3/2})$$

in Proposition 6. Also note that $\tilde{T}$ differs from T by the Jacobian term.

This proposition means that the three variable addition tests have lower power than the Cox test because the former tests use only the nonlinear effect. The latter test makes use of heteroskedasticity, nonnormality, the term of $O_p(n^{-3/2})$, and the Jacobian of the transformation, in addition to the nonlinear effect, in distinguishing between the linear and loglinear models. The proof is almost the same as that of Propositions 2 and 3 and hence is omitted.

The result that the LM test has higher power than the Cox test might appear, at first glance, to be somewhat unusual. Note that the difference in the powers of these tests depends upon the intercept parameter ϕ, which can take arbitrary values by changing the unit of measurement of the variables. The LM test seems to detect the deviation of $(y^\lambda - 1)/\lambda$ from $\log y$ for small positive λ under the alternative of the linear model; for fixed λ, this deviation increases as the absolute value of $\phi + \delta x$ increases, because y is expressed as $y = 1 + \phi + \delta x + u$, where $u = O_p(n^{-1/2})$ and $x = O(n^{-1/2})$.

7. DISCUSSION

Several extensions of the main results are possible. One is the extension to time series analysis, especially to nonstationary time series processes. The recent flourishing of dynamic models in applied econometrics has made the choice of the transformation of economic variables more important; for example, the heteroskedasticity in the loglinear model may be mistaken as the effect of a random walk process. In particular, the choice between tests of unit roots conducted in levels or logarithms (i.e., where the dependent variable is in level differences or log differences), cointegrating relations among variables estimated in levels or logarithms, and the associated error correction models expressed in levels or logarithms has become an issue of considerable importance in the econometric analysis of time series data. Discussions of these issues can be found in Granger and Hallman (1991) and Franses and McAleer (1998).

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

REFERENCES

Andrews, D.F. (1971) A note on the selection of data transformations. *Biometrika* 58, 249–254.
Bera, A.K. & M. McAleer (1983) Further results on testing linear and log-linear regression models. Paper presented to the SSRC Econometric Study Group Conference on Model Specification and Testing, Warwick.
Bera, A.K. & M. McAleer (1989) Nested and non-nested procedures for testing linear and log-linear regression models. *Sankhya B* 51, 212–224.
Bickel, P.J. & K.A. Doksum (1981) An analysis of transformations revisited. *Journal of the American Statistical Association* 76, 296–310.
Box, G.E.P. & D.R. Cox (1964) An analysis of transformations. *Journal of the Royal Statistical Society B* 26, 211–243.
Cox, D.R. (1961) Tests of separate families of hypotheses. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 105–123. Berkeley: University of California Press.
Cox, D.R. (1962) Further results on tests of separate families of hypotheses. *Journal of the Royal Statistical Society B* 24, 406–424.
Davidson, R. & J.G. MacKinnon (1984) Model specification tests based on artificial linear regressions. *International Economic Review* 25, 485–502.
Davidson, R. & J.G. MacKinnon (1985) Testing linear and loglinear regressions against Box–Cox alternatives. *Canadian Journal of Economics* 18, 499–517.
Franses, P.H. & M. McAleer (1998) Testing for unit roots and nonlinear transformations. *Journal of Time Series Analysis* 19, 147–164.
Godfrey, L.G., M. McAleer, & C.R. McKenzie (1988) Variable addition and Lagrange multiplier tests for linear and logarithmic regression models. *Review of Economics and Statistics* 70, 492–503.
Godfrey, L.G. & M.R. Wickens (1981) Testing linear and log-linear regressions for functional form. *Review of Economic Studies* 48, 487–496.
Gourieroux, C. & A. Monfort (1994) Testing non-nested hypotheses. In R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics*, vol. 4. Amsterdam: North-Holland.
Granger, C.W.J. & J. Hallman (1991) Nonlinear transformations of integrated time series. *Journal of Time Series Analysis* 12, 207–224.
Kobayashi, M. (1994) Power of tests for nonlinear transformation in regression analysis. *Econometric Theory* 10, 357–371.
MacKinnon, J.G., H. White, & R. Davidson (1983) Tests for model specification in the presence of alternative hypotheses: Some further results. *Journal of Econometrics* 21, 53–70.
Pesaran, B. & M.H. Pesaran (1995) A non-nested test of level differenced versus log-differenced stationary models. *Econometric Review* 14, 213–227.
Pesaran, M.H. & B. Pesaran (1993) A simulation approach to the problem of computing Cox's statistic for testing non-nested models. *Journal of Econometrics* 57, 377–392.

APPENDIX: ALGEBRAIC DETAILS

For ease of computation and generalization, the models are defined in a more general form in Assumption A.

Assumption A. The null hypothesis is defined as

$$F(y_t) = \alpha + \beta w_t + v_t, \qquad \textbf{(A.1)}$$

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

where $v_t \sim \text{NID}(0,\sigma^2)$, $\sum w_t = 0$, $w = O(\sigma)$, $\sigma = O(n^{-1/2})$ as $n \to \infty$, and $F(x)$ is a continuously differentiable function with $F'(x) > 0$ and $F''(x) \neq 0$ in the parameter space of γ. The alternative hypothesis is defined as

$$y_t = \gamma + \delta x_t + u_t, \tag{A.2}$$

where $u_t \sim \text{NID}(0,\omega^2)$, $\sum x_t = 0$, $x = O(\sigma)$, $\omega = O(n^{-1/2})$ as $n \to \infty$. It is also assumed that $w_t = h(x_t) - \sum h(x_t)/n$ for some continuously differentiable function $h(x)$ such that $h(0)$ and $h'(0) = 1$.

Under this assumption, the inverse of $F(\ )$ is well defined, because $F'(\ )$ is nonzero. It can be assumed that $F'(\ ) > 0$ without loss of generality, because the models remain essentially unchanged by interchanging $F(\ )$ and $-F(\ )$.

The loglinear and linear models are defined by specifying $F(y) = \log y$, $\gamma > 0$, and $w = \log(1 + x) - \sum \log(1 + x)/n$ in (A.1) and (A.2).

Model (A.1) has several advantages over model (1) in that the results can be made quite general: (1) it includes the case where the same regressor is used in both models, by setting $x = w$ in (A.1) and (A.2); (ii) it includes the nonnested case where the dependent variables are identical (i.e., $F(y_t) = y_t$) but the regressors are not (i.e., $w_t \neq x_t$); (iii) the distribution of the log-likelihood under the alternative can be easily obtained from the null distribution by replacing $F(\ )$ with the inverse function $F^{-1}(\ )$ and interchanging the roles of the models; (iv) the case where the linear model is the null can be dealt with by treating $F(y)$ as the original dependent variable in (2). Although the linear model is the alternative in this paper, the results also hold when the linear model is the null and the loglinear model is the alternative.

A.1. The Variable Addition Tests. The power of the PE test is derived as follows. By using the expansion

$$F(y) = F(\gamma + \delta x + u) = F(\gamma) + F_1(\delta x + u) + F_2(\delta x + u)^2/2 + O_p(\sigma^3),$$

where the ith derivative of $F(\ )$ evaluated at γ is denoted by F_i $(i = 1,2)$, it follows that

$$\text{PE} = \sum \lambda[F_1(\delta\xi + u) + F_2\lambda(\delta x + u)^2/2]/(\hat{\sigma}^2 \sum \lambda^2)^{1/2} + O_p(\sigma)$$

$$= n^{1/2}[F_1\delta \sum \lambda\xi/n + F_1 \sum \lambda u/n + \tfrac{1}{2}F_2\delta^2 \sum \lambda x^2/n]/(F_1^2\omega^2 \sum \lambda^2/n)^{1/2} + O_p(\sigma),$$

where $\lambda_t = w_t^2 - w_t \sum w_t^3/\sum w_t^2 - \sum w_t^2/n$, $\xi = x - w = -h_2x^2 + O(\sigma^3)$, and $h_2 = h''(0)$. Note that $\hat{\sigma}^2$ converges to $\omega^2F_1^2$ under the alternative, that $F_1\delta \sum \lambda\xi/n$ and $F_2\delta^2 \sum \lambda x^2/n$ are $O(\sigma^4)$, and that $F_1 \sum \lambda u/n$ is $O_p(\sigma^4)$, which is normally distributed with mean 0 and variance $F_1^2\omega^2 \sum \lambda^2/n^2 = n^{-1}F_1^2\omega^2(m_{wwww} - m_{ww}^2 - m_{www}^2/m_{ww})$, where $m_{ww} = \sum w^2/n$, $m_{www} = \sum w^3/n$, and $m_{wwww} = \sum w^4/n$. Note that the difference between $m_{wwww} - m_{ww}^2 - m_{www}^2/m_{ww}$ and $H = m_{xxxx} - m_{xx}^2 - m_{xxx}^2/m_{xx}$ is smaller than $O(\sigma^4)$ and is negligible for the analysis. The expansions derived previously lead to the following proposition.

PROPOSITION A.1. *The asymptotic distribution of the* PE *test statistic under the alternative is normal with mean*

$$n^{1/2}\tfrac{1}{2}H[\delta^2(F_2/F_1) - \delta h_2]/(\omega H^{1/2}) \tag{A.3}$$

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

and variance 1, *where* $m_{xxxx} = \sum x^4/n$, $m_{xxx} = \sum x^3/n$, $m_{xx} = \sum x^2/n$, $H = m_{xxxx} - m_{xx}^2 - m_{xxx}^2/m_{xx}$, *and* $h_2 = h''(0)$.

Noting that $h_2 = -1$ and $F_2/F_1 = -1/\gamma$ for the loglinear model, Proposition 1 can be easily obtained from Proposition A.1.

In the case where $x = w$, that is, when the same regressor is used in each of the models, the mean of the test statistic under the alternative is $n^{1/2}(F_2/F_1)\delta^2 H^{1/2}/(2\omega)$, which was given in Kobayashi (1994).

A.2. The Cox Test. The asymptotic mean and variance of the log-likelihood under the null are obtained as follows. From the orthogonality between $y - \hat{\gamma} - \hat{\delta}x$ and $(\hat{\gamma} - \gamma_0) - (\hat{\delta} - \delta_0)x$, it follows that

$$\hat{\omega}^2 = \sum (y - \hat{\gamma} - \hat{\delta}x)^2/n$$

$$= \sum (y - \gamma_0 - \delta_0 x)^2/n - (\hat{\gamma} - \gamma_0)^2 - (\hat{\delta} - \delta_0)^2 \sum x^2/n, \tag{A.4}$$

where $\gamma_0 = G(\alpha)$ and $\delta_0 = G_1\beta$, to which the ML estimators $\hat{\gamma}$ and $\hat{\delta}$ converge.

Under the null, the dependent variable is expanded as

$$y = G(\alpha) + G_1 e + G_2 e^2/2 + G_3 e^3/6 + \cdots +$$

$$= G(\alpha) + G_1\beta x - G_1\beta\xi + G_1 v + G_2 e^2/2 + G_3 e^3/6 + O_p(n^{-2}) \tag{A.5}$$

from (A.1) and (A.2), where $e = \beta w + v = \beta x - \beta\xi + v = O(\sigma)$, $x - w = \xi = x^2/2 - \frac{1}{2}\sum x^2/n + O(\sigma^3) = O(\sigma^2)$, and $G_i = G^{(i)}(\alpha)$ $(i = 1,2,3)$.

It follows from (A.5) that

$$\hat{\gamma} - \gamma_0 = \sum y/n - \gamma_0$$

$$= G_1 m_v + G_2 m_{ee}/2 + O_p(\sigma^3)$$

$$= G_1(m_v + g_2 m_{ee}/2) + O_p(\sigma^3), \tag{A.6}$$

$$\hat{\delta} - \delta_0 = \sum xy \Big/ \sum x^2 - \delta_0$$

$$= (G_1 m_{vx} - G_1\beta m_{x\xi} + G_2 m_{xee}/2)/m_{xx} + O_p(\sigma^3)$$

$$= G_1(m_{vx} - \beta m_{x\xi} + g_2 m_{xee}/2)/m_{xx} + O_p(\sigma^2). \tag{A.7}$$

It also follows from (A.5) that

$$\sum (y - \gamma_0 - \delta_0 x)^2/n = G_1^2[\beta^2 m_{\xi\xi} - 2\beta m_{\xi v} + m_{vv}$$

$$+ g_2^2 m_{eeee}/4 - g_2\beta m_{\xi ee} + g_2 m_{eev} + g_3 m_{veee}/3] + O_p(\sigma^5), \tag{A.8}$$

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

where $g_2 = G_2/G_1$ and $g_3 = G_3/G_1$. Then, from (A.4) and (A.6)–(A.8), the estimator of the variance of (A.2) is expressed as

$$\begin{aligned}\hat{\omega}^2 &= G_1^2[m_{vv} - 2\beta m_{\xi v} + \beta^2 m_{\xi\xi} \\ &\quad + g_2^2 m_{eeee}/4 - g_2\beta m_{\xi ee} + g_2 m_{eev} + g_3 m_{veee}/3 \\ &\quad - (m_v^2 + g_2 m_v m_{ee} + g_2^2 m_{ee}^2/4) \\ &\quad - (m_{vx}^2 + \beta^2 m_{x\xi}^2 + g_2^2 m_{xee}^2/4 - 2m_{vx}\beta m_{x\xi} + g_2 m_{xee} m_{vx} - g_2\beta m_{x\xi} m_{xee})/m_{xx}] \\ &= G_1^2[(m_{vv} - m_v^2 - m_{vx}^2/m_{xx}) - 2\beta(m_{\xi v} - \beta m_{vx} m_{x\xi}/m_{xx}) \\ &\quad + \beta^2(m_{\xi\xi} - m_{x\xi}^2/m_{xx}) + g_2^2 m_{eeee}/4 - g_2\beta m_{\xi ee} + g_2 m_{eev} + g_3 m_{veee}/3 \\ &\quad - (g_2 m_v m_{ee} + g_2^2 m_{ee}^2/4) - (g_2 m_{xee} m_{vx} - g_2\beta m_{x\xi} m_{xee} + g_2^2 m_{xee}^2/4)/m_{xx}].\end{aligned} \tag{A.9}$$

The ML estimator of the variance of (A.1) is expressed as

$$\hat{\sigma}^2 = m_{vv} - m_v^2 - m_{vw}^2/m_{ww} = m_{vv} - m_v^2 - m_{vx}^2/m_{xx} + O(\sigma^5). \tag{A.10}$$

The last term in the log-likelihood ratio is expressed as

$$\begin{aligned}(1/n)\sum \log(F'(y)) &= -(1/n)\sum \log(G'(F(y))) \\ &= -[\log G_1 + g_2 m_v + \tfrac{1}{2}(g_3 - g_2^2)(\beta_2^2 m_{ww} + m_{vv})] + O_p(n^{-3/2}),\end{aligned} \tag{A.11}$$

because

$$-\log(G'(F(y)) = \log F'(y)$$

and

$$\log G'(y) = \log G_1 + g_2(\beta w + v) + g_3(\beta w + v)^2/2 - \tfrac{1}{2}g_2^2(\beta w + v)^2 + O_p(n^{-3/2}).$$

Then, from (A.9)–(A.11), it follows that

$$\begin{aligned}T/n &= \tfrac{1}{2}\log(\hat{\omega}^2/\hat{\sigma}^2) + \sum \log F'(y)/n \\ &= \tfrac{1}{2}[-2\beta(m_{v\xi} - m_{x\xi} m_{vx}/m_{xx}) + \beta^2(m_{\xi\xi} - m_{x\xi}^2/m_{xx}) \\ &\quad + g_2^2 m_{eeee}/4 - g_2\beta m_{\xi ee} + g_2 m_{eev} + g_3 m_{veee}/3 - (g_2 m_v m_{ee} + g_2^2 m_{ee}^2/4) \\ &\quad - (g_2 m_{xee} m_{vx} - g_2\beta m_{x\xi} m_{xee} + g_2^2 m_{xee}^2/4)/m_{xx}]/\sigma^2 \\ &\quad - g_2 m_v - \tfrac{1}{2}(g_3 - g_2^2)m_{ee} + O_p(n^{-3/2}).\end{aligned} \tag{A.12}$$

Equation (A.12) may be rewritten as

$$T/n = A + B + O_p(\sigma^3),$$

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

where the nonstochastic term A is

$$A = \tfrac{1}{2}[\beta^2(m_{\xi\xi} - m_{x\xi}^2/m_{xx}) + g_2^2 m_{eeee}/4 - g_2\beta m_{\xi ee} + g_3 m_{veee}/3 - g_2^2 m_{ee}^2/4 - (-g_2\beta m_{x\xi} m_{xee} + g_2^2 m_{xee}^2/4)/m_{xx}]/\sigma^2 - \tfrac{1}{2}(g_3 - g_2^2)m_{ee}$$

and the stochastic term B is

$$B = \tfrac{1}{2}[-2\beta(m_{v\xi} - m_{x\xi} m_{vx}/m_{xx}) + g_2 m_{eev} - g_2 m_v m_{ee} - g_2 m_{xee} m_{vx}/m_{xx}]/\sigma^2 - g_2 m_v.$$

Substitution yields

$$\begin{aligned}
m_{ee} &= \beta^2 m_{xx} + \sigma^2 + O(\sigma^3),\\
m_{eev} &= \beta^2 m_{wwv} + 2\beta m_{wvv} + m_{vvv} = \beta^2 m_{xxv} + 2\beta_{xvv} + m_{vvv} = O_p(\sigma^3),\\
m_{\xi ee} &= \beta^2 m_{\xi xx} + \sigma^2 m_\xi = \beta^2 m_{\xi xx} + O_p(\sigma^5),\\
m_{xee} &= \beta^2 m_{xww} + 2\beta m_{xwv} + m_{xvv} = \beta^2 m_{xxx} + O(\sigma^4),\\
m_{xxee} &= \beta^2 m_{xxww} + m_{xxvv} + 2\beta m_{xxwv} = \beta^2 m_{xxxx} + \sigma^2 m_{xx} + O_p(\sigma^5),\\
m_{veee} &= \beta^3 m_{vxxx} + 3\beta^2 m_{vvxx} + 3\beta m_{vvvx} + m_{vvvv} = 3(\beta^2\sigma^2 m_{xx} + \sigma^4) + O_p(\sigma^5),\\
m_{eeee} &= \beta^4 m_{xxxx} + 6\beta^2 m_{xx}\sigma^2 + 3\sigma^4 + O_p(\sigma^5).
\end{aligned}$$

It then follows that

$$\begin{aligned}
A &= \tfrac{1}{2}[\beta^2(m_{\xi\xi} - m_{x\xi}^2/m_{xx}) + g_2^2(\beta^4 m_{xxxx} + 6\beta^2 m_{xx}\sigma^2 + 3\sigma^4)/4 - g_2\beta^3 m_{\xi xx}\\
&\quad + g_3(\beta^2\sigma^2 m_{xx} + \sigma^4) - g_2^2(\beta^2 m_{xx} + \sigma^2)^2/4\\
&\quad - (-g_2\beta m_{x\xi}(\beta^2 m_{xxx}) + g_2^2(\beta^2 m_{xxx})^2/4)/m_{xx}]/\sigma^2\\
&\quad - \tfrac{1}{2}(g_3 - g_2^2)(\beta^2 m_{xx} + \sigma^2)\\
&= [\tfrac{1}{8}g_2^2\beta^4(m_{xxxx} - m_{xx}^2 - m_{xxx}^2/m_{xx}) - \tfrac{1}{2}g_2\beta^3(m_{\xi xx} - m_{x\xi} m_{xxx}/m_{xx})\\
&\quad + \beta^2(m_{\xi\xi} - m_{x\xi}^2/m_{xx}) + \beta^2\sigma^2 m_{xx}(\tfrac{1}{2}g_3 + \tfrac{3}{4}g_2^2 - \tfrac{1}{4}g_2^2 - \tfrac{1}{2}(g_3 - g_2^2))\\
&\quad + \sigma^4(g_2^2\tfrac{3}{8} - g_2^2\tfrac{1}{8} + \tfrac{1}{2}g_3 - \tfrac{1}{2}(g_3 - g_2^2))]/\sigma^2\\
&= [g_2^2\beta^4 H/8 - g_2\beta^3 J/2 + \beta^2 K/2]/\sigma^2 + g_2^2\beta^2 m_{xx} + \tfrac{3}{4}g_2^2\sigma^2,
\end{aligned}$$

where

$$H = m_{xxxx} - m_{xx}^2 - m_{xxx}^2/m_{xx}, \qquad J = m_{\xi xx} - m_{x\xi} m_{xxx}/m_{xx}, \qquad K = m_{\xi\xi} - m_{x\xi}^2/m_{xx}$$

and

$$B = \tfrac{1}{2}[-2\beta(m_{v\xi} - m_{x\xi} m_{vx}/m_{xx}) + g_2(\beta^2 m_{xxv} + 2\beta m_{xvv} + m_{vvv}) - g_2 m_v(\beta^2 m_{xx} + \sigma^2) - g_2\beta^2 m_{xxx} m_{vx}/m_{xx}]/\sigma^2 - g_2 m_v. \qquad \textbf{(A.13)}$$

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

Equation (A.13) can be expressed as

$$B = (1/n)g_2 \sum (c_1 v^3 + c_2 v^2 + c_3 v)/\sigma^2,$$

where

$$c_1 = \tfrac{1}{2}, \qquad c_2 = \beta x, \qquad c_3 = \beta\tau/g_2 + \tfrac{1}{2}\beta^2\eta - \tfrac{3}{2}\sigma^2,$$

$$\eta = x^2 - m_{xx} - x m_{xxx}/m_{xx}, \qquad \tau = \xi - x m_{x\xi}/m_{xx}.$$

Noting that

$$\sum \tau = 0, \qquad \sum \tau x = 0, \qquad \sum \eta = 0, \qquad \sum \eta x = 0,$$

$$\sum c_3^2/n = \beta^4 H/4 - \beta^3 J/g_2 + \beta^2 K/g_2^2 + \tfrac{9}{4}\sigma^2,$$

the variance of B is given by

$$\mathrm{var}(B) = g_2^2[3\sigma^6/2 + 2\beta^2 m_{xx}\sigma^4 + (\beta^4 H/4 - \beta^3 J/g_2 + \beta^2 K/g_2^2)\sigma^2]/(n\sigma^4).$$

Substitution of $J = -h_2 H/2$ and $K = h_2^2 H/4$ leads to the following proposition.

PROPOSITION A.2. *The asymptotic distribution of T/n is normally distributed with mean μ_N and variance σ_N^2, where*

$$\mu_N = \tfrac{1}{8}\beta^2 H(g_2\beta + h_2)^2/\sigma^2 + g_2^2\beta^2 m_{xx} + \tfrac{3}{4}g_2^2\sigma^2,$$

$$\sigma_N^2 = 2\mu_N/n,$$

$$h_2 = h''(0).$$

Then the Cox test statistic for (A.1) *against* (A.2) *is* $(T/n - \mu_N)/\sigma_N$, *where the ML estimators of β and σ^2 are substituted in μ_N and σ_N^2.*

Proposition 2 for testing the linear and loglinear models is obtained upon substituting

$$g_2 = G''(\alpha)/G'(\alpha) = 1 \quad \text{and} \quad h_2 = -1.$$

The distribution of T/n under the alternative hypothesis of the linear model can be obtained easily from Proposition A.2. Substituting $F(y)$ for y in (A.1) and $y = G(F(y))$ for $F(y)$ in (A.2), and interchanging the regressors and parameters of the two models, namely, $(\alpha, \beta, \sigma, w, f_2)$ by $(\gamma, \delta, \omega, x, -f_2)$, leads to the following proposition.

PROPOSITION A.3. *The asymptotic distribution of T/n under the alternative hypothesis of the linear model* (A.2) *is normal with mean*

$$\mu_A = -\tfrac{1}{8}H\delta^2(f_2\delta - h_2)^2/\omega^2 + f_2^2\delta^2 m_{xx} + \tfrac{3}{4}f_2^2\omega^2$$

and variance $\sigma_A^2 = 2\mu_A/n$, where $f_2 = F''(\gamma)/F'(\gamma)$.

The power of the test statistic $(T/n - \mu_N)/\sigma_N$ is derived from Propositions A.2 and A.3, by noting that $f_2 = F''(\gamma)/F'(\gamma) = -F'(\gamma)G''(F(\gamma))/G'(F(\gamma))$, where $\hat{\alpha}$ converges to $F(\gamma)$, $\hat{\beta}$ to $\delta F'(\gamma) = \delta F_1$, σ^2 to $F_1^2\omega^2$, and hence $G''(\hat{\alpha})/G'(\hat{\alpha})$ to $G''(F(\gamma))/G'(F(\gamma)) =$

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press

$-f_2/F_1$, where F_1 is denoted by $F'(\gamma)$. Note that, under the alternative, the estimated μ_N and σ_N^2 converge to $-\mu_A$ and σ_A^2, respectively. Thus, the test statistic follows the asymptotic normal distribution with mean $-2\mu_A/\sigma_A = -n^{1/2}(2\mu_A)^{1/2}$ and variance 1. The power of the Cox test statistic is given in the following proposition.

PROPOSITION A.4. *Under the alternative hypothesis, the asymptotic distribution of the Cox test statistic* $(T/n - \mu_N)/\sigma_N$ *is normal with mean*

$$-n^{1/2}[\tfrac{1}{4}\delta^2 H(f_2\delta - h_2)^2/\omega^2 + 2f_2^2\delta^2 m_{xx} + \tfrac{3}{2}f_2^2\omega^2]^{1/2}$$

and variance 1.

The power of the Cox test for the loglinear model against the linear model, given in Proposition 3, is obtained by substituting $F'(\gamma) = 1/\gamma$, $f_2 = -1/\gamma$, and $h_2 = -1$.

A.3. The Lagrange Multiplier Test. The asymptotic distribution of the LM test statistic under the null hypothesis of the loglinear model is given as follows. The first-order derivative of the likelihood evaluated at the ML estimators under the null hypothesis is

$$S = \sum \log y/n - (1/n)\sum \hat{v}[(\log y)^2 - \hat{\beta}q^2]/(2\hat{\sigma}^2),$$

where $q = \log(1 + x)$. Reexpressing (5) as $\log y = \rho + \beta w + v$, $\rho = \alpha + \beta \sum q/n$, and $w = q - \sum q/n$, the first term is expressed as $\sum \log y/n = \rho + m_v$. Noting that the residual in the second term is expanded as $\hat{v} = v - m_v - w m_{vw}/m_{ww}$ and $(\log y)^2 = \rho^2 + \beta^2 w^2 + v^2 + 2\rho\beta w + 2\beta w v + 2\rho v$, after some algebra S may be rewritten as

$$\begin{aligned} S = \rho(1 - m_{vv}/\sigma^2) + m_v - \{&\beta(1-\beta)m_{vww} + m_{vvv} + 2\beta m_{wvv} \\ &- m_v[\beta(1-\beta)m_{ww} + m_{vv}] \\ &- (m_{vw}/m_{ww})\beta(1-\beta)m_{www}\}/(2\hat{\sigma}^2) + O_p(n^{-3/2}). \end{aligned}$$

Then the asymptotic distribution of S under the null hypothesis is normal with mean 0 and variance

$$(1/n)[\tfrac{3}{2}\sigma^2 + \tfrac{1}{4}\beta^2(1-\beta)^2 H/\sigma^2 + 2(\alpha^2 + \beta^2 m_{xx})].$$

Under the alternative hypothesis of the linear model, some algebraic manipulation leads to the expansion

$$\begin{aligned} S = \phi(1 - m_{uu}/\omega^2) + m_u, \\ - [\delta(1-\delta)m_{xuu} + m_{uuu} + 2\delta m_{xuu} \\ - m_u(\delta(1-\delta)m_{xx} + \omega^2) - (m_{xu}/m_{xx})\delta(1-\delta)m_{xxx}]/(2\omega^2) \\ + \tfrac{3}{2}\omega^2 + \delta^2(1-\delta)^2 H/(4\omega^2) + 2(\phi^2 + \delta^2 m_{xx}). \end{aligned}$$

Noting that, under the alternative, $\hat{\omega}^2$ converges to σ^2, $\hat{\beta}$ to δ, and $\hat{\alpha}$ to ϕ, it follows that the asymptotic distribution of the test statistic $S/(\mathrm{var}(S))^{1/2}$ is normal with mean

$$n^{1/2}[\tfrac{3}{2}\omega^2 + \delta^2(1-\delta)^2 H/(4\omega^2) + 2(\phi^2 + \delta^2 m_{xx})]^{1/2}$$

and variance 1, which is Proposition 4.

https://doi.org/10.1017/S0266466699151053 Published online by Cambridge University Press