Log Income versus Linear Income: An Application of the Encompassing Principle

Luigi Ermini and David F. Hendry *
University of Hawaii at Manoa,
and
Nuffield College, Oxford;

January 20, 1995

Abstract

An open question in empirical economics is whether models should be estimated by using the logarithm of the relevant economic variables or their actual, or 'linear', values. This paper applies the principle of encompassing to suggest specification and mis-specification tests of log versus linear individual equations fitted to $\mathsf{I}(1)$ data, and illustrates the analysis for US quarterly disposable income. The finite-sample properties of the encompassing tests are examined in a Monte Carlo experiment customized to the parameter values found in the empirical analysis.

Keywords: encompassing, log transformations
JEL Classification: 211

1 Introduction

An open question in empirical economics is whether models should be estimated by using the logarithm of the relevant economic variables or their actual, or 'linear', values. The use of log values is particularly favored by macro-econometricians, as a way to linearize exponentially growing trends apparently present in many variables of interest (GNP, consumption, etc.). The use of linear values is favored by theoretical economists, on the grounds that economic agents make decisions about 'objects' of choice defined in terms of actual and not log values.

A number of considerations bear on this important issue. First, if a model is characterized by both constant structural parameters and a constant error variance, then linearity in levels seems unlikely in an economic variable, especially if it is integrated of order one (denoted $\mathsf{I}(1)$) or higher. In fact, as the error standard deviation must be in absolute amounts (e.g. \$100m), it would become negligible relative to the variance of the observed magnitude. Secondly, most economic variables are inherently positive, and constant-parameter linear models cannot ensure such a condition whereas it is binding for log-linear representations. Thirdly, elasticities are often found to be more constant empirically than propensities, again favoring a log-linear model.

Nevertheless, these considerations are not definitive. Indeed, a substantial literature has developed ways of analyzing the problem of discriminating between log and linear equations, including contributions by Cox (1961), Sargan (1964), Aneuryn-Evans and Deaton (1980), Godfrey and Wickens (1982), Bera and McAleer (1983), Davidson and MacKinnon (1985) and Hendry and Richard (1989) inter alia, as well as many papers on the related issue of testing for incorrect functional forms (see Godfrey, 1988, for a survey).

*Financial support from the UK Economic and Social Research Council under grant R00024954 is gratefully acknowledged. We are indebted to Jurgen Doornik, Neil Ericsson, Grayham Mizon, Adrian Pagan and two referees for helpful comments.

Most studies have assumed stationary processes, but many economic time series appear to be $\mathsf{I}(1)$. It is thus interesting to study ways to test log versus linear equations fitted to $\mathsf{I}(1)$ data. This paper offers an approach based on the principle of encompassing (formalized in Hendry and Richard, 1982, Mizon, 1984, and Mizon and Richard, 1986), and on the statistical properties of univariate transformations of $\mathsf{I}(1)$ processes studied by Ermini and Granger (1993). The analysis is illustrated with US quarterly disposable income.

First, the paper discusses the possibility of conducting a specification encompassing test based on the fact that the log random walk fits US quarterly disposable income parsimoniously, in the sense of homoskedastic residuals and time-invariant drift, whereas the linear random walk exhibits both heteroskedasticity and time-increasing drift, with time patterns accurately predictable from the log model. The formal null is a homoskedastic constant-drift, log-normal random walk, and the alternative exhibits features that the linear model should possess when the log model is true. Although conceptually attractive, deriving the asymptotic distribution of this test is complex in the presence of explosive regressors combined with explosive heteroskedasticity.[1] However, we present a Monte Carlo study of its performance in finite samples, customized to the parameter values found in the empirical analysis, which shows reasonable size and good power properties. Interestingly, when applied to US income, the null is not rejected by this test.

The paper then performs a number of mis-specification tests suggested by the encompassing principle. In these tests, the null is the homoskedastic, constant-drift linear random walk model, which is tested for mis-specification against alternative linear models exhibiting heteroskedasticity, exponential drift, and other features, as predicted for the linear model if the log random walk model is true. The linear model is systematically rejected by these tests. The combination of these two sets of tests thus corroborates the log model against the linear model. The Monte Carlo experiment also examines the finite-sample properties of these tests under both null and alternative.

The plan of the paper is as follows. Section 2 summarizes the encompassing principle in the context of the present analysis. Section 3 describes the two competing models, log and linear. Section 4 discusses the specification and mis-specification encompassing tests, and section 5 the customized Monte Carlo experiment. Section 6 applies the tests suggested by the encompassing principle to US income. Section 7 offers some concluding remarks.

2 The principle of encompassing

The idea underlying the principle of encompassing is that economic time series are generated by an unknown, extremely complex, and probably unknowable, process, or data generating process (DGP). The econometrician seeks to mimic salient features of the DGP through simplified representations, obtained by introducing a set of conjectures about the unknown structure of the DGP. These simplified representations are called statistical generating mechanisms (SGM) to emphasize that they must have enough structure to enable the econometrician to actually generate data with them.

To each conjectured SGM there corresponds an empirical model that describes how the SGM actually performs, given the DGP and choice of estimator. For example, suppose that y_t is generated by $\alpha + \beta x_t + u_t$, with a bivariate distribution for $\{u_t, x_t\}$ exhibiting, inter alia, the properties that $u_t|x_t \sim \mathsf{IN}(0, \sigma_u^2)$. Suppose, however, that the econometrician conjectures the SGM that y_t is generated by $\gamma x_t + e_t$, believing $e_t \sim \mathsf{IN}(0, \sigma_e^2)$ and $\mathsf{E}[x_t e_t] = 0$.[2] Since

[1] This specification encompassing test could be performed if desired by simulation estimation as in Mizon and Lu (1993).

[2] The conjecture as stated is incomplete, in that it does not enable data to be generated. To do so, the

$\mathsf{E}[y_t] = \mu_y = \alpha + \beta\mu_x$ when $\mathsf{E}[x_t] = \mu_x$, letting $\mathsf{E}[x_t^2] = m_{xx}$:

$$\mathsf{E}\left[y_t x_t\right] = \alpha\mu_x + \beta\mathsf{E}\left[x_t^2\right], \quad \text{so} \quad \gamma = \beta + \alpha\mu_x m_{xx}^{-1}.$$

When $\mathsf{E}[x_t e_t] = 0$ is enforced by choice of estimator, the empirical model – in population values – corresponding to this SGM is:

$$y_t = \left(\beta + \alpha\mu_x m_{xx}^{-1}\right) x_t + v_t \tag{1}$$

where:

$$v_t = u_t + \alpha\left(1 - \mu_x m_{xx}^{-1} x_t\right) \quad \text{so that} \quad \mathsf{E}[x_t v_t] = 0.$$

However, when $\alpha = \mu_y - \beta\mu_x \neq 0$, v_t is not distributed as $\mathsf{IN}(0, \sigma_u^2)$. Indeed, $\mathsf{E}[v_t] = \alpha(1 - \mu_x^2 m_{xx}^{-1})$, and if $\mathsf{E}[x_t x_{t-1}] = r_1$:

$$\mathsf{E}[v_t v_{t-1}] = \alpha^2\mathsf{E}\left[\left(1 - \mu_x m_{xx}^{-1} x_t\right)\left(1 - \mu_x m_{xx}^{-1} x_{t-1}\right)\right] = \alpha^2\left[1 - 2\mu_x^2 m_{xx}^{-1} + \mu_x^2 m_{xx}^{-2} r_1\right].$$

When $\mathsf{E}[x_t]$ and $\mathsf{E}[x_t^2]$ are non-constant, then the empirical model even has non-constant coefficients. Of course, on discovering such specification 'problems', the SGM may well be revised, but the same principle reapplies, so a revised empirical model results.

Although an econometrician may want to compare SGMs, given the data, he or she can only compare empirical models. During the procedure of modeling, some features of the DGP may be lost, so different empirical models – corresponding to different conjectures – may capture different sets of salient features. It follows that whereas all these empirical models are encompassed by the DGP, some may be non-nested with respect to others. Given any two empirical models, one of which is the model of choice of the econometrician (the 'null' model), and the other a rival model – perhaps constructed after the occurrence of some new pieces of information/theory – encompassing is a way of testing whether the rival model captures features of the DGP that are not already captured by the null model. If the rival model does not offer any new insights about the DGP, then it is said to be encompassed by the null model. The latter could also be encompassed by the rival model, in which case the two models are isomorphic. In this case, the null model is still a valid hypothesis, as it is not successfully challenged by the rival. In terms of our example, the (non-nested) SGM that $y_t = \rho + w_t$, claiming $w_t \sim \mathsf{IN}(0, \sigma_w^2)$, could not encompass, nor be encompassed by, (1) when $\alpha \neq 0$ since the first model excludes the intercept and the second excludes x_t. However, when $\alpha = 0$ but $\mu_x \neq 0$, then $\gamma = \beta$ so the first encompasses the second with $\rho = \beta\mu_x$.

Encompassing tests are based on two fundamental implications of the encompassing principle. First, if the null model encompasses the rival, then it ought to be able to explain the predictions of the rival model. Alternatively, as the encompassed model does not capture all the features of the DGP which are captured by the encompassing model, the null model ought to be able to predict some mis-specifications of the rival model. Both types of prediction are of interest below.

To exposit the idea, suppose that the null model is the first conjecture of the example above, and that the rival model is the second conjecture. Then, the definition of encompassing, whereby the null model encompasses the rival if and only if the parameters of the rival can be predicted from the parameters of the null model, entails that ρ must be predicted to be $\gamma\mu_x$. This is a testable prediction, and forms the basis for the specification encompassing test: the implication that the difference $\rho - \gamma\mu_x$ must be zero can be tested by seeing if $\hat{\rho} - \hat{\gamma}\widehat{\mu_x}$, where $\hat{}$ stands for estimated values, is significantly different from zero.

econometrician must specify the bivariate distribution of (y_t, x_t). However, for the sake of exposition, we consider the conditional analysis.

More generally, denote the DGP $h(\mathbf{y};\boldsymbol{\theta})$ of the stochastic process y_t by M_0, the maintained model $f(\mathbf{y};\boldsymbol{\delta})$ by M_1, and the rival model $g(\mathbf{y};\boldsymbol{\mu})$ by M_2, with $\boldsymbol{\theta}$, $\boldsymbol{\delta}$ and $\boldsymbol{\mu}$ vectors of dimension n, m and p respectively. As the maintained and rival models are reparameterizations of the DGP, their parameter vectors $\boldsymbol{\delta}$ and $\boldsymbol{\mu}$ must be obtained from $\boldsymbol{\theta}$ through suitable mappings $\mathbb{R}^n \to \mathbb{R}^m$: $\boldsymbol{\delta} = \boldsymbol{\phi}_{10}(\boldsymbol{\theta})$, and $\mathbb{R}^n \to \mathbb{R}^p$: $\boldsymbol{\mu} = \boldsymbol{\phi}_{20}(\boldsymbol{\theta})$. If some components of $\boldsymbol{\delta}$ or $\boldsymbol{\mu}$ are not identifiable through $\boldsymbol{\theta}$, then the corresponding model could not be a reparameterization of the DGP.

Then, the null encompasses the rival model if there exists a mapping $\mathbb{R}^m \to \mathbb{R}^p$: $\boldsymbol{\mu} = \boldsymbol{\phi}_{21}(\boldsymbol{\delta})$, that is if the parameterization $\boldsymbol{\mu}$ of the rival model can be predicted from the parameterization $\boldsymbol{\delta}$ of the null model. This condition implies that:

$$\boldsymbol{\phi}_{20}(\boldsymbol{\theta}) = \boldsymbol{\phi}_{21}\left(\boldsymbol{\phi}_{10}(\boldsymbol{\theta})\right). \tag{2}$$

Similarly, M_2 encompasses M_1 if there exists a mapping $\mathbb{R}^p \to \mathbb{R}^m$: $\boldsymbol{\delta} = \boldsymbol{\phi}_{12}(\boldsymbol{\mu})$, which implies that:

$$\boldsymbol{\phi}_{10}(\boldsymbol{\theta}) = \boldsymbol{\phi}_{12}\left(\boldsymbol{\phi}_{20}(\boldsymbol{\theta})\right). \tag{3}$$

Finally, the maintained and rival models are isomorphic, $\mathsf{M}_1 \leftrightarrow \mathsf{M}_2$, when:

$$\boldsymbol{\phi}_{20}(\boldsymbol{\theta}) = \boldsymbol{\phi}_{21}\left(\boldsymbol{\phi}_{10}(\boldsymbol{\theta})\right) = \boldsymbol{\phi}_{21}\left(\boldsymbol{\phi}_{12}\left[\boldsymbol{\phi}_{20}(\boldsymbol{\theta})\right]\right), \forall \boldsymbol{\theta}, \tag{4}$$

so that $\boldsymbol{\phi}_{21}\boldsymbol{\phi}_{12}$ must be the identity transformation.

The specification encompassing test is based on the argument that if M_1 encompasses M_2 then $\boldsymbol{\phi}(\boldsymbol{\delta}) = \boldsymbol{\mu} - \boldsymbol{\phi}_{21}(\boldsymbol{\delta})$ must be zero, and this can be tested by seeing if the statistic:

$$\boldsymbol{\phi}\left(\hat{\boldsymbol{\delta}}\right) = \hat{\boldsymbol{\mu}} - \boldsymbol{\phi}_{21}\left(\hat{\boldsymbol{\delta}}\right) \tag{5}$$

is significantly different from zero.

Alternatively, one can perform a mis-specification encompassing test, based on the implication that the encompassing model ought to be able to predict the mis-specification of the encompassed model. One way to conceive such a test is to extend M_2 parametrically to be consistent with the predictions of the maintained model, and to allow for possible violations. For example, if M_2 assumes white-noise errors, but M_1 predicts that it should manifest autocorrelated errors, then M_2 is extended to have (say) an autoregressive error, and the latter is tested for significance. A mis-specification test predicts both a non-zero value for the added parameter, and a specification test predicts the numerical magnitude it should have on average when M_1 is the DGP. Thus, M_2 should not only be false, it should be wrong in precisely the way predicted by M_1.

3 The two competing models

The data used here are US quarterly, seasonally adjusted, constant dollars 1982, total disposable income from 1951(1) to 1989(2) (see FIG. 1).[3] Both the series in linear values y_t and in log values $\log y_t$ appear to be integrated of order one.[4]

[3] The data is the series GYD.82 of the Citibank Database. The series also contains data from 1947(1), but this subsample was not included to simplify the estimation of a homoskedastic log model.

[4] A Dickey-Fuller test yielded a statistic with unrestricted drift ϕ_3 (see Dickey and Fuller, 1981) of 5.05 for y_t, and 1.68 for $\log y_t$, against a 10%-level critical value of approximately 5.4 (sample size 166). For the linear case, a Perron (1989) test was also performed, with apparent change of growth rate in mid-1966. The unit-root hypothesis is again not rejected, with a t-value of -3.14 for the coefficient of y_{t-1}, against a 5%-level critical value of -3.94 for a sample fraction $\lambda = 0.4$.

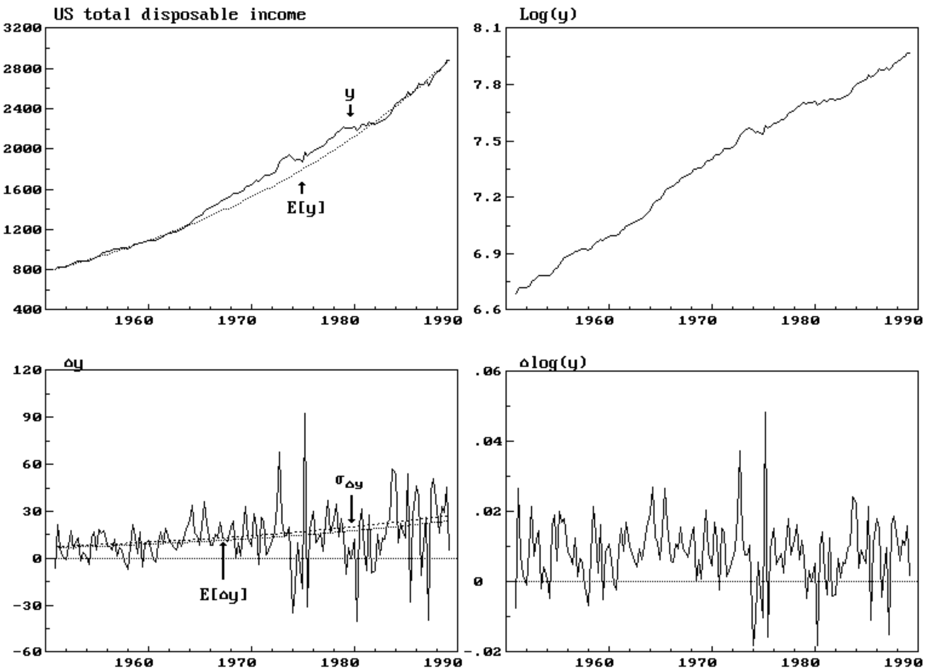

Fig. 1 Time series of US total disposable income

The increasing heteroskedasticity over time of the linear changes is visually clear, and the linear level representation has an apparent quadratic trend. The lines for $\mathsf{E}[y]$, $\sigma_{\Delta y}$ and $\mathsf{E}[\Delta y]$ are discussed below following equations (9) and (10).

To select the best model – in the mean-square sense – to represent both series, a set of ARMA(p, q) models were fitted to the first differences of both series, with $p + q \leq 4$. As the random walk model for the log-linear series was not rejected against the ARMA alternatives, the two following models were selected (standard errors in parenthesis):

$$\Delta \log y_t = \underset{(0.00076)}{0.0083} + \epsilon_t, \quad \hat{\sigma}_{\log y} = 0.0094. \tag{6}$$

$$\Delta y_t = \underset{(1.44)}{13.51} + \eta_t, \quad \hat{\sigma}_y = 17.91. \tag{7}$$

Several features of these estimates merit comment.

(a) *Log random walk.* The recursive estimates of the drift from (6) are shown in FIG. 2a (see Doornik and Hendry, 1994), from which the drift appears fairly constant over time. The residual standard deviation of the log model appears substantially homoskedastic (FIG. 2b), other than the 'spikes' in 1972(4), 1974(1) and 1975(2). These are adequately treated as additive outliers to the growth rate. However, the break-point Chow test fails briefly in mid-sample (FIG. 2d), possibly due to a higher growth rate of disposable income during the Vietnam-war period followed by slower growth after the 'oil-crisis'. We also applied the two constancy statistics in Hansen (1992), denoted V and J for variance and joint constancy, and neither was significant in any specification (with or without impulse dummies for the additive outliers).

The non-normality test was based on the χ^2_2 statistic proposed by Doornik and Hansen (1994), which yielded $\hat{\xi} = 27.4$. This rejects normality, although these residuals are much closer to normality than the residuals from the linear random walk (discussed below). When the three additive outliers are removed, the non-normality test ceases to be significant at the 5% level ($\hat{\xi} = 5.6$). Indeed, in that case, the log model also satisfies tests for residual serial correlation, heteroskedasticity (see White, 1980), and ARCH (see Engle, 1982), providing further support for the choice of the log model as the maintained model.

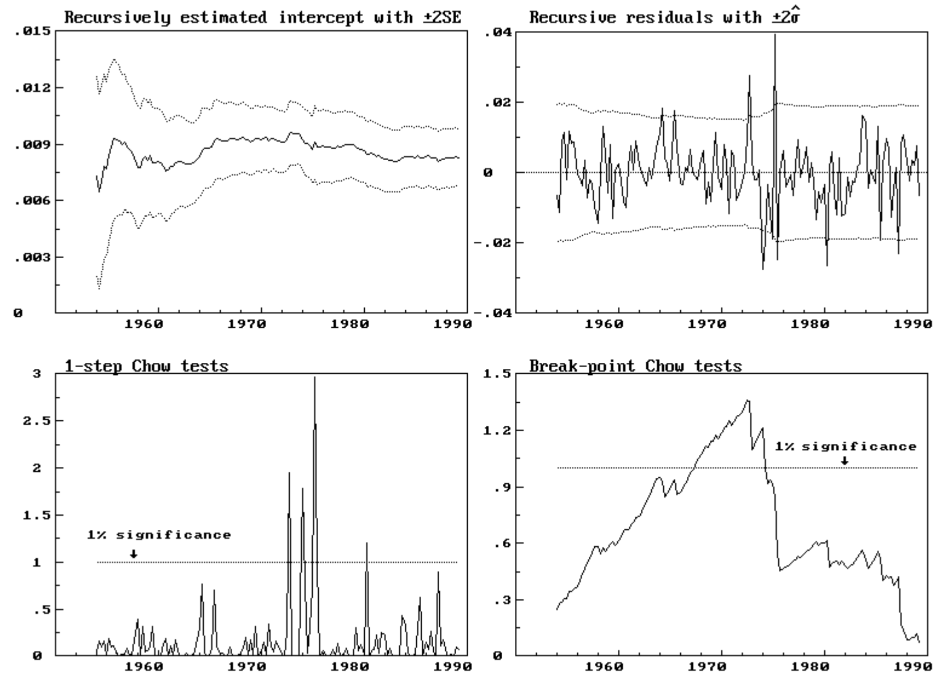

Fig. 2 Recursive statistics of the log model

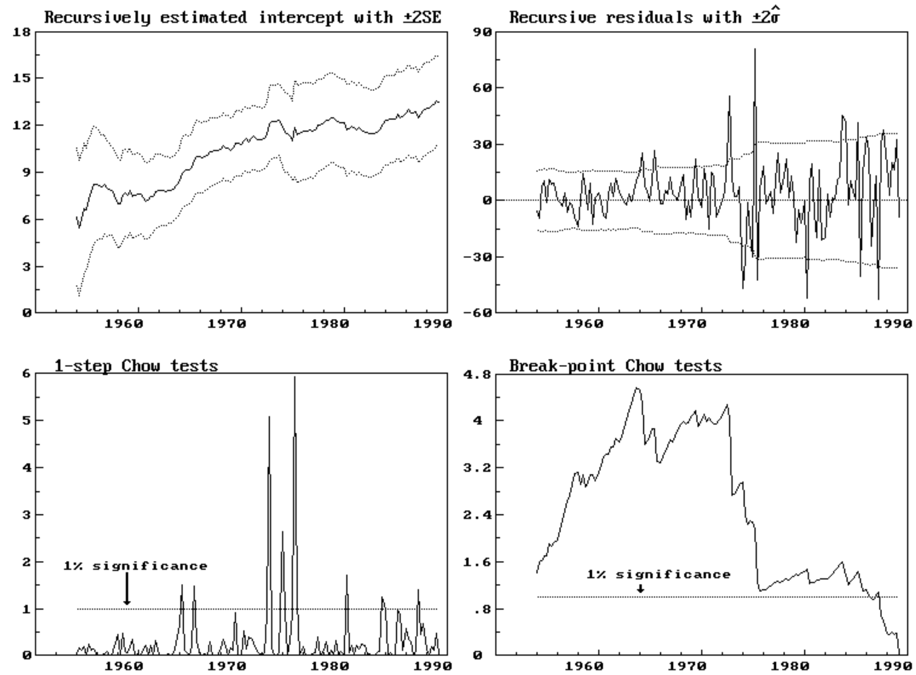

Fig. 3 Recursive statistics of the linear model

(b) *Linear random walk.* Unlike the log-linear model, (7) could be rejected against an equation which also included Δy_{t-3} ($\mathsf{p} < 0.01$). However, noting the discussion in section 2, we persist with (7) as the rival for (6). FIG. 3a reports the recursive estimates of the drift of model (7), which is increasing with time. A similar pattern holds for the residual standard deviation (FIG. 3b); it is seen that this model exhibits increasing heteroskedasticity, as is predicted from the log model (see below). In effect, the break-point Chow (1960) test fails at almost every point (FIG. 3d), indicating substantial non-constancy of the linear random-walk model. The V

and J statistics in Hansen (1992) both also reject the constancy of (7) at the 1% level.

The statistic for non-normality is $\hat{\xi} = 39.7$. Removing the three outliers does not make this statistic insignificant ($\hat{\xi} = 19.7$), so the non-normality is more extensive than in the log model. We comment on the other mis-specification tests below.

From this evidence, it appears that the log model captures more parsimoniously – in the sense of homoskedastic residuals and constant drift – the salient features of US disposable income.

4 Encompassing tests

Let the null model M_1 be the log-normal random walk with drift, $\Delta \log y_t = m + \epsilon_t$, with time-invariant distribution $\Delta \log y_t \sim \mathsf{IN}(m, \sigma^2)$, and with initial value $\mu = \log y_0$; the parameter vector of the null model is thus $\boldsymbol{\alpha} = [m, \sigma, \mu]'$. Let the rival model M_2 be the linear-normal random walk with drift, $\Delta y_t = \gamma + \eta_t$, with distribution claimed to be $\Delta y_t \sim \mathsf{IN}(\gamma, \tau^2)$, and with initial value y_0; its parameter vector is thus $\boldsymbol{\beta} = [\gamma, \tau, y_0]'$. Throughout, $y_t > 0 \ \forall t$. The claimed M_2 distribution is not necessarily time-invariant, nor even necessarily independent, under M_1. To test whether M_1 encompasses M_2 we must know how the parameters of the linear model are related to the parameters of the log model when the latter is treated as the DGP. This relation is established in Ermini and Granger (1993). In particular, we are interested in $\mathsf{E}[\Delta y_t]$ and $\mathsf{V}[\Delta y_t]$.[5]

First, conditioning on y_0:

$$\log y_t = \log y_{t-1} + m + \epsilon_t = \mu + mt + \sum_{j=0}^{t-1} \epsilon_{t-j},$$

so:

$$y_t = \exp\left(\mu + mt + \sum_{j=0}^{t-1} \epsilon_{t-j}\right) = e^{\mu} e^{mt} \exp\left(\sum_{j=0}^{t-1} \epsilon_{t-j}\right).$$

As $\epsilon_t \sim \mathsf{IN}(0, \sigma^2)$:

$$\mathsf{E}\left[y_t\right] = e^{\mu} e^{mt} \mathsf{E}\left[\exp\left(\sum_{j=0}^{t-1} \epsilon_{t-j}\right)\right] = e^{\mu} e^{mt} \exp\left(\frac{\sigma^2}{2} t\right) = e^{\mu} e^{\lambda t}, \tag{8}$$

where $\lambda = (m + \sigma^2/2)$. This is the line shown as $\mathsf{E}\left[y_t\right]$ on FIG. 1a, using the estimates in (6). As can be seen, it provides a good match. From (8):

$$\mathsf{E}\left[\Delta y_t\right] = e^{\mu} e^{\lambda t} - e^{\mu} e^{\lambda (t-1)} = \psi e^{\lambda t} \tag{9}$$

where $\psi = e^{\mu}(1 - e^{-\lambda})$. This predicts the mean of the linear model given that the log model is the DGP.

Next, to derive the variance prediction:

$$\mathsf{E}\left[y_t^2\right] = e^{2\mu} e^{2mt} \mathsf{E}\left[\exp\left(2 \sum_{j=0}^{t-1} \epsilon_{t-j}\right)\right] = e^{2\mu} e^{2mt} \exp\left(2\sigma^2 t\right) = e^{2\mu} e^{2\left(m+\sigma^2\right)t},$$

[5] Different ARIMA parametrizations have been found in the literature that best fit US disposable income data. For example, Watson (1986) found that both the ARI(1,1) and the IMA(1,1) models fit US quarterly GNP in logs. Encompassing tests and mis-specification tests similar to those described in this paper for the random walk case can be extended to these other representations by making direct use of the stochastic properties of exponential transformations of ARI(1,1) and IMA(1,1) models in logs, also reported in Ermini and Granger (1993).

so that:

$$\mathsf{V}[y_t] = \mathsf{E}\left[y_t^2 - e^{2\mu}e^{2\lambda t}\right] = e^{2\mu}e^{2\lambda t}\left(e^{\sigma^2 t} - 1\right).$$

Since:

$$\mathsf{E}[y_t y_{t-1}] = e^{2\mu}e^{2mt}e^{-m}\mathsf{E}\left[\exp\left(\sum_{j=0}^{t-1}\epsilon_{t-j} + \sum_{i=1}^{t-1}\epsilon_{t-i}\right)\right] = e^{2\mu}e^{2\left(m+\sigma^2\right)t}e^{-\left(\lambda+\sigma^2\right)},$$

then:

$$\mathsf{C}[y_t, y_{t-1}] = \mathsf{E}[y_t y_{t-1}] - e^{2\mu}e^{-\lambda}e^{2\lambda t} = e^{2\mu}e^{2\lambda t}e^{-\lambda}\left(e^{\sigma^2(t-1)} - 1\right),$$

and as $\mathsf{V}[\Delta y_t] = \mathsf{V}[y_t] + \mathsf{V}[y_{t-1}] - 2\mathsf{C}[y_t, y_{t-1}]$:

$$\mathsf{V}[\Delta y_t] = e^{2\mu}e^{2\lambda t}\left[e^{\sigma^2 t}\left[1 + e^{-2\delta} - 2e^{-\left(\lambda+\sigma^2\right)}\right] - \left(1 - e^{-\lambda}\right)^2\right] = \phi_1 e^{2\delta t} - \phi_2 e^{2\lambda t}, \tag{10}$$

where $\delta = (\lambda + \sigma^2/2) = (m + \sigma^2)$, and:

$$\phi_1 = e^{2\mu}\left(1 - 2e^{-\left(\lambda+\sigma^2\right)} + e^{-2\delta}\right) \quad \text{and} \quad \phi_2 = e^{2\mu}\left(1 - e^{-\lambda}\right)^2.$$

This predicts the variance of the linear model under a log DGP. Log-normality is necessary for these results.

For the purpose of this paper, the most important feature of (9) and (10) is that the mean and variance of the first differences of the linear random walk are both functions of time. Particularly, for $m + \sigma^2/2 > 0$, both $\mathsf{E}[\Delta y_t]$ and $\mathsf{V}[\Delta y_t]$ are exponentially increasing, as shown above in FIG. 1c. However, with the estimated values $\hat{m} = 0.0083$ and $\hat{\sigma} = 0.0094$ in (6), the exponential growth rate of these time functions is small.

For M_1 to encompass M_2 under stationarity, the following condition must hold (see Sargan, 1964):

$$\hat{\sigma} < \frac{\hat{\tau}}{\bar{y}} \quad \text{or} \quad \hat{\sigma}^2\bar{y}^2 < \hat{\tau}^2, \tag{11}$$

where $\bar{y} = T^{-1}\sum_{t=1}^{T} y_t$, and T is the sample length. For $\mathsf{I}(1)$ processes, it is less clear that the inequality must hold. A heuristic argument for (11) is presented in the appendix, which also shows that the inequality holds empirically at all sample sizes. From (6) and (7), for example, for the full sample:

$$\hat{\sigma} = 0.0094 < \frac{\hat{\tau}}{\bar{y}} = \frac{17.91}{1683} = 0.0106. \tag{12}$$

The specification encompassing test is based on the argument that if the null model is true, then the linear model must have the two properties (9) and (10). Not only does this entail that the linear random walk (7) with constant drift and homoskedastic innovations is grossly mis-specified, but more importantly that from the null model we should be able to predict accurately the way the linear model is mis-specified. From (9):

$$\Delta y_t = \mathsf{E}[\Delta y_t] + \omega_t = \rho\left[\psi e^{\lambda t}\right] + \omega_t,$$

so that in the regression:

$$\Delta y_t = \pi + \rho\left[\hat{\psi}e^{\hat{\lambda}t}\right] + \eta_t, \tag{13}$$

the exponential regressor $\exp(\hat{\lambda}t)$, constructed with the estimates of the log model (6), should be significant, so H_0^1: $\rho = 0$ entails that M_1 does not encompass M_2. As the unconditional mean $\mathsf{E}[\Delta y_t]$ predicted from (9) has no constant, the hypothesis 'M_1 encompasses M_2' may be corroborated only if $\rho = 1$ and $\pi = 0$. Test statistics for either or both:

$$\mathsf{H}_0^2 : \pi = 0 \ \text{ versus } \ \mathsf{H}_a^2 : \pi \neq 0, \quad \text{or} \quad \mathsf{H}_0^3 : \rho = 1 \ \text{ versus } \ \mathsf{H}_a^3 : \rho \neq 1 \tag{14}$$

have non-standard distributions under M_1, their complexity deriving from $\hat{\lambda}$ being an estimate and not the true value, from the regressand being heteroskedastic according to (10), and from the regressor being explosive. The joint test is denoted $\mathsf{F}_{\mathsf{enc}}$ below, although it does not have an F-distribution under encompassing. Even when $\psi e^{\lambda t}$ is known, (10) shows that the standard deviation needed to induce homoskedasticity increases faster than $\mathsf{E}[\Delta y_t]$, so the equation is asymptotically singular. However, in small samples that problem does not arise, so we examine this test below, using Monte Carlo calibrated critical values (bootstrap-based values could be used empirically).[6]

Mis-specification tests of (7) suggested by the encompassing principle include that η_t should be heteroskedastic and non-normal, and possibly autocorrelated, and that a linear trend should be significant if added. These tests are well behaved when the linear model is the DGP. Prior to implementing both types of tests on US income, we undertake a Monte Carlo study of their behavior in finite samples, customized to the present empirical analysis.

5 A Monte Carlo study

To analyze the finite-sample properties of these tests, three sets of Monte Carlo experiments were undertaken using Ox (see Doornik, 1994) and PC-NAIVE (see Hendry, Neale and Ericsson, 1991). In the first and third sets, the log random walk comprised the DGP; in the second set, the linear random walk is the DGP. The first and second sets used estimated parameters, whereas the third set used known parameters to facilitate recursive Monte Carlo.

The parameter values of the Monte Carlo are centered around those of the empirical study, namely $m = 0.01$ and $\sigma = 0.01$ (μ was set at 0.01, but the test rejection frequencies are invariant to its value). First, four experiments were examined at $(m, \sigma) = (0.01, 0.01)$; $(0.01, 0.02)$; $(0.02, 0.01)$; $(0.02, 0.02)$, denoted (a), (b), (c) and (d) respectively. The maximum sample size was set at $T = 200$. Equation (13) was estimated, adding $\hat{\psi}\exp(\hat{\lambda}t)$ where $\hat{m} = T^{-1}\log(y_T/y_1)$ and $\hat{\sigma}$ were estimated in each replication. Tests were conducted of H_0^1: $\rho = 0$ ($\mathsf{t}_{\rho=0}$); H_0^2: $\pi = 0$ ($\mathsf{t}_{\pi=0}$); H_0^3: $\rho = 1$ ($\mathsf{t}_{\rho=1}$); and H_0^2 and H_0^3 holding jointly ($\mathsf{F}_{\mathsf{enc}}$); these tests should be well behaved when a model like (7) is the DGP. We also examined the power of autocorrelation and normality mis-specification tests in equations like (7), and heteroskedasticity in (13).

Table 1 Specification encompassing test rejection frequencies under log DGP

Alternative		(a) 5%	(a) 1%	(b) 5%	(b) 1%	(c) 5%	(c) 1%	(d) 5%	(d) 1%
$T = 100$	$\mathsf{t}_{\pi=0}$	0.049	0.012	0.038	0.008	0.045	0.012	0.038	0.007
	$\mathsf{t}_{\rho=1}$	0.055	0.015	0.072	0.025	0.052	0.015	0.060	0.015
	$\mathsf{F}_{\mathsf{enc}}$	0.023	0.003	0.029	0.011	0.022	0.003	0.026	0.005
$T = 200$	$\mathsf{t}_{\pi=0}$	0.030	0.005	0.016	0.001	0.029	0.004	0.026	0.004
	$\mathsf{t}_{\rho=1}$	0.060	0.019	0.114	0.041	0.060	0.011	0.122	0.045
	$\mathsf{F}_{\mathsf{enc}}$	0.020	0.001	0.054	0.015	0.018	0.004	0.062	0.019

Tables 1 and 2 report the outcomes of the first set of tests, the first block at $T = 100$, the

[6] A specification encompassing test could also be performed in relation to the autocorrelation function of the linear model (perhaps by Monte Carlo experiments as in Mizon and Lu (1993).). Ermini and Granger (1993) find that the exponential transformation of a random walk process produces an $\mathsf{I}(0)$ process with an autocorrelation function identical to that of an AR(1) process, with coefficient of autoregression equal to $\exp(-\hat{\sigma}^2/2)$. Here, this test has no power to distinguish between the hypothesis of M_1 encompassing M_2 (which implies an autoregression coefficient less than one in absolute value), and the hypothesis of non-encompassing, implied by a unit root. In fact, with the estimates from the log model, this coefficient equals 0.9953, which – although consistent with the prediction of the theory – is statistically indistinguishable from the unit root.

second at $T = 200$.[7] The test null-rejection frequencies are reasonable, with a slight tendency to under-reject compared to the nominal significance level, except for $\mathsf{t}_{\rho=1}$, but little improvement as T increases. However, in finite samples, the tests do not manifest serious deficiencies despite the explosive heteroskedasticity.

Table 2 Mis-specification encompassing test rejection frequencies under log DGP

	Alternative	(a) 5%	(a) 1%	(b) 5%	(b) 1%	(c) 5%	(c) 1%	(d) 5%	(d) 1%
$T = 100$	$\mathsf{t}_{\rho=0}$	0.747	0.542	1.000	1.000	0.341	0.180	0.970	0.929
	autocorrelation	0.136	0.054	0.982	0.954	0.080	0.026	0.471	0.328
	non-normality	0.445	0.262	0.999	0.989	0.390	0.217	0.935	0.844
	heteroskedasticity	0.854	0.644	1.000	1.000	0.848	0.665	1.000	0.996
$T = 200$	$\mathsf{t}_{\rho=0}$	1.000	0.999	1.000	1.000	0.837	0.698	1.000	1.000
	autocorrelation	0.677	0.536	1.000	1.000	0.244	0.129	0.919	0.864
	non-normality	0.997	0.990	1.000	1.000	0.977	0.947	1.000	1.000
	heteroskedasticity	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

The power of the variable-addition test is high except when the drift is small compared to the error standard deviation. The non-normality also does well but the residual autocorrelation is really only detectable when the drift is large, even at the larger sample size. The heteroskedasticity test based on (13) is particularly powerful, and provides a good reason for using the encompassing implications in this context. As shown below (see Table 4), its actual size is close to its nominal.

The converse issue is the power of the specification encompassing tests to reject the log model when the linear model is the DGP. Thus, in the second set, we reverse the roles of the two models, generating data from, and still fitting, the linear model, basing the tests on the predictions the log model would have made.[8] Tables 3 and 4 record the outcomes, using $y_0 = 1000$, $\gamma = 15$, 30, and $\tau = 15$, 30.

Table 3 Specification encompassing test rejection frequencies under linear DGP

	Alternative	(a) 5%	(a) 1%	(b) 5%	(b) 1%	(c) 5%	(c) 1%	(d) 5%	(d) 1%
$T = 100$	$\mathsf{t}_{\pi=0}$	0.688	0.464	1.000	1.000	0.238	0.088	0.939	0.816
	$\mathsf{t}_{\rho=1}$	0.694	0.476	1.000	1.000	0.263	0.111	0.949	0.842
	$\mathsf{F}_{\mathsf{enc}}$	0.537	0.308	1.000	1.000	0.141	0.049	0.884	0.729
$T = 200$	$\mathsf{t}_{\pi=0}$	1.000	0.996	1.000	1.000	0.733	0.520	1.000	1.000
	$\mathsf{t}_{\rho=1}$	1.000	0.997	1.000	1.000	0.793	0.590	1.000	1.000
	$\mathsf{F}_{\mathsf{enc}}$	0.998	0.983	1.000	1.000	0.641	0.428	1.000	1.000

As can be seen, the powers of the specification tests to reject encompassing are high unless the relative drift is small, given that their sizes are about right from Table 1. In turn, Table 4 reveals that the sizes of all the mis-specification tests are close to their nominal values. Such evidence suggests that valid inference will result when the linear model is the DGP, with good power to reject that specification when the log model generated the data.

Finally, experiments were conducted recursively at every sample size from 20 to 200, using $1,000$ replications so that the standard errors of test rejection frequencies are always less than

[7] The heteroskedasticity test is based on equation (13).

[8] Jalan (1989) showed that the actual size of the $\mathsf{t}_{\rho=0}$ test was approximately the nominal at the 5% level when the DGP is the linear model (with a large positive initial value), and the term $\exp(\hat{\lambda}t)$ is added, using sample estimates.

Table 4 Mis-specification encompassing test rejection frequencies under linear DGP

Alternative	(a) 5%	(a) 1%	(b) 5%	(b) 1%	(c) 5%	(c) 1%	(d) 5%	(d) 1%
$T = 100$ $\mathsf{t}_{\rho=0}$	0.054	0.011	0.056	0.014	0.042	0.016	0.049	0.009
autocorrelation	0.040	0.006	0.060	0.011	0.047	0.010	0.050	0.011
non-normality	0.048	0.013	0.062	0.016	0.048	0.017	0.046	0.015
heteroskedasticity	0.045	0.006	0.052	0.009	0.041	0.007	0.043	0.007
$T = 200$ $\mathsf{t}_{\rho=0}$	0.049	0.009	0.047	0.008	0.041	0.011	0.047	0.013
autocorrelation	0.040	0.008	0.044	0.005	0.045	0.009	0.045	0.003
non-normality	0.050	0.014	0.043	0.012	0.049	0.017	0.042	0.014
heteroskedasticity	0.044	0.010	0.042	0.007	0.050	0.011	0.046	0.012

0.015. Three models were estimated: equations (7), (13) and an equation obtained from (13) by correcting for the heteroskedasticity predicted by (10) (i.e., by scaling down the relevant variables at each sampling point by $\sqrt{(\mathsf{V}[\Delta y_t])}$). Here, the population values of λ (rather than estimated values) were used. The outcomes provide a slightly favorable picture of the performance of the $\mathsf{t}_{\rho=0}$ test, but not by much in large samples in comparison to the results recorded above (cross-checking the size of this 'infeasible' test showed it to be similar to that reported above).

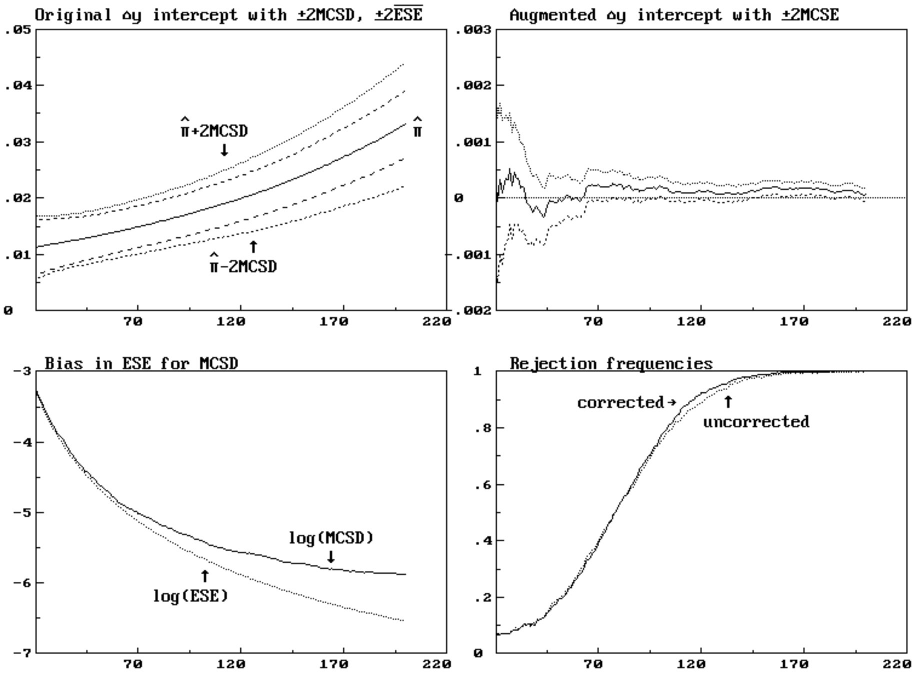

Fig. 4 Recursive Monte Carlo statistics

Figure 4a shows the mean Monte Carlo value of $\hat{\pi}$ together with $\pm 2\overline{ESE}$ (the mean Monte Carlo estimated standard error) and $\pm 2MCSD$ (the Monte Carlo sampling standard deviation of the individual $\hat{\pi}_i$ across replications). The upward trend reflects equation (9), the downward bias in the $\overline{ESE}$ for the $MCSD$ is due to the residual heteroskedasticity (see (10)), and the average outcome is representative of FIG. 3a holding in random replications. The average estimated equation standard error increased monotonically with the sample size, from 0.011 at $T = 20$ to 0.038 at $T = 200$, and the break-point Chow (1960) test rejected the null of constancy 100% of the time at all sample sizes except the last few (i.e., tests of a break between

$T = 195$ and $T = 200$). This evidence is again consistent with the US income results when the log model is valid.

Figure 4b records the mean behavior of $\hat{\pi}$ when $\psi\exp(\lambda t)$ is added as a regressor (as in (13)), together with $\pm 2MCSE$ (the Monte Carlo standard error, which measures the uncertainty in the estimate of the mean of $\hat{\pi}$, and is defined as $MCSD/\sqrt{M}$ for M replications). There is no trend in the estimate of the intercept, and none in the estimated coefficient of $\psi\exp(\lambda t)$ (not shown). However, the average ESE is a downward biased estimator of $MCSD$ due to the heteroskedasticity, as can be seen in FIG. 4c which graphs the logs of these two quantities to highlight percentage differences. Similar outcomes obtained for the models with heteroskedasticity corrections.

The more interesting issue concerns the rejection frequency of H_0^1: $\rho = 0$ for a 5% nominal significance level, and the corresponding rejection of the null that the intercept equals zero. The former, denoted P_a, is shown in FIG. 4d for both heteroskedasticity-corrected and uncorrected regressions, and reveals similar and typically shaped power functions, rapidly rising to high power for sample size exceeding 100 observations. The null H_0^2: $\pi = 0$ was rejected 4.6% of the time without a heteroskedasticity correction, and 9.3% with, so the transform was unhelpful in this Monte Carlo experiment.[9]

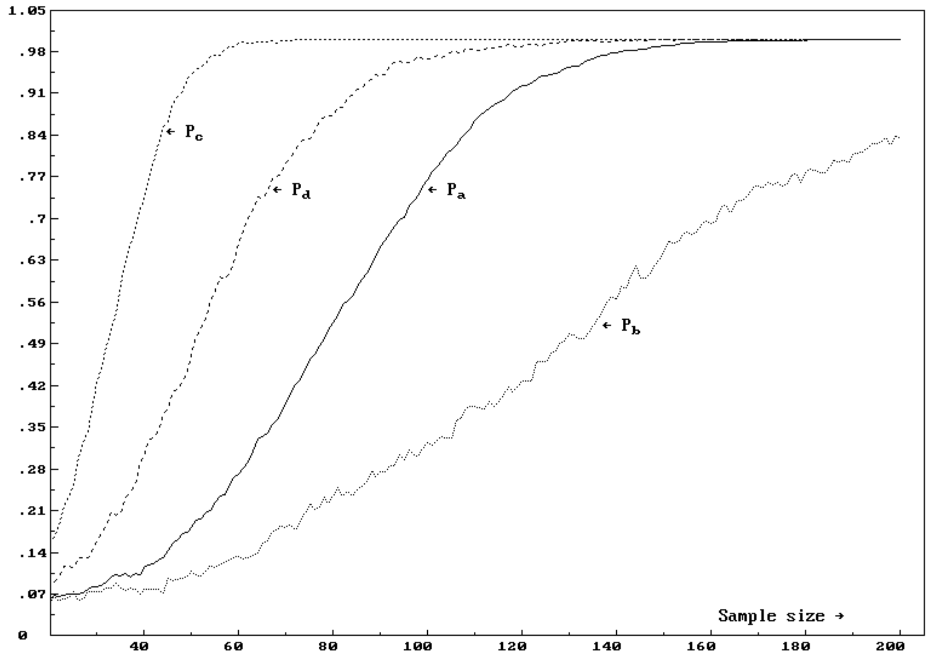

Fig. 5 Recursive Monte Carlo rejection frequencies

The effects of varying the design parameters were as might be anticipated from (9) and (10): increasing m improved the mis-specification test performance (P_c), increasing σ reduced it (P_b), and increasing both but keeping their ratio constant also improved the performance of the test (P_d). Figure 5 shows the test powers across all sample sizes and experiments for the heteroskedastic-corrected version of (13). At a given sample size, power is much higher at the larger λ value, and much lower at the larger σ (of 2% of y rather than 1% of y, which corresponds to a relatively noisy series), but the power loss can be offset by a larger sample.

The other aspect of the mis-specification encompassing test is that the intercept should not be significant after including $\exp(\lambda t)$. Graphs of the rejection frequencies of the intercept in the

[9] This may be due to the fact that, since the variance is an exponentially increasing function of t, the estimates of the heteroskedasticity-corrected intercept are inconsistent.

four experiments are not very revealing as they all cluster around the 5% level, and show little tendency to vary with changes in the sample size. Instead, we report the overall mean values of these rejection frequencies averaged across both M and T together with standard errors across T of the average at every sample size, in each of the four experiments for equations of the form (13). The outcomes were as follows:

$$\begin{bmatrix} \text{Experiment} & \text{size} & \text{standard error} \\ (a) & 0.052 & 0.007 \\ (b) & 0.048 & 0.008 \\ (c) & 0.050 & 0.008 \\ (d) & 0.052 & 0.007 \end{bmatrix}$$

These are close to earlier sizes based on estimated parameters, and also close to the nominal size of 0.05, with standard errors near the theoretical standard errors across replications of 0.007 (based on $[\mathsf{p}(1-\mathsf{p})]^{1/2}/\sqrt{M}$). There is little variation in the rejection frequency with changes in the design parameters of the experiments.

The behavior of the empirical example is consistent with these Monte Carlo findings for a null of a log-linear random walk when a linear difference is incorrectly fitted. A random replication from the artificial data was analyzed in detail and revealed the same pattern as that found for US disposable income.

6 Encompassing tests applied to US disposable income

The encompassing principle, and the statistical properties of the exponential transformation of $\mathsf{I}(1)$ processes (particularly, (9) and (10)), suggest applying both sets of specification and mis-specification tests. The power of the tests can be increased if we do not restrict the alternatives to the exact predictions of the linear model under the log model discussed in section 4. For example, the null would be rejected not only against an alternative exhibiting an exponentially increasing drift of type (13), but also against an alternative exhibiting a linearly increasing drift with time slope of the order of $(y_T - y_1)/T$. This also permits a relaxation of the normality restriction imposed by (9) and (10).

First, linear model mis-specification tests were conducted, using the equation in (15), with the three impulse dummies denoted by D_i for the outliers discussed above:

$$\widehat{\Delta y_t} = \underset{(2.94)}{4.16} + \underset{(0.07)}{0.20}\,\Delta y_{t-3} + \underset{(15.2)}{54.6}\,D_1 - \underset{(15.2)}{49.1}\,D_2 + \underset{(15.2)}{79.0}\,D_3 + \underset{(0.028)}{0.066}\,t \tag{15}$$

where $\hat{\sigma}_{\Delta y} = 15.1$. Since $\bar{y} = 1683$, the implied proportional equation standard error of the linear model is 0.0090. Table 5 records the tests on the residuals.[10]

Table 5 US income linear model mis-specification tests

Alternative	Test	Value
autocorrelation	$F(5,141)$	0.2
ARCH	$F(4,138)$	1.8
heteroskedasticity	$F(10,135)$	3.80^{**}
non-normality	$\chi^2(2)$	20.9^{**}
RESET	$F(1,145)$	0.03
constancy	V, J	2.29^{**}, 3.27^{**}

[10] * and ** denote significance at the 5% and 1% level respectively.

From these results, it is clear that (15) is not an adequate specification, even ignoring the improbable implication of a constantly increasing absolute change.

By way of comparison, the log-random walk delivered:

$$\widehat{\Delta \log y_t} = \underset{(0.0007)}{0.008} + \underset{(0.008)}{0.029}\, D_1 - \underset{(0.008)}{0.026}\, D_2 + \underset{(0.008)}{0.040}\, D_3 \tag{16}$$

with $\hat{\sigma}_{\Delta \log y} = 0.0081$. Table 6 records the tests on the residuals.

Table 6 US income log model mis-specification tests

Alternative	Test	Value
autocorrelation	$F(5,142)$	0.66
ARCH	$F(2,142)$	0.35
heteroskedasticity	$F(8,138)$	1.3
non-normality	$\chi^2(2)$	4.1
RESET	$F(1,146)$	1.13
omitted trend	$F(1,146)$	0.73
constancy	V,J	0.18, 0.92

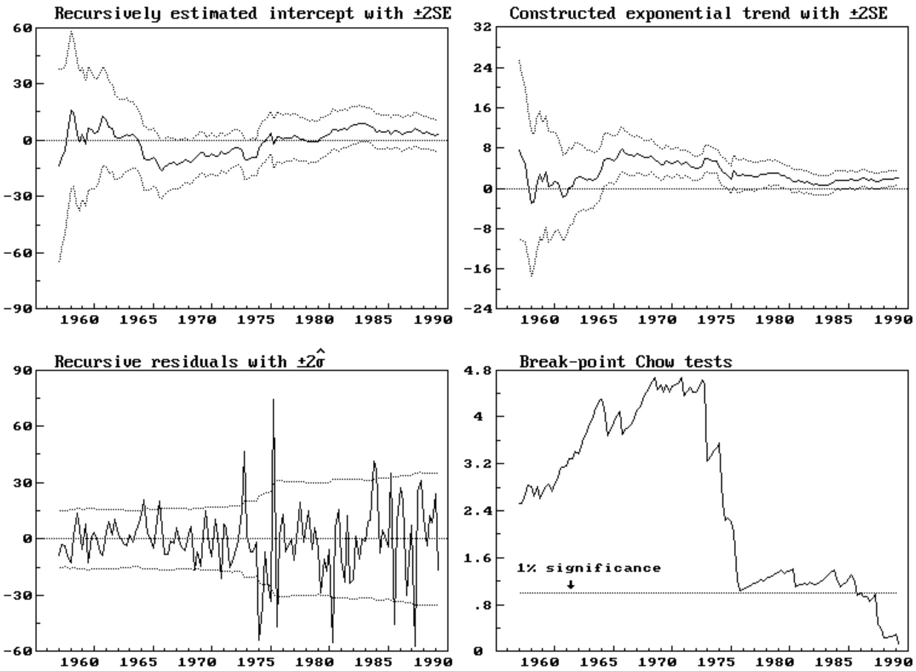

Fig. 6 Recursive statistics with the constructed exponential trend

Finally, equations (17) and (18) report the regression results after adding the constructed exponential trend, with and without the heteroskedasticity correction:

$$\Delta y_t = \underset{(4.12)}{3.06} + \underset{(0.29)}{0.77}\, \mathsf{E}[\Delta y_t] + v_t, \quad \hat{\sigma}_{\Delta y} = 17.55; \tag{17}$$

and:

$$\Delta y_t^* = \underset{(3.07)}{1.72}\, \sigma_{\Delta y}^{-1} + \underset{(0.28)}{0.88}\, \mathsf{E}[\Delta y_t] \sigma_{\Delta y}^{-1} + v_t^*, \quad \hat{\sigma}_{\Delta y^*} = 0.78 \tag{18}$$

where $\Delta y_t/\sigma_{\Delta y} = \Delta y_t^*$. As can be seen, the constant is insignificant in (17), whereas the coefficient of $\mathsf{E}[\Delta y_t]$ is significantly different from zero, but not from unity (its encompassing prediction). The joint test of $\rho = 1$ and $\pi = 0$ (in the notation of (13)) yields $\mathsf{F}^2_{152} = 0.32$. This encompassing hypothesis is equally easily satisfied in (18), where constancy and homoskedasticity hold after partialling out the three dummies ($\mathsf{F}^2_{149} = 0.39$).

Figure 6 shows the impact of adding the constructed exponential trend on the constancy statistics of the linear model. As FIG. 6a reveals, the intercept was never significant, whereas (FIG. 6b) the exponential trend was. The two sets of Chow statistics in FIG. 6c–d remain as for the linear case, in line with the theory predictions. Adding three impulse dummies for the outliers does not alter this conclusion, whereas scaling Δy, the intercept and the constructed exponential by $\sigma_{\Delta y}$ yields a relatively constant outcome, close to that of the log model.

7 Conclusions

This paper applies the principle of encompassing to suggest a number of tests of log versus linear individual equations fitted to $\mathsf{I}(1)$ data, and illustrates the analysis for US quarterly disposable income. The paper discusses both specification and mis-specification encompassing tests, and examines their finite-sample properties in a Monte Carlo experiment customized to the parameter values found in the empirical analysis.

Regarding the tests suggested by the principle of encompassing, in the first set of tests, the null is the homoskedastic, constant-drift log random walk model, and the alternative exhibits features that the linear model should possess if the log model is true. The null of encompassing in this first set of tests is not rejected empirically, matching the Monte Carlo.

In the second set of tests, the null is the homoskedastic, constant-drift linear random walk model, which is tested for mis-specification against alternative linear models exhibiting heteroskedasticity, exponential drift, and other features, as predicted for the linear model if the log model is true. The null of these tests is systematically rejected for US income and again matches the Monte Carlo findings. The combination of the results of these two sets of tests thus corroborate the log model against the linear model.

References

Aneuryn-Evans, G., and Deaton, A. S. (1980). Testing linear versus logarithmic regression models. *Review of Economic Studies*, **47**, 275–91.

Bera, A. K., and McAleer, M. (1983). Nested and non-nested procedures for testing linear and log-linear regression models. *Sankhya*, **51**, 212–24. Series B.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, **28**, 591–605.

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.

Davidson, R., and MacKinnon, J. G. (1985). Testing linear and log-linear regressions against Box-Cox alternatives. *Canadian Journal of Economics*, **18**, 499–517.

Dickey, D. A., and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, **49**, 1057–1072.

Doornik, J. A. (1994). The Ox language. Mimeo, Nuffield College, University of Oxford.

Doornik, J. A., and Hansen, H. (1994). A practical test for univariate and multivariate normality. Discussion paper, Nuffield College.

Doornik, J. A., and Hendry, D. F. (1994). *PcGive 8: An Interactive Econometric Modelling System.* London: International Thomson Publishing, and Belmont, CA: Duxbury Press.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity, with estimates of the variance of United Kingdom inflations. *Econometrica*, **50**, 987–1007.

Ermini, L., and Granger, C. W. J. (1993). Some generalizations on the algebra of I(1) processes. *Journal of Econometrics*, **58**, 369–84.

Godfrey, L. G. (1988). *Misspecification Tests in Econometrics.* Cambridge: Cambridge University Press.

Godfrey, L. G., and Wickens, M. R. (1982). Tests of misspecification using locally equivalent alternative models. In Chow, G. C., and Corsi, P. (eds.), *Evaluating the Reliability of Macro-Economic Models*, Ch. 6. New York: John Wiley.

Hansen, B. E. (1992). Testing for parameter instability in linear models. *Journal of Policy Modeling*, **14**, 517–533.

Hendry, D. F., Neale, A. J., and Ericsson, N. R. (1991). *PC-NAIVE, An Interactive Program for Monte Carlo Experimentation in Econometrics. Version 6.0.* Oxford: Institute of Economics and Statistics, University of Oxford.

Hendry, D. F., and Richard, J.-F. (1982). On the formulation of empirical models in dynamic econometrics. *Journal of Econometrics*, **20**, 3–33. Reprinted in Granger C. W. J. (ed.) (1990), *Modelling Economic Series.* Oxford: Clarendon Press and in Hendry D. F. (1993), *Econometrics: Alchemy or Science?* Oxford: Blackwell Publishers.

Hendry, D. F., and Richard, J.-F. (1989). Recent developments in the theory of encompassing. In Cornet, B., and Tulkens, H. (eds.), *Contributions to Operations Research and Econometrics. The XXth Anniversary of CORE*, pp. 393–440. Cambridge, MA: MIT Press.

Jalan, J. (1989). Monte carlo evidence on an encompassing test for log vs linear in nonstationary cases. unpublished paper, Economics Department, University of California, San Diego.

Mizon, G. E. (1984). The encompassing approach in econometrics. In Hendry, D. F., and Wallis, K. F. (eds.), *Econometrics and Quantitative Economics*, pp. 135–172. Oxford: Basil Blackwell.

Mizon, G. E., and Lu, M. (1993). Monte carlo encompassing. Mimeo, Economics Department, Southampton University, Southampton.

Mizon, G. E., and Richard, J.-F. (1986). The encompassing principle and its application to non-nested hypothesis tests. *Econometrica*, **54**, 657–678.

Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis. *Econometrica*, **57**, 1361–1401.

Sargan, J. D. (1964). Wages and prices in the United Kingdom: A study in econometric methodology (with discussion). In Hart, P. E., Mills, G., and Whitaker, J. K. (eds.), *Econometric Analysis for National Economic Planning*, Vol. 16 of *Colston Papers*, pp. 25–63. London: Butterworth Co. Reprinted as pp. 275–314 in Hendry D. F. and Wallis K. F. (eds.) (1984). *Econometrics and Quantitative Economics.* Oxford: Basil Blackwell, and as pp. 124–169 in Sargan J. D. (1988), *Contributions to Econometrics*, Vol. 1, Cambridge: Cambridge University Press.

Watson, W. M. (1986). Univariate detrending methods with stochastic trends. *Journal of Monetary Economics*, **18**, 49–75.

White, H. (1980). A heteroskedastic-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, **48**, 817–838.

8 Appendix

For M_1 to encompass M_2 under stationarity, the following condition must hold (see Sargan, 1964):

$$\hat{\sigma} < \frac{\hat{\tau}}{\bar{y}} \quad \text{or} \quad \hat{\sigma}\bar{y} < \hat{\tau}, \tag{19}$$

where $\bar{y} = T^{-1}\sum_{t=1}^{T} y_t$, and T is the sample length. When $\log y_t$ is an $\mathsf{I}(1)$ process, the implication is less clear since y_t is explosive. A heuristic argument for (19) is as follows. The units of y_t are independent of those of $\hat{\sigma}$ so we consider whether on M_1:

$$\frac{\sigma^2 \left(\mathsf{E}\left[\bar{y}\right]\right)^2}{\mathsf{E}\left[\hat{\tau}^2\right]} < 1, \tag{20}$$

where, as $\hat{\tau}^2 = T^{-1}\sum_{t=1}^{T}(\Delta y_t - \overline{\Delta y})^2$:

$$\mathsf{E}\left[\hat{\tau}^2\right] = T^{-1}\sum_{t=1}^{T}\left(\mathsf{V}\left[\Delta y_t\right] + \left(\mathsf{E}\left[\Delta y_t\right]\right)^2 - \mathsf{E}\left[\left(\overline{\Delta y}\right)^2\right]\right),$$

and:

$$T^{-1}\sum_{t=1}^{T}\mathsf{V}\left[\Delta y_t\right] = T^{-1}\sum_{t=1}^{T}\left(\phi_1 e^{2\delta t} - \phi_2 e^{2\lambda t}\right),$$

where:

$$T^{-1}\sum_{t=1}^{T}\left(\mathsf{E}\left[\Delta y_t\right]\right)^2 = T^{-1}\phi_2\sum_{t=1}^{T} e^{2\lambda t}.$$

Since:

$$\sum_{t=1}^{T} e^{\lambda t} = \left(\frac{e^{\lambda T} - 1}{1 - e^{-\lambda}}\right), \tag{21}$$

so:

$$T^{-1}\sum_{t=1}^{T}\left(\mathsf{V}\left[\Delta y_t\right] + \left(\mathsf{E}\left[\Delta y_t\right]\right)^2\right) = T^{-1}\phi_1\sum_{t=1}^{T} e^{2\delta t} = T^{-1}\phi_1 e^{2\delta T}\frac{1 - e^{-2\delta T}}{1 - e^{-2\delta}}$$

Also, as:

$$\mathsf{E}\left[y_T y_1\right] = e^{2\mu}e^{mT}e^{m}\mathsf{E}\left[\exp\left(\sum_{j=0}^{T-1}\epsilon_{T-j} + \epsilon_1\right)\right] = e^{2\mu}e^{\lambda T}e^{\lambda}e^{\sigma^2},$$

and since:

$$\mathsf{E}\left[\ \overline{\Delta y}\ \right] = T^{-1}\mathsf{E}\left[y_T - y_1\right] = T^{-1}e^{\mu}e^{\lambda T}\left[1 - e^{-\lambda T}\right],$$

then:

$$\mathsf{E}\left[\left(\overline{\Delta y}\right)^2\right] = T^{-2}\mathsf{E}\left[y_T^2 + y_1^2 - 2y_T y_1\right] = T^{-2}e^{2\mu}\left[e^{2\delta T} + e^{2\delta} - 2e^{\lambda T}e^{\lambda}e^{\sigma^2}\right]. \tag{22}$$

Thus:

$$\mathsf{E}\left[\bar{y}\right] = e^{\mu}T^{-1}\sum_{t=1}^{T} e^{\lambda t} = T^{-1}e^{\mu}e^{\lambda T}\left(\frac{1 - e^{-\lambda T}}{1 - e^{-\lambda}}\right).$$

Finally, using (21):

$$\mathsf{E}\left[\hat{\tau}^2\right] = T^{-1}\left[\phi_1 e^{2\delta T}\frac{1 - e^{-2\delta T}}{1 - e^{-2\delta}}\right] - T^{-2}e^{2\mu}\left[e^{2\delta T} + e^{2\delta} - 2e^{\lambda T}e^{\lambda}e^{\sigma^2}\right].$$

Consequently from (20) (dropping the asymptotically negligible T^{-2} term):

$$\frac{\sigma^2 \left(\mathsf{E}\left[\bar{y}\right]\right)^2}{\mathsf{E}\left[\hat{\tau}^2\right]} \simeq \frac{\sigma^2 T^{-2} e^{2\mu} e^{2\lambda T}\left(1-e^{-\lambda T}\right)^2\left(1-e^{-2\delta}\right)}{T^{-1}\phi_1 e^{2\delta T}\left(1-e^{-2\delta T}\right)\left(1-e^{-\lambda}\right)^2} = \frac{\sigma^2 e^{2\mu}\left(1-e^{-2\delta}\right)}{\phi_1\left(1-e^{-\lambda}\right)^2}\cdot\frac{\left(1-2e^{-\lambda T}+e^{-2\lambda T}\right)}{Te^{\sigma^2 T}\left(1-e^{-2\delta T}\right)}.$$

As $T \rightarrow \infty$, the second fraction tends to 0, so the inequality is bound to be satisfied.

FIG.7 shows recursive estimates of $\bar{y}$ with $\mathsf{E}\left[\bar{y}\right]$; $\overline{\Delta y}$ with $\mathsf{E}[\overline{\Delta y}]$; $\hat{\tau}^2$ with $\mathsf{E}\left[\hat{\tau}^2\right]$; and $\hat{\sigma}\bar{y}/\hat{\tau}$ with $\sigma\mathsf{E}\left[\bar{y}\right]/\mathsf{E}\left[\hat{\tau}\right]$: the expected values are calculated numerically using (6) for estimates of the parameters. These numerical calculations corroborate reasonably except for the Sargan inequality, which holds everywhere but overpredicts: mismatches may reflect the need for impulse dummies in the linear model. Nevertheless, the approximate analytic inequality holds.

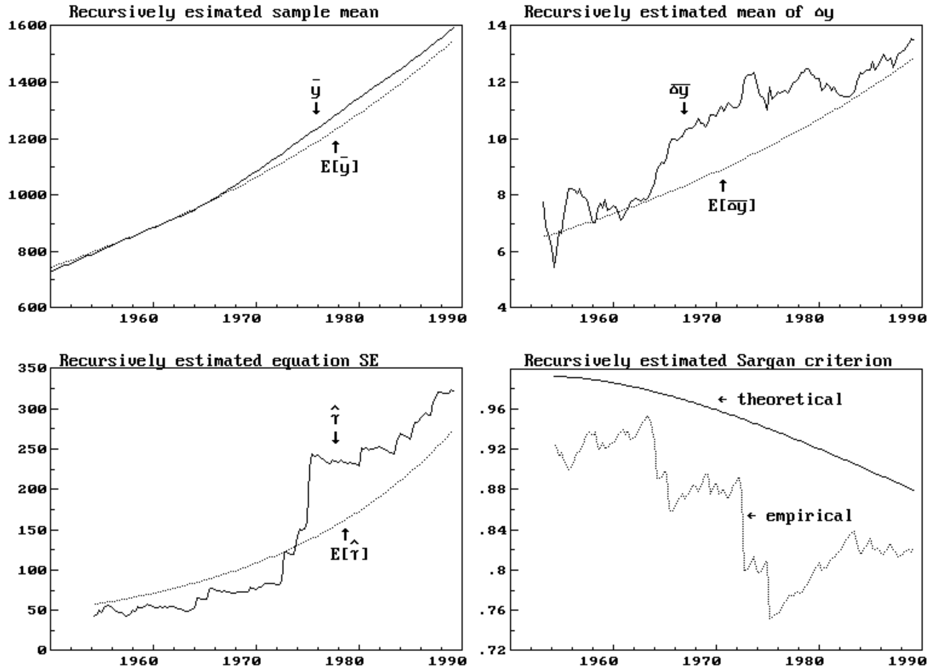

Fig. 7 Predicted moments for US total disposable income

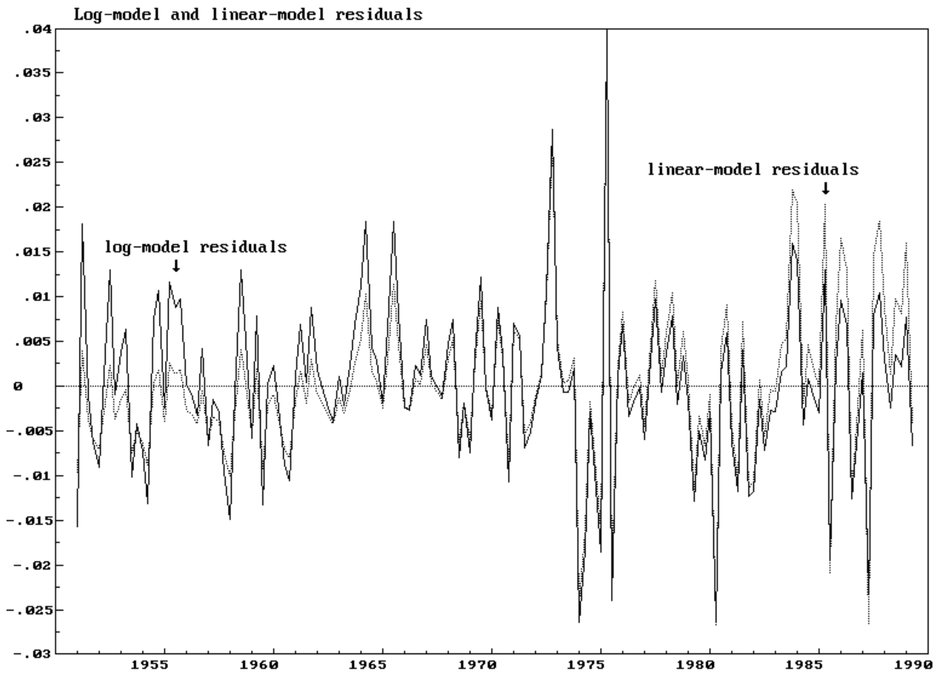

Fig. 8 Residuals of the linear and log models

The residuals of the two models are plotted jointly in FIG. 8 to reveal the heteroskedasticity of the linear representation.