MONEY DEMAND FUNCTION ESTIMATION BY NONLINEAR COINTEGRATION

YOUNGSOO BAE[†] AND ROBERT M. DE JONG[‡]

Department of Economics, Ohio State University, USA

August 2004

SUMMARY

Conventionally, the money demand function is estimated using a regression of the logarithm of money demand on either the interest rate or the logarithm of the interest rate. This equation is presumed to be a cointegrating regression. In this paper, we aim to combine the logarithmic specification, which models the liquidity trap better than a linear model, with the assumption that the interest rate itself is an integrated process. The proposed technique is robust to serial correlation in the errors. For the US, our new technique results in larger coefficient estimates than previous research suggested, and produces superior out-of-sample prediction.

JEL Classification: E41; C22

Keywords: Liquidity Trap, Money Demand, Nonlinear Cointegration

---

[†] Correspondence to: 410 Arps Hall, 1945 North High Street, Columbus OH, 43210, USA. Email: bae.35@osu.edu. Telephone: 614-292-5765. Fax: 614-292-3906. I would like to thank Masao Ogaki for invaluable advices and encouragement.

[‡] Email: dejong@ecolan.sbs.ohio-state.edu

1 INTRODUCTION

1.1 Aim of This Paper

In the literature on the estimation of the long-run money demand function, two functional forms are typically used. Those functional forms are

$$m_t = \beta_0 + \beta_1 r_t + u_t \tag{1}$$

and

$$m_t = \beta_0 + \beta_1 \ln(r_t) + u_t, \tag{2}$$

where m_t denotes the logarithm of real money demand and r_t is the nominal interest rate. These functional forms have been used in previous studies of the long-run money demand function estimation. For example, see Lucas (1988), Hoffman and Rasche (1991), Stock and Watson (1993), Ball (1998), and Anderson and Rasche (2002).

It is believed by many researchers that the nominal interest rate is an integrated process. This assumption was made in e.g. Stock and Watson (1993), Ball (1998), Anderson and Rasche (2001), and Hu and Phillips (2002). See Hu and Phillips (2002, p.5) for an argument for nonstationarity in the nominal interest rate. In addition, this assumption might have some advantages over the alternative assumption that the logarithm of the nominal interest rate is an integrated process. In the latter assumption, the nominal interest rate is an exponential function of an integrated process, which implies that its percentage change has a stationary distribution. This might be an appropriate assumption for some macroeconomic variables, such as GDP and CPI. For the nominal interest rate, however, if we consider the fact that the Federal Reserve Board (FRB) usually adjusts its target interest rate by multiples of 25 basis points, not by certain percentage of the current interest rate level, it might be more appropriate to assume that its difference has a stationary distribution, which in turn implies that the nominal interest rate itself is an integrated process.

For the estimation and the statistical inference on the money demand function, cointegration methods have been used. However, to use conventional linear cointegration methods, such as Phillips and Hansen (1992)'s "Fully Modified OLS" (FMOLS) and Stock and Watson (1993)'s "Dynamic OLS" (DOLS), we need to make different assumptions for different functional forms; that is, for Equation (1), r_t must be assumed to be an integrated process, and for Equation (2), $\ln(r_t)$ must be assumed to be an integrated process. The approach in this paper seeks to estimate the functional form of Equation (2) under the assumption that the interest rate is an integrated process; that is,

we aim to reconcile the assumption that interest rate is an integrated process with the logarithmic functional form. To this end, we will develop a new nonlinear "fully modified" type estimation technique.

1.2   Functional Forms and Money Demand

In the literature on money demand, the choice of functional form has important implications for the presence of the liquidity trap. The liquidity trap is the phenomenon where the public's money demand becomes indefinite at a low interest rate; that is, the public is willing to hold any amount of money at a low interest rate, because it is indifferent between money and other financial assets.[1] In the functional form of Equation (1), there is no liquidity trap, while the liquidity trap exists in the functional form of Equation (2). This is because in the latter specification, the money demand increases into infinity as the interest rate approaches to zero. Laidler (1985, p.53) argued that *"[Keynes's analysis]... suggest that it [money demand] cannot be treated as a simple, stable, approximately linear, negative relationship with respect to the rate of interest."* The presence of the liquidity trap has important implications for monetary policy. As Krugman (1998) pointed out,[2] the liquidity trap makes the traditional monetary policy impotent when interest rate is close to zero. Furthermore, the possibility of the liquidity trap makes a central bank's optimal policy different even at the normal range of the interest rate. Nishiyama (2003) showed that in the presence of the liquidity trap the optimal monetary policy is to set a small but positive inflation rate as its target, to avoid the possibility of falling into the liquidity trap. This is because a central bank cannot set a negative nominal interest rate if there is the liquidity trap. For the Japanese economy, the presence of the liquidity trap has been an important issue for a decade.

   In addition, the specification of Equation (2) has two more advantages over the specification of Equation (1). First, macroeconomic data show a clear nonlinear relationship between money demand and the interest rate, as is illustrated in Figure 1. On the other hand, a graph of the relationship between money demand and the logarithm of the interest rate appears to be linear. Second, as argued by Anderson and Rasche (2001), the elasticity of money demand with respect to the interest rate should not be an increasing function of the interest rate, but it should be a decreasing or at least

---

[1] Laidler (1988, p.53) mentioned that "This is the doctrine of the liquidity trap, which argues that the interest elasticity of the demand for money can, at low levels of the rate of interest, take the value infinity."

[2] Krugman (1998) mentioned that "A liquidity trap may be defined as a situation in which conventional monetary policies have become impotent, because nominal interest rates are at or near zero."

a non-increasing function. This is because as the public reduces its money holdings, each successive reduction should not be less difficult. While the interest elasticity of money demand is an increasing function of the interest rate in the specification of Equation (1), it is constant in the specification of Equation (2). Considering all these aspects, we may conclude that the specification of Equation (2) is more appropriate than the specification of Equation (1).

1.3   Linear Cointegration Methods

The use of standard linear cointegration methods, such as Phillips and Hansen (1992)'s "Fully Modified OLS" (FMOLS) and Stock and Watson (1993)'s "Dynamic OLS" (DOLS), for the estimation of the functional form of Equation (2) implies the assumption that $\ln(r_t)$ is an integrated process. However, if r_t is an integrated process, then the logarithm of r_t is not an integrated process in any meaningful sense, and vice versa. Because of this problem, linear cointegration methods do not provide a consistent framework to estimate both Equation (1) and (2) under the assumption that the interest rate is an integrated process. Recently there have been new developments of nonlinear cointegration methods; notably, Park and Phillips (1999, 2001), Chang et al. (2001), and de Jong (2002). However, currently available nonlinear cointegration methods are not suitable for the analysis of the model of Equation (2). The techniques of Park and Phillips (1999, 2001) and Chang et al. (2001) can be used in principle, but only if we assume that the error u_t is a martingale difference sequence. Because of the presence of serial correlation and possible endogeneity, this assumption is not acceptable for estimation of the long-run money demand function estimation. De Jong (2002) relaxes the martingale difference sequence assumption, however, his technique cannot be used for the specification of Equation (2) because of the unboundedness of the logarithm function near zero.

This paper develops a new nonlinear "fully modified" type cointegration method that can be used in the functional form of Equation (2) under the assumption that r_t is an integrated process. Of course, the assumption of a logarithmic specification in combination with the presence of a unit root in r_t will give rise to the inner inconsistency that integrated processes can take on negative values. This is not a problem only for Equation (2). In Equation (1), the interest rate cannot be an integrated process, because it is always positive. This is true for some other macroeconomic variables as well, such as the unemployment rate. Therefore, in the sequel of this paper it is presumed that Equation (2) is approximated by a logarithm function of the absolute value of the interest rate. Since asymptotic theory provides us with nothing more than approximations to true coefficient distributions, this paper should be viewed as providing a different means of conducting inference in

the money demand function of Equation (2), using a different approximation of the limit distributions of coefficients.

The paper is organized as follows. In Section 2, a new nonlinear cointegration method is proposed and its asymptotic properties are established. In Section 3, the US long-run money demand function is estimated using both the newly proposed technique and more conventional techniques, and the estimation results are presented. The conclusions can be found in Section 4.

2   NONLINEAR COINTEGRATING REGRESSION

2.1   Model and Assumptions

In this paper, the nonlinear cointegrating regression

$$y_t = \beta_0 + \beta_1 \ln |x_t| + u_t \tag{3}$$

is considered, where x_t is an integrated process and u_t is a stationary process. Note that Equation (3) is different from Equation (2); i.e. the absolute value of an integrated process $|x_t|$ goes into a logarithm function in Equation (3), not x_t. Its estimation and statistical inference are not simple tasks. To illustrate analytical difficulties that Equation (3) poses, we investigate the properties of the OLS estimator. The OLS estimator $\overline{\beta}_1$ satisfies

$$\sqrt{n}(\overline{\beta}_1 - \beta_1) = \frac{\frac{1}{\sqrt{n}} \sum_{t=1}^n (\ln |x_t| - \overline{\ln |x_t|}) u_t}{\frac{1}{n} \sum_{t=1}^n (\ln |x_t| - \overline{\ln |x_t|})^2} = \frac{\frac{1}{\sqrt{n}} \sum_{t=1}^n (\ln \left|\frac{x_t}{\sqrt{n}}\right| - \overline{\ln \left|\frac{x_t}{\sqrt{n}}\right|}) u_t}{\frac{1}{n} \sum_{t=1}^n (\ln \left|\frac{x_t}{\sqrt{n}}\right| - \overline{\ln \left|\frac{x_t}{\sqrt{n}}\right|})^2},$$

where $\overline{X_t}$ indicates the sample average of X_t. Note that if u_t is a martingale difference sequence as in Park and Phillips (2001), it can be shown that $\sqrt{n}(\overline{\beta}_1 - \beta_1)$ is $O_p(1)$. However, if u_t has serial correlation, and u_t and Δx_t are correlated with each other, de Jong (2002) has shown that for continuously differentiable function $T(\cdot)$,

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n T(\frac{x_t}{\sqrt{n}}) u_t \overset{d}{\longrightarrow} \int_0^1 T(W(r)) dU(r) + \Lambda \int_0^1 T'(W(r)) dr$$

where $W(r)$ and $U(r)$ are the limiting Brownian processes associated with $(\Delta x_t,\ u_t)$, and Λ is the correlation parameter. Since $\int_0^1 \frac{1}{|W(r)|} dr$ is undefined, it can be conjectured that if $\Lambda \neq 0$, $\frac{1}{\sqrt{n}} \sum_{t=1}^n \ln \left|\frac{x_t}{\sqrt{n}}\right| u_t$ has no well-defined limit, and therefore $\sqrt{n}(\overline{\beta}_1 - \beta_1)$ does not converge in distribution.

Throughout the paper, the following assumptions are maintained.

5

Assumption 1 *Let* $\Delta x_t = w_t$ *with* $x_0 = O_p(1)$. u_t *and* w_t *are linear processes*

$$u_t = \sum_{i=0}^{\infty} \phi_{1,i} \varepsilon_{1,t-i}$$

$$w_t = \sum_{i=0}^{\infty} \phi_{2,i} \varepsilon_{2,t-i},$$

where $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})$ *is a sequence of independent and identically distributed (i.i.d.) random variables with mean zero. The long-run covariance matrix of* $\eta_t = (u_t, w_t)$, Ω, *is given by*

$$\Omega = C_0 + \sum_{j=1}^{\infty} (C_j + C_j') = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}$$

where $C_j = E(\eta_t \eta_{t-j}')$. Ω_{22} *is nonsingular.*

Note that u_t and w_t are allowed to be weakly dependent and correlated with each other.

Assumption 2 $\sum_{i=0}^{\infty} i |\phi_{j,i}| < \infty$ *and* $E|\varepsilon_{j,t}|^p < \infty$ *for some* $p > 2$ *for* $j = 1, 2$.

Assumption 3 *The distribution of* $(\varepsilon_{1,t}, \varepsilon_{2,t})$ *is absolutely continuous with respect to the Lebesgue measure and has characteristic function* $\psi(s)$ *for which* $\lim_{|s| \to \infty} |s|^q \psi(|s|) = 0$ *for some* $q > 1$.

Under Assumption 1, 2 and 3, $\frac{1}{\sqrt{t}} \sum_{i=1}^{t} w_i$ has a density $f_t(\cdot)$ that is uniformly bounded over t. Also, define

$$U_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} u_t$$

$$W_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} w_t,$$

where $[s]$ denotes the largest integer not exceeding s. Then we can assume without loss of generality that by the Skorokhod representation theorem,

$$\sup_{r \in [0,1]} |W_n(r) - W(r)| = o(1)$$

and

$$\sup_{r \in [0,1]} |U_n(r) - U(r)| = o(\delta_n),$$

where δ_n is a deterministic sequence that is $o(n^{-\frac{p-2}{2p}})$; see Park and Phillips (1999, p.271).

6

2.2   Estimation

In this section, a "Nonlinear Cointegration Least Squares" (NCLS) estimator $\widetilde{\beta}_1$ is defined and its asymptotic properties are established. This estimator solves the problems associated with the unboundedness of the logarithm function near zero that make the OLS estimator intractable for analysis. Let k_n be an integer-valued positive sequence such that $k_n \to \infty$ and $k_n n^{-\frac{p-2}{3p+\eta}} \to 0$ for some $\eta > 0$, and $n_j = \left[\frac{nj}{k_n}\right]$ for $j = 0, 1, 2, \ldots, k_n$. Let $z_t = \ln |x_{n_{j-1}+1}|$ for $n_{j-1} + 1 \le t \le n_j$ for $j = 1, 2, \ldots, k_n$. Then the NCLS estimator $\widetilde{\beta}_1$ is

$$\widetilde{\beta}_1 = \frac{\sum_{t=1}^n z_t (y_t - \overline{y_t})}{\sum_{t=1}^n z_t (\ln |x_t| - \overline{\ln |x_t|})} = \frac{\sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln |x_{n_{j-1}+1}| (y_t - \overline{y_t})}{\sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln |x_{n_{j-1}+1}| (\ln |x_t| - \overline{\ln |x_t|})}.$$

This is an IV estimator that uses z_t as the instrumental variable for $\ln |x_t|$. Note that $k_n \to \infty$ at a slower rate than n, and $k_n = o(n^{1/3})$ when $p = \infty$.

Two technical results that are needed in the sequel are stated below.

Theorem 1 *Under Assumption 1, 2 and 3,*

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln \left| \frac{x_{n_{j-1}+1}}{\sqrt{n}} \right| u_t \xrightarrow{d} \int_0^1 \ln |W(r)| dU(r).$$

Theorem 2 *Under Assumption 1, 2 and 3,*

$$\frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln \left| \frac{x_{n_{j-1}+1}}{\sqrt{n}} \right| \ln \left| \frac{x_t}{\sqrt{n}} \right| \xrightarrow{d} \int_0^1 [\ln |W(r)|]^2 dr.$$

Based on Theorem 1 and 2, the asymptotic distribution of $\widetilde{\beta}_1$ is obtained in the following theorem.

Theorem 3 *Under Assumption 1, 2 and 3,*

$$\sqrt{n}(\widetilde{\beta}_1 - \beta_1) \xrightarrow{d} \frac{\int_0^1 \ln |W(r)| dU(r) - U(1) \int_0^1 \ln |W(r)| dr}{\int_0^1 [\ln |W(r)|]^2 dr - \left[\int_0^1 \ln |W(r)| dr \right]^2}.$$

The result of Theorem 3 is reminiscent of the limit theory for the least squares estimator in a linear cointegration regression. In both cases, the result holds without exogeneity assumption of any kind, and the limit distribution is not directly suitable for inference in some empirical settings. If $W(\cdot)$ and

$U(\cdot)$ are orthogonal, then the usual "t- and F-statistics" are valid because they achieve the correct significance level conditionally on $W(\cdot)$. However, if they are not orthogonal, which is likely in the long-run money demand function case, this result cannot be used for the statistical inference, since the above limit distribution depends on the correlation between $U(\cdot)$ and $W(\cdot)$.

2.3 Fully Modified Type Estimation of the NCLS Estimator

In this section, a "fully modified" type estimation procedure for the NCLS estimator is proposed. A modification is needed to establish the asymptotic conditional normality. The regression Equation (3) can be written as

$$y_t^\dagger = \beta_0 + \beta_1 \ln |x_t| + u_t^\dagger, \tag{4}$$

where $y_t^\dagger = y_t - \Omega_{21}' \Omega_{22}^{-1} \Delta x_t$ and $u_t^\dagger = u_t - \Omega_{21}' \Omega_{22}^{-1} \Delta x_t$. Note that now the limiting Brownian process associated with $u_t^\dagger$, $U^\dagger(\cdot)$, and $W(\cdot)$ are orthogonal.[3]

The proposed estimation procedure is now as follows.

1. Calculate the residual, $\widehat{u}_t$, from a regression of y_t on an intercept and $\ln |x_t|$ by the NCLS estimation method.

2. Get a HAC estimate $\widehat{\Omega}$ by using $(\widehat{u}_t, \Delta x_t)$.

3. Calculate $\widehat{y}_t^\dagger$ in a way analogous to FMOLS,

$$\widehat{y}_t^\dagger = y_t - \widehat{\Omega}_{21}' \widehat{\Omega}_{22}^{-1} \Delta x_t.$$

4. The "fully modified" NCLS estimator $\widehat{\beta}_1$ is defined as the NCLS estimator that is calculated using the modified dependent variable $\widehat{y}_t^\dagger$, instead of y_t, i.e.

$$\widehat{\beta}_1 = \frac{\sum_{t=1}^{n} z_t (\widehat{y}_t^\dagger - \overline{\widehat{y}_t^\dagger})}{\sum_{t=1}^{n} z_t (\ln |x_t| - \overline{\ln |x_t|})} = \frac{\sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln |x_{n_{j-1}+1}| (\widehat{y}_t^\dagger - \overline{\widehat{y}_t^\dagger})}{\sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln |x_{n_{j-1}+1}| (\ln |x_t| - \overline{\ln |x_t|})}.$$

---

[3] This follows because

$$\lim_{n \to \infty} E \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{[rn]} u_t^\dagger \right) \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{[rn]} \Delta x_t \right)$$

$$= \lim_{n \to \infty} E \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{[rn]} u_t \right) \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{[rn]} \Delta x_t \right) - \Omega_{21}' \Omega_{22}^{-1} E \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{[rn]} \Delta x_t \right)^2$$

$$= \Omega_{21}' - \Omega_{21}' \Omega_{22}^{-1} \Omega_{22} = 0.$$

To establish the asymptotic distribution of $\widehat{\beta}_1$, the following additional assumption is needed.

Assumption 4 *Let $k(\cdot)$ and $\{b_n : n \geq 1\}$ be a kernel function and a sequence of bandwidth parameters that are used for the HAC estimator $\widehat{\Omega}$ in step 2.*

(i) $k(0) = 1, k(\cdot)$ is continuous at zero, and $\sup_{x \geq 0}|k(x)| < \infty$.

(ii) $\int_0^\infty \overline{k}(x)dx < \infty$, where $\overline{k}(x) = \sup_{y \geq x}|k(y)|$.

(iii) $b_n \subseteq (0, \infty)$ and $\lim_{n \to \infty} \left(\dfrac{1}{b_n} + \dfrac{b_n}{\sqrt{n}} \right) = 0$.

Assumption 4 is Assumptions A3 and A4 in Jansson (2002, p.1450). Note that Assumption 4 (i) and (ii) hold for the 15 kernels, including the Bartlett kernel, studied by Ng and Perron (1996), and Assumption 4 (iii) holds whenever the bandwidth expansion rate coincides with the optimal rate in Andrews (1991); see Jansson (2002, p. 1450-1451).

The following results establishes the consistency of the HAC estimator $\widehat{\Omega}$.

Theorem 4 *Let $\widehat{\Omega}$ be the HAC estimator in Step 2. Under Assumption 1, 2, 3 and 4, $\widehat{\Omega} \xrightarrow{p} \Omega$.*

The asymptotic distribution of the "fully modified" NCLS estimator $\widehat{\beta}_1$ is obtained in the following theorem.

Theorem 5 *Under Assumption 1, 2, 3 and 4,*

$$\sqrt{n}(\widehat{\beta}_1 - \beta_1) \xrightarrow{d} \frac{\displaystyle\int_0^1 \ln|W(r)|dU^\dagger(r) - U^\dagger(1) \int_0^1 \ln|W(r)|dr}{\displaystyle\int_0^1 [\ln|W(r)|]^2 dr - \left[\int_0^1 \ln|W(r)|dr\right]^2}.$$

Since $W(r)$ and $U^\dagger(r)$ are orthogonal in Theorem 5, it follows that conditionally on $W(\cdot)$,

$$\sqrt{n}(\widehat{\beta}_1 - \beta_1) \xrightarrow{d} N\left(0, \frac{\Omega_{11}^\dagger}{\displaystyle\int_0^1 [\ln|W(r)|]^2 dr - \left[\int_0^1 \ln|W(r)|dr\right]^2}\right),$$

where $\Omega_{11}^\dagger$ is the long-run variance of $u_t^\dagger$. This implies that the usual "t- and F-statistics" are valid. For estimation of $\Omega_{11}^\dagger$, a consistent estimator $\widehat{\Omega}_{11}^\dagger$ is provided by

$$\widehat{\Omega}_{11}^\dagger = \widehat{\Omega}_{11} - \widehat{\Omega}_{21}'\widehat{\Omega}_{22}^{-1}\widehat{\Omega}_{21}.$$

Note that since $\widehat{\Omega} \xrightarrow{p} \Omega$ by Theorem 4, $\widehat{\Omega}_{11}^\dagger$ is a consistent estimator for $\Omega_{11}^\dagger$.

3   US MONEY DEMAND

In this section, the US long-run money demand functions of Equation (1) and (2) are estimated by both linear and nonlinear cointegration methods. Specifically, we use "Static OLS" (SOLS), DOLS, FMOLS, and NCLS estimation techniques.

3.1   Unit Root Test Results

We first conduct unit root tests for the nominal interest rate and the logarithm of the nominal interest rate by both the Augmented Dickey-Fuller (ADF) test and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test (Kwiatkowski et al. 1992). Test results are presented in Table 1. The ADF test results show that the null hypotheses that r_t and $\ln(r_t)$ are integrated processes without drift cannot be rejected, and the KPSS test results show that the null hypotheses that r_t and $\ln(r_t)$ are stationary processes are rejected. Therefore, both tests confirm the hypotheses that r_t and $\ln(r_t)$ are integrated processes.

3.2   Estimation Results

The same data set as Ball (1998), Stock and Watson (1993) and Lucas (1978) is used, however, it was extended up to 1997.[4] M_1[5], Net National Products (NNP), NNP deflator, and 6-months commercial paper rate are used as money, output, price, and interest rate, respectively.

Coefficient estimates and their asymptotic standard deviations are presented in Table 2. When the estimation results are compared between linear cointegration methods and the NCLS method in the functional form of Equation (3), the NCLS estimates are larger in absolute value than all other linear cointegration estimates when the bandwidth parameter k_n is small. Note that the theory for the NCLS estimator is valid for $k_n = o(n^{\frac{1}{3}})$ at best, which suggests that a small value for k_n may be appropriate. Also, note that for $k_n = 5$, all NCLS point estimates are outside all the 95% confidence intervals suggested by the SOLS, DOLS and FMOLS estimates.

To further address the question which functional form is most appropriate and which estimation technique is to be preferred, we investigate out-of-sample prediction performance in terms of sum of squared forecast errors. Results are presented in Table 3 and 4. Between the functional

---

[4] The reason data extension is stopped in 1997 is that 6-months commercial paper rate was discontinued at that point.

[5] After 1995, the official M1 statistics excludes the Sweep Account. As shown by Dutkowsky and Cynamon (2003), this might have an impact on cointegration relationship. Therefore, we added the Sweep Account to M1 from 1994.

forms of Equation (1) and (3), Equation (3) outperforms Equation (1) in all combinations of different estimation methods and prediction methods. Within the functional form of Equation (3), the NCLS estimator produces superior prediction performance compared to other conventional linear cointegration estimators. The superior out-of-sample prediction performance of the NCLS estimator in combination with the logarithmic specification suggests that the NCLS technique produces the most desirable estimation results, and this reinforces our belief that the conventional techniques may underestimate the impact of the interest rate on money demand.

4   CONCLUSION

This paper investigates two different functional forms of the US long-run money demand function by linear and nonlinear cointegration methods. Since different functional forms have different implications for the presence of the liquidity trap and effectiveness of the traditional monetary policy, the choice of functional form is a important issue. Therefore, its estimation should be conducted in a consistent framework so that estimation results can be comparable. Since a logarithm function involves a nonlinear relationship between money demand and the interest rate, nonlinear cointegration methods are more appropriate methods. This paper proposes a new nonlinear cointegration method that is suitable for a more general time series situation, such as the long-run money demand function estimation, where serial correlation in the errors is present. Estimation results show that the interest rate coefficient estimates are larger in absolute value for the NCLS estimator than for other conventional linear cointegration estimators. Also, out-of-sample prediction performances for our NCLS technique are superior to those of the conventional linear estimation techniques. Therefore, our conclusion is that the conventional techniques may underestimate the impact of the interest rate on money demand.

5   Appendix

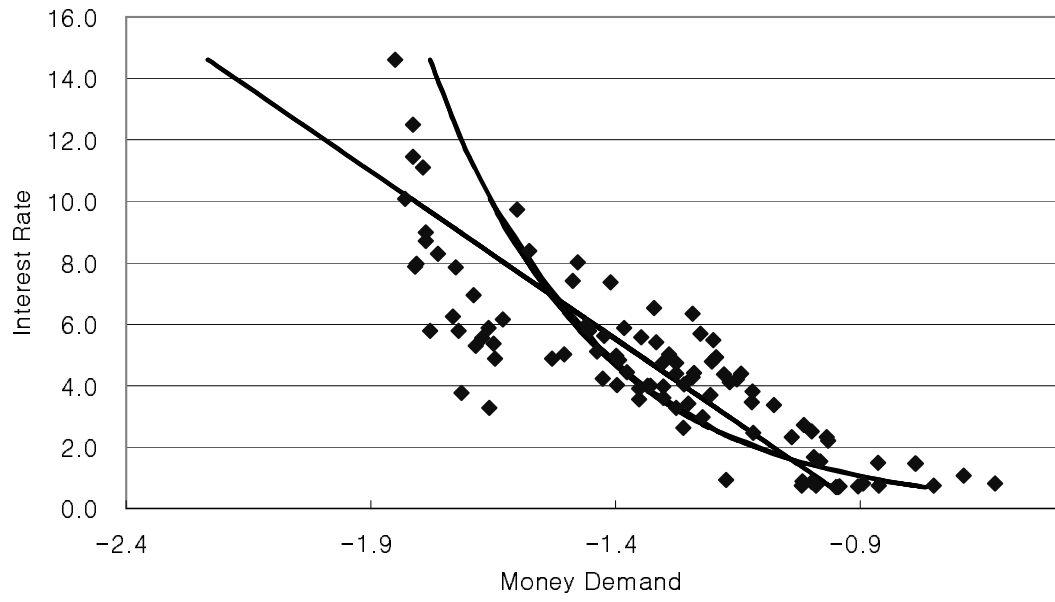## 5.1   Estimation Results

Figure 1: Money Demand - Linear vs. Logarithm

Table 1: Unit Root Test for Interest Rate[1]

	r_t	Δr_t	$\ln(r_t)$	$\Delta \ln(r_t)$
ADF test[2]	-0.07	-1.39**	-0.05	-1.04**
KPSS test[3]	0.64*	0.07	0.50*	0.13

1) Reject the null hypothesis at 5%(*) and 1%(**).

2) H_0: I(1) vs. H_1: I(0). The order of the first difference terms is 4.

3) H_0: I(0) vs. H_1: I(1).

Table 2: Estimation[1] Results of Interest Rate Coefficient β_1

sample period	SOLS	DOLS[2]	FMOLS	NCLS ($k_n = 20$)	NCLS ($k_n = 10$)	NCLS ($k_n = 5$)
	$m_t = \beta_0 + \beta_1 r_t + u_t$					
1900-1997	-0.0914	-0.1138	-0.1050	–	–	–
	(0.0100)	(0.0074)	(0.0097)			
1900-1945	-0.0751	-0.0824	-0.0807	–	–	–
	(0.0114)	(0.0131)	(0.0114)			
1946-1997	-0.0910	-0.1054	-0.1045	–	–	–
	(0.0149)	(0.0097)	(0.0138)			
	$m_t = \beta_0 + \beta_1 \ln(r_t) + u_t$					
1900-1997	-0.3326	-0.3793	-0.3557	-0.3548	-0.4165	-0.5254
	(0.0400)	(0.0352)	(0.0383)	(0.0392)	(0.0414)	(0.0522)
1900-1945	-0.1869	-0.1898	-0.1895	-0.1902	-0.2276	-0.2659
	(0.0247)	(0.0283)	(0.0251)	(0.0251)	(0.0278)	(0.0339)
1946-1997	-0.4863	-0.5140	-0.5089	-0.5196	-0.5908	-0.6212
	(0.0474)	(0.0397)	(0.0447)	(0.0463)	(0.0485)	(0.0514)

1) An HAC estimator with Bartlett kernel and bandwidth parameter of 4 are used.

2) The order of the leads and lags is 4.

Table 3: Out-of-sample Prediction Performance Results using estimates from 1900-1989[1]

	SOLS	DOLS	FMOLS	NCLS ($k_n = 5$)
	$m_t = \beta_0 + \beta_1 r_t + u_t$			
$\sum_{t=1990}^{1997} [m_t - \widehat{m}_t]^2$	1.0439	0.9529	1.0191	–
	$m_t = \beta_0 + \beta_1 \ln(r_t) + u_t$			
$\sum_{t=1990}^{1997} [m_t - \widehat{m}_t]^2$	0.7411	0.6486	0.7056	0.5400

1) The same kernel, bandwidth and order of leads and lags as in Table 2 are used.

Table 4: Out-of-sample One-step-ahead Prediction Performance Results[1]

	SOLS	DOLS	FMOLS	NCLS ($k_n = 5$)
	$m_t = \beta_0 + \beta_1 r_t + u_t$			
$\sum_{t=1994}^{1997} [m_t - \widehat{m}_t]^2$	0.3021	0.2774	0.2790	–
$\sum_{t=1993}^{1997} [m_t - \widehat{m}_t]^2$	0.5319	0.5345	0.5246	–
$\sum_{t=1992}^{1997} [m_t - \widehat{m}_t]^2$	0.7799	0.7963	0.7825	–
$\sum_{t=1991}^{1997} [m_t - \widehat{m}_t]^2$	0.9242	0.9202	0.9165	–
$\sum_{t=1990}^{1997} [m_t - \widehat{m}_t]^2$	0.9758	0.9423	0.9517	–
	$m_t = \beta_0 + \beta_1 \ln(r_t) + u_t$			
$\sum_{t=1994}^{1997} [m_t - \widehat{m}_t]^2$	0.1864	0.1567	0.1693	0.0834
$\sum_{t=1993}^{1997} [m_t - \widehat{m}_t]^2$	0.3377	0.3096	0.3213	0.2452
$\sum_{t=1992}^{1997} [m_t - \widehat{m}_t]^2$	0.5017	0.4689	0.4838	0.4025
$\sum_{t=1991}^{1997} [m_t - \widehat{m}_t]^2$	0.6121	0.5632	0.5871	0.4689
$\sum_{t=1990}^{1997} [m_t - \widehat{m}_t]^2$	0.6846	0.6165	0.6506	0.4906

1) The same kernel, bandwidth and order of leads and lags as in Table 2 are used.

5.2   Proofs

Proof of Theorem 1: Let k_n and n_j be same as in Section 2.2. For notational simplicity, let $v_{t+1} = u_t$ for $t = 0, 1, 2, \ldots, n$, and define $V_n(r) = \frac{1}{\sqrt{n}} \sum_{t=0}^{[nr]} v_{t+1}$. Let $r_j = \frac{n_j+1}{n}$ for $j = 0, 1, 2, \ldots, k_n$. Consider

$$G_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln \left| \frac{x_{n_{j-1}+1}}{\sqrt{n}} \right| u_t = \frac{1}{\sqrt{n}} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln \left| \frac{x_{n_{j-1}+1}}{\sqrt{n}} \right| v_{t+1}$$

$$= \sum_{j=1}^{k_n} \ln |W_n(r_{j-1})| \left(V_n(r_j) - V_n(r_{j-1}) \right).$$

Then G_n can be written as

$$G_n = (G_n - P_n) + (P_n - I) + I$$

where

$$P_n = \sum_{j=1}^{k_n} \ln |W(r_{j-1})| (V(r_j) - V(r_{j-1})) = \sum_{j=1}^{k_n} \int_{r_{j-1}}^{r_j} \ln |W(r_{j-1})| dV(r)$$

and

$$I = \int_0^1 \ln |W(r)| dV(r) = \int_0^1 \ln |W(r)| dU(r).$$

Then the theorem can be proved by showing that

$$|G_n - P_n| \xrightarrow{p} 0$$

and

$$|P_n - I| \xrightarrow{p} 0.$$

First, by summation by parts as in Davidson (1994, p. 512), $G_n - P_n$ can be written as

$$G_n - P_n = \sum_{j=1}^{k_n} (\ln |W_n(r_{j-1})| - \ln |W(r_{j-1})|)(V_n(r_j) - V_n(r_{j-1})) \tag{1}$$

$$+ \ln |W(r_{k_n})|(V_n(r_{k_n}) - V(r_{k_n})) \tag{2}$$

$$- \sum_{j=1}^{k_n} (V_n(r_j) - V(r_j))(\ln |W(r_j)| - \ln |W(r_{j-1})|). \tag{3}$$

It will be shown that all three terms converge to 0 in probability. For (1),

$$\left[\sum_{j=1}^{k_n} (\ln |W_n(r_{j-1})| - \ln |W(r_{j-1})|)(V_n(r_j) - V_n(r_{j-1})) \right]^2$$

$$\leq \sum_{j=1}^{k_n} [\ln |W_n(r_{j-1})| - \ln |W(r_{j-1})|]^2 \sum_{j=1}^{k_n} [V_n(r_j) - V_n(r_{j-1})]^2.$$

15

Now note that

$$E\sum_{j=1}^{k_n}[V_n(r_j) - V_n(r_{j-1})]^2 = O(1)$$

because Assumption 1, 2 and 3 imply that $E[V_n(r_j) - V_n(r_{j-1})]^2 \leq C \cdot (r_j - r_{j-1})$. Therefore it suffices to show that

$$\sum_{j=1}^{k_n}[\ln|W_n(r_{j-1})| - \ln|W(r_{j-1})|]^2 = o_p(1).$$

Let $T_\varepsilon(x) = \ln|x|I(|x| > \varepsilon) + \ln(\varepsilon)I(|x| \leq \varepsilon)$, $Q_\varepsilon(x) = (\ln(\varepsilon) - \ln|x|)I(|x| < \varepsilon)$, and $\varepsilon_n = n^{-\frac{p-2}{3p}}$. First note that since $(a+b)^2 \leq 2a^2 + 2b^2$, it follows that

$$(a+b+c)^2 \leq 2a^2 + 2(b+c)^2 \leq 2a^2 + 4b^2 + 4c^2.$$

Then by letting $a = T_{\varepsilon_n}(W_n) - T_{\varepsilon_n}(W)$, $b = Q_{\varepsilon_n}(W_n)$, and $c = Q_{\varepsilon_n}(W)$, we have the following,

$$\sum_{j=1}^{k_n}[\ln|W_n(r_{j-1})| - \ln|W(r_{j-1})|]^2$$

$$\leq 2\sum_{j=1}^{k_n}[T_{\varepsilon_n}(W_n(r_{j-1})) - T_{\varepsilon_n}(W(r_{j-1}))]^2 + 4\sum_{j=1}^{k_n}Q_{\varepsilon_n}(W_n(r_{j-1}))^2 + 4\sum_{j=1}^{k_n}Q_{\varepsilon_n}(W(r_{j-1}))^2.$$

Since $T_{\varepsilon_n}(\cdot)$ is Lipschitz-continuous with Lipschitz-coefficient $\frac{1}{\varepsilon_n}$, it follows that

$$\sum_{j=1}^{k_n}[T_{\varepsilon_n}(W_n(r_{j-1})) - T_{\varepsilon_n}(W(r_{j-1}))]^2 \leq k_n\left(\frac{\delta_n}{\varepsilon_n}\right)^2 = o_p(n^{\frac{p-2}{3p+\eta}} \cdot n^{-\frac{p-2}{p}} \cdot n^{\frac{2(p-2)}{3p}})$$

$$= o_p(n^{\frac{p-2}{3p+\eta} - \frac{p-2}{3p}}) = o_p(1).$$

Recall that $f_t(\cdot)$ denotes the density of $\frac{x_t}{\sqrt{t}}$ and that under Assumption 1, 2 and 3, $\sup_t \sup_{y\in\mathbb{R}} f_t(y) < \infty$. Then note that by using the substitution $y = \frac{\sqrt{n}}{\sqrt{n_{j-1}+1}}W_n(r_{j-1})$,

$$E\sum_{j=1}^{k_n}Q_{\varepsilon_n}(W_n(r_{j-1}))^2 = \sum_{j=1}^{k_n}\int_{-\infty}^{\infty}Q_{\varepsilon_n}(\frac{\sqrt{n_{j-1}+1}}{\sqrt{n}}y)^2 f_{n_{j-1}+1}(y)dy$$

$$\leq \left[\sup_t \sup_{y\in\mathbb{R}} f_t(y)\right]\sum_{j=1}^{k_n}\int_{-\infty}^{\infty}Q_{\varepsilon_n}(x)^2\frac{\sqrt{n}}{\sqrt{n_{j-1}+1}}dx$$

$$= \left[\sup_t \sup_{y\in\mathbb{R}} f_t(y)\right]\left[\sum_{j=1}^{k_n}\frac{\sqrt{n}}{\sqrt{n_{j-1}+1}}\right]\left[\int_{-\infty}^{\infty}Q_{\varepsilon_n}(x)^2dx\right]$$

$$= O(k_n\varepsilon_n) = O(n^{\frac{p-2}{3p+\eta} - \frac{p-2}{3p}}) = o(1).$$

A similar argument holds for $\sum_{j=1}^{k_n}Q_{\varepsilon_n}(W(r_{j-1}))^2$. For (2), note that

$$\ln|W(r_{k_n})|(V_n(r_{k_n}) - V(r_{k_n})) = O_p(\delta_n) = o_p(1)$$

16

where δ_n is a deterministic sequence that is $o(n^{-\frac{p-2}{2p}})$. A similar argument holds for (3). Therefore we have established that $|G_n - P_n| \xrightarrow{p} 0$. Next, $P_n - I$ can be written as

$$P_n - I = (P_n - P_{n\varepsilon}) + (P_{n\varepsilon} - I_{n\varepsilon}) + (I_{n\varepsilon} - I)$$

where

$$P_{n\varepsilon} = \sum_{j=1}^{k_n} T_\varepsilon(W(r_{j-1}))(V(r_j) - V(r_{j-1})) = \sum_{j=1}^{k_n} \int_{r_{j-1}}^{r_j} T_\varepsilon(W(r_{j-1}))dV(r)$$

and

$$I_{n\varepsilon} = \int_0^1 T_\varepsilon(W(r))dV(r).$$

Then it will be shown that

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} E|P_n - P_{n\varepsilon}|^2 = 0, \tag{I}$$

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} E|P_{n\varepsilon} - I_{n\varepsilon}|^2 = 0, \tag{II}$$

and

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} E|I_{n\varepsilon} - I|^2 = 0. \tag{III}$$

For (I),

$$\limsup_{n \to \infty} E|P_n - P_{n\varepsilon}|^2 = \limsup_{n \to \infty} E\left|\sum_{j=1}^{k_n} \int_{r_{j-1}}^{r_j} [\ln|W(r_{j-1})| - T_\varepsilon(W(r_{j-1}))]dV(r)\right|^2$$

$$= \limsup_{n \to \infty} \sum_{j=1}^{k_n} \int_{r_{j-1}}^{r_j} E[T_\varepsilon(W(r_{j-1})) - \ln|W(r_{j-1})|]^2 dr$$

$$= \limsup_{n \to \infty} \sum_{j=1}^{k_n} \int_{r_{j-1}}^{r_j} E[\ln|\varepsilon| - \ln|W(r_{j-1})|]^2 I(|W(r_{j-1})| \le \varepsilon)dr$$

$$= \limsup_{n \to \infty} \frac{1}{k_n} \sum_{j=1}^{k_n} E[\ln|\varepsilon| - \ln|W(r_{j-1})|]^2 I(|W(r_{j-1})| \le \varepsilon)$$

$$= \limsup_{n \to \infty} \frac{1}{k_n} \sum_{j=1}^{k_n} \int_{-\infty}^{\infty} [\ln|\varepsilon| - \ln|\sqrt{r_{j-1}}z|]^2 I(|\sqrt{r_{j-1}}z| \le \varepsilon)\phi(z)dz,$$

17

and by using the substitution $y = \sqrt{r_{j-1}}z$, it follows that

$$\limsup_{n\to\infty} E|P_n - P_{n\varepsilon}|^2 = \limsup_{n\to\infty} \frac{1}{k_n} \sum_{j=1}^{k_n} \int_{-\varepsilon}^{\varepsilon} [\ln|\varepsilon| - \ln|y|]^2 \frac{1}{\sqrt{r_{j-1}}} \phi(y) dy$$

$$\leq \limsup_{n\to\infty} \frac{1}{k_n} \frac{1}{\sqrt{2\pi}} \sum_{j=1}^{k_n} \int_{-\varepsilon}^{\varepsilon} [\ln|\varepsilon| - \ln|y|]^2 \frac{1}{\sqrt{r_{j-1}}} dy$$

$$= \frac{1}{\sqrt{2\pi}} \left[\limsup_{n\to\infty} \frac{1}{k_n} \sum_{j=1}^{k_n} \frac{1}{\sqrt{r_{j-1}}} \right] \int_{-\varepsilon}^{\varepsilon} [\ln|\varepsilon| - \ln|y|]^2 dy$$

$$\to 0 \text{ as } \varepsilon \to 0.$$

Note that $\limsup_{n\to\infty} \frac{1}{k_n} \sum_{j=1}^{k_n} \frac{1}{\sqrt{r_{j-1}}} = 2$. For (II),

$$\limsup_{n\to\infty} E|P_{n\varepsilon} - I_{n\varepsilon}|^2 = \limsup_{n\to\infty} E \left| \sum_{j=1}^{k_n} \int_{r_{j-1}}^{r_j} T_\varepsilon(W(r_{j-1})) dV(r) - \int_0^1 T_\varepsilon(W(r)) dV(r) \right|^2$$

$$\leq \limsup_{n\to\infty} \sum_{j=1}^{k_n} \int_{r_{j-1}}^{r_j} E|T_\varepsilon(W(r_{j-1})) - T_\varepsilon(W(r))|^2 dr = 0$$

for any $\varepsilon > 0$, because $T_\varepsilon(\cdot)$ is uniformly continuous. Finally, for (III),

$$\limsup_{n\to\infty} E|I_{n\varepsilon} - I|^2 = E \left| \int_0^1 T_\varepsilon(W(r)) dV(r) - \int_0^1 \ln|W(r)| dV(r) \right|^2$$

$$= E \left| \int_0^\varepsilon \ln|W(r)| dV(r) \right|^2$$

$$= \int_0^\varepsilon E[\ln|W(r)|]^2 dr$$

$$= \int_0^\varepsilon \left[\int_{-\infty}^\infty [\ln|\sqrt{r}z|]^2 \phi(z) dz \right] dr$$

$$= \left[\int_0^\varepsilon \frac{1}{\sqrt{r}} dr \right] \left[\int_{-\infty}^\infty [\ln|x|]^2 \phi(x) dx \right] \to 0 \text{ as } \varepsilon \to 0.$$

This is because

$$\int_{-\infty}^\infty [\ln|x|]^2 \phi(x) dx = \int_{-1}^1 [\ln|x|]^2 \phi(x) dx + \int_{\mathbb{R}\setminus[-1,1]} [\ln|x|]^2 \phi(x) dx$$

$$\leq \frac{1}{\sqrt{2\pi}} \int_{-1}^1 [\ln|x|]^2 dx + \int_{\mathbb{R}\setminus[-1,1]} |x|^2 \phi(x) dx < \infty.$$

Note that $[\ln|x|]^2 < |x|^2$ if $|x| > 1$, and that the logarithm function and its square are locally integrable. Also note that

$$\int_{\mathbb{R}\setminus[-1,1]} |x|^2 \phi(x) dx < \int_{-\infty}^\infty |x|^2 \phi(x) dx = 1,$$

because $\phi(\cdot)$ is the pdf of the standard normal. We have established that $|P_n - I| \xrightarrow{p} 0$, which complete the proof of the theorem. ∎

To prove Theorem 2, we begin with the following lemma.

Lemma 1 *As $n \to \infty$,*

$$\frac{1}{n} \sum_{j=1}^{n} \ln \left| \frac{x_t}{\sqrt{n}} \right| \xrightarrow{d} \int_0^1 \ln |W(r)| dr,$$

$$\frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln \left| \frac{x_{n_{j-1}+1}}{\sqrt{n}} \right| \xrightarrow{d} \int_0^1 \ln |W(r)| dr,$$

$$\frac{1}{n} \sum_{j=1}^{n} \left[\ln \left| \frac{x_t}{\sqrt{n}} \right| \right]^2 \xrightarrow{d} \int_0^1 [\ln |W(r)|]^2 dr,$$

and

$$\frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \left[\ln \left| \frac{x_{n_{j-1}+1}}{\sqrt{n}} \right| \right]^2 \xrightarrow{d} \int_0^1 [\ln |W(r)|]^2 dr.$$

Proof of Lemma 1: See de Jong (2004) or Pötscher (2004). ∎

Proof of Theorem 2: Consider

$$H_n = \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln \left| \frac{x_t}{\sqrt{n}} \right| \ln \left| \frac{x_{n_{j-1}+1}}{\sqrt{n}} \right|.$$

The theorem can be proved by showing that

$$H_{n\varepsilon} \xrightarrow{d} H_\varepsilon \text{ as } n \to \infty \text{ for all } \varepsilon > 0, \tag{I}$$

$$H_\varepsilon \xrightarrow{d} I \text{ as } \varepsilon \to 0, \tag{II}$$

and

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} E|H_n - H_{n\varepsilon}| = 0 \tag{III}$$

where

$$H_{n\varepsilon} = \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} T_\varepsilon \left(\frac{x_t}{\sqrt{n}} \right) T_\varepsilon \left(\frac{x_{n_{j-1}+1}}{\sqrt{n}} \right),$$

$$H_\varepsilon = \int_0^1 [T_\varepsilon(W(r))]^2 dr,$$

19

and

$$I = \int_0^1 [\ln |W(r)|]^2 dr.$$

Note that (I) holds by the continuous mapping theorem, and (II) holds by the occupation times formula; see de Jong (2004, p.633-635). For (III), it follows that

$$\limsup_{n \to \infty} E|H_n - H_{n\varepsilon}|$$

$$\le \limsup_{n \to \infty} E \left| \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \left[\ln \left| \frac{x_t}{\sqrt{n}} \right| - T_\varepsilon \left(\frac{x_t}{\sqrt{n}} \right) \right] \ln \left| \frac{x_{n_{j-1}+1}}{\sqrt{n}} \right| \right| \tag{4}$$

$$+ \limsup_{n \to \infty} E \left| \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} T_\varepsilon \left(\frac{x_t}{\sqrt{n}} \right) \left[\ln \left| \frac{x_{n_{j-1}+1}}{\sqrt{n}} \right| - T_\varepsilon \left(\frac{x_{n_{j-1}+1}}{\sqrt{n}} \right) \right] \right|. \tag{5}$$

For (4), we have by the Cauchy-Schwartz inequality,

$$\limsup_{n \to \infty} E \left| \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \left[\ln \left| \frac{x_t}{\sqrt{n}} \right| - T_\varepsilon \left(\frac{x_t}{\sqrt{n}} \right) \right] \ln \left| \frac{x_{n_{j-1}+1}}{\sqrt{n}} \right| \right|^2$$

$$\le \left[\limsup_{n \to \infty} \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} E \left[\ln \left| \frac{x_t}{\sqrt{n}} \right| - T_\varepsilon \left(\frac{x_t}{\sqrt{n}} \right) \right]^2 \right]$$

$$\cdot \left[\limsup_{n \to \infty} \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} E \left[\ln \left| \frac{x_{n_{j-1}+1}}{\sqrt{n}} \right| \right]^2 \right].$$

It will be shown that the first term converges to zero as $\varepsilon \to 0$, and the second term is bounded. For the first term, let $Q_\varepsilon(x) = (\ln(\varepsilon) - \ln|x|)I(|x| < \varepsilon)$ and $f_t(\cdot)$ be the density of $\frac{x_t}{\sqrt{t}}$. Then by using the substitution $y = \frac{x_t}{\sqrt{t}}$, we have

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} E \left[\ln \left| \frac{x_t}{\sqrt{n}} \right| - T_\varepsilon \left(\frac{x_t}{\sqrt{n}} \right) \right]^2$$

$$= \limsup_{n \to \infty} \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \int_{-\infty}^{\infty} [Q_\varepsilon(\frac{\sqrt{t}}{\sqrt{n}}y)]^2 f_t(y) dy$$

$$\le \left[\sup_t \sup_{y \in \mathbb{R}} f_t(y) \right] \limsup_{n \to \infty} \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \int_{-\infty}^{\infty} [Q_\varepsilon(\frac{\sqrt{t}}{\sqrt{n}}y)]^2 dy$$

$$\le \left[\sup_t \sup_{y \in \mathbb{R}} f_t(y) \right] \left[\limsup_{n \to \infty} \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \frac{\sqrt{n}}{\sqrt{t}} \right] \left[\int_{-\infty}^{\infty} [Q_\varepsilon(x)]^2 dx \right] \to 0 \text{ as } \varepsilon \to 0.$$

Note that $f_t(y)$ is uniformly bounded over t, and that $\limsup_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \frac{\sqrt{n}}{\sqrt{t}} = 2$. Also note that

$$\lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} [Q_\varepsilon(x)]^2 dx = \lim_{\varepsilon \to 0} \int_{-\varepsilon}^{\varepsilon} [\ln|x| - \ln(\varepsilon)]^2 dx = 0$$

since the logarithm function and its square are locally integrable. For the second term, again by using the substitution $y = \frac{x_{n_{j-1}+1}}{\sqrt{n_{j-1}+1}}$, we have

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} E\left[\ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right|\right]^2 = \limsup_{n\to\infty} \frac{1}{k_n} \sum_{j=1}^{k_n} E\left[\ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right|\right]^2$$

$$= \limsup_{n\to\infty} \frac{1}{k_n} \sum_{j=1}^{k_n} \int_{-\infty}^{\infty} \left[\ln\left|\frac{\sqrt{n_{j-1}+1}}{\sqrt{n}}y\right|\right]^2 f_{n_{j-1}+1}(y)dy$$

$$= \limsup_{n\to\infty} \frac{1}{k_n} \sum_{j=1}^{k_n} \frac{\sqrt{n}}{\sqrt{n_{j-1}+1}} \int_{-\infty}^{\infty} [\ln|x|]^2 f_{n_{j-1}+1}(x)dx$$

$$= \limsup_{n\to\infty} \frac{1}{k_n} \sum_{j=1}^{k_n} \frac{\sqrt{n}}{\sqrt{n_{j-1}+1}} \left[\int_{-1}^{1} [\ln|x|]^2 f_{n_{j-1}+1}(x)dx + \int_{\mathbb{R}\setminus[-1,1]} [\ln|x|]^2 f_{n_{j-1}+1}(x)dx\right]$$

$$\leq \left[\limsup_{n\to\infty} \frac{1}{k_n} \sum_{j=1}^{k_n} \frac{\sqrt{n}}{\sqrt{n_{j-1}+1}}\right] \left[\sup_t \sup_{x\in[-1,1]} f_t(x)\right] \left[\int_{-1}^{1} [\ln|x|]^2 dx\right]$$

$$+ \left[\limsup_{n\to\infty} \frac{1}{k_n} \sum_{j=1}^{k_n} \frac{\sqrt{n}}{\sqrt{n_{j-1}+1}}\right] \left[\sup_t \int_{\mathbb{R}\setminus[-1,1]} |x|^2 f_t(x)dx\right]$$

$$< \infty.$$

Note that the first inequality holds because $[\ln|x|]^2 < |x|^2$ if $|x| > 1$. Also note that

$$\left[\sup_t \sup_{x\in[-1,1]} f_t(x)\right] \left[\int_{-1}^{1} [\ln|x|]^2 dx\right] < \infty$$

because $f_t(x)$ is uniformly bounded over t and the logarithm function and its square are locally integrable, and that

$$\sup_t \int_{\mathbb{R}\setminus[-1,1]} |x|^2 f_t(x)dx < \sup_t \int_{-\infty}^{\infty} |x|^2 f_t(x)dx = \sup_t E\left|\frac{x_t}{\sqrt{t}}\right|^2 < \infty$$

by Assumption 1, 2 and 3. For (5), we have again by the Cauchy-Schwartz inequality,

$$\limsup_{n\to\infty} E\left|\frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} T_\varepsilon\left(\frac{x_t}{\sqrt{n}}\right) \left[\ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right| - T_\varepsilon\left(\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right)\right]\right|^2$$

$$\leq \left[\limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} E\left[T_\varepsilon\left(\frac{x_t}{\sqrt{n}}\right)\right]^2\right]$$

$$\cdot \left[\limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} E\left[\ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right| - T_\varepsilon\left(\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right)\right]^2\right].$$

Similarly, it can be shown that the second term converges to zero as $\varepsilon \to 0$, and that the first term is bounded. This finishes the proof. ∎

Proof of Theorem 3: The NCLS estimator $\widetilde{\beta}_1$ satisfies

$$\sqrt{n}(\widetilde{\beta}_1 - \beta_1) = \frac{\frac{1}{\sqrt{n}} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} (\ln|x_{n_{j-1}+1}| - \overline{\ln|x_{n_{j-1}+1}|})(u_t - \overline{u_t})}{\frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} (\ln|x_{n_{j-1}+1}| - \overline{\ln|x_{n_{j-1}+1}|})(\ln|x_t| - \overline{\ln|x_t|})},$$

where $\overline{X_t}$ indicate the sample average of X_t. For the numerator, we have

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \left[\ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right| - \overline{\ln\left|\frac{x_t}{\sqrt{n}}\right|} \right] [u_t - \overline{u_t}]$$

$$= \frac{1}{\sqrt{n}} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right| [u_t - \overline{u_t}]$$

$$= \left[\frac{1}{\sqrt{n}} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right| u_t \right] - [\sqrt{n}\overline{u_t}] \left[\frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right| \right].$$

Note that

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right| u_t \xrightarrow{d} \int_0^1 \ln|W(r)| dU(r)$$

by Theorem 1, and that

$$\sqrt{n}\overline{u_t} \xrightarrow{d} U(1)$$

and

$$\frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right| \xrightarrow{d} \int_0^1 \ln|W(r)| dr$$

by Lemma 1. For the denominator,

$$\frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} [\ln|x_{n_{j-1}+1}| - \overline{\ln|x_{n_{j-1}+1}|}][\ln|x_t| - \overline{\ln|x_t|}]$$

$$= \left[\frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right| \ln\left|\frac{x_t}{\sqrt{n}}\right| \right] - \left[\overline{\ln\left|\frac{x_t}{\sqrt{n}}\right|} \right] \left[\frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right| \right].$$

Note that

$$\frac{1}{n} \sum_{j=1}^{k_n} \sum_{t=n_{j-1}+1}^{n_j} \ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right| \ln\left|\frac{x_t}{\sqrt{n}}\right| \xrightarrow{d} \int_0^1 (\ln|W(r)|)^2 dr$$

by Theorem 2, and that

$$\overline{\ln\left|\frac{x_t}{\sqrt{n}}\right|} \xrightarrow{d} \int_0^1 \ln|W(r)| dr$$

22

and

$$\frac{1}{n}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j}\ln\left|\frac{x_{n_{j-1}+1}}{\sqrt{n}}\right|\xrightarrow{d}\int_0^1\ln|W(r)|dr$$

by Lemma 1. ∎

Proof of Theorem 4: By Corollary 3 in Jansson (2002, p.1452), it suffices to show that under Assumption 1, 2, 3 and 4, Assumption A5(ii) in Jansson (2002), which is restated below, is satisfied.

A5(ii): $\sqrt{n}(\widehat{\theta}-\theta)=O_p(1)$ and
$$\sup_{t\geq 1}E\left(\sup_{\theta\in\mathcal{N}}||\frac{\partial V_t(\theta)}{\partial\theta'}||^2\right)<\infty\text{ for some neighborhood }\mathcal{N}\text{ of }\theta.$$

By letting $V_t(\theta)=u_t(\beta)$, $\theta=\beta_1$, and $\widehat{\theta}=\widetilde{\beta}_1$, we have

$$V_t(\theta)=(y_t-\overline{y_t})-\theta\left(\ln\left|\frac{x_t}{\sqrt{n}}\right|-\overline{\ln\left|\frac{x_t}{\sqrt{n}}\right|}\right)+\overline{u}.$$

Note that $\sqrt{n}(\widetilde{\beta}_1-\beta_1)=O_p(1)$ by Theorem 3, and that

$$\sup_{t\geq 1}E\left(\sup_{\theta\in\mathcal{N}}||\frac{\partial V_t(\theta)}{\partial\theta}||^2\right)=\sup_{t\geq 1}E\left[\ln\left|\frac{x_t}{\sqrt{n}}\right|-\overline{\ln\left|\frac{x_t}{\sqrt{n}}\right|}\right]^2$$

$$=\sup_{t\geq 1}E\left[\ln\left|\frac{x_t}{\sqrt{t}}\right|-\overline{\ln\left|\frac{x_t}{\sqrt{t}}\right|}+\ln\left|\frac{\sqrt{t}}{\sqrt{n}}\right|-\overline{\ln\left|\frac{\sqrt{t}}{\sqrt{n}}\right|}\right]^2$$

$$\leq 2\sup_{t\geq 1}E\left[\ln\left|\frac{x_t}{\sqrt{t}}\right|-\overline{\ln\left|\frac{x_t}{\sqrt{t}}\right|}\right]^2+2\sup_{t\geq 1}E\left[\ln\left|\frac{\sqrt{t}}{\sqrt{n}}\right|-\overline{\ln\left|\frac{\sqrt{t}}{\sqrt{n}}\right|}\right]^2$$

$$<\infty.$$

Note that the first term is finite since under Assumption 1, 2 and 3, $\frac{x_t}{\sqrt{t}}$ has the density $f_t(\cdot)$ that is uniformly bounded over t, and the logarithm and its square are locally integrable functions. Also note that the second term is finite too. ∎

Proof of Theorem 5: The fully modified NCLS estimator $\widehat{\beta}_1$ satisfies

$$\sqrt{n}(\widehat{\beta}_1-\beta_1)=\frac{\frac{1}{\sqrt{n}}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j}(\ln|x_{n_{j-1}+1}|-\overline{\ln|x_{n_{j-1}+1}|})(\widehat{u}_t^\dagger-\overline{\widehat{u}_t^\dagger})}{\frac{1}{n}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j}(\ln|x_{n_{j-1}+1}|-\overline{\ln|x_{n_{j-1}+1}|})(\ln|x_t|-\overline{\ln|x_t|})}$$

$$=\frac{\frac{1}{\sqrt{n}}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j}(\ln|x_{n_{j-1}+1}|-\overline{\ln|x_{n_{j-1}+1}|})[u_t^\dagger+(\widehat{u}_t^\dagger-u_t^\dagger)]}{\frac{1}{n}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j}(\ln|x_{n_{j-1}+1}|-\overline{\ln|x_{n_{j-1}+1}|})(\ln|x_t|-\overline{\ln|x_t|})}.$$

Since by Theorem 3,

$$\frac{\frac{1}{\sqrt{n}}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j}(\ln|x_{n_{j-1}+1}|-\overline{\ln|x_{n_{j-1}+1}|})u_t^{\dagger}}{\frac{1}{n}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j}(\ln|x_{n_{j-1}+1}|-\overline{\ln|x_{n_{j-1}+1}|})(\ln|x_t|-\overline{\ln|x_t|})}$$

$$\xrightarrow{d}\frac{\int_0^1\ln|W(r)|dU^{\dagger}(r)-U^{\dagger}(1)\int_0^1\ln|W(r)|dr}{\int_0^1(\ln|W(r)|)^2dr-\left[\int_0^1\ln|W(r)|dr\right]^2},$$

it suffices to show that

$$\frac{\frac{1}{\sqrt{n}}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j}(\ln|x_{n_{j-1}+1}|-\overline{\ln|x_{n_{j-1}+1}|})(\hat{u}_t^{\dagger}-u_t^{\dagger})}{\frac{1}{n}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j}(\ln|x_{n_{j-1}+1}|-\overline{\ln|x_{n_{j-1}+1}|})(\ln|x_t|-\overline{\ln|x_t|})}$$

$$=\frac{\frac{1}{\sqrt{n}}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j}(\ln|x_{n_{j-1}+1}|-\overline{\ln|x_{n_{j-1}+1}|})[\Omega_{21}'\Omega_{22}^{-1}-\hat{\Omega}_{21}'\hat{\Omega}_{22}^{-1}]\Delta x_t}{\frac{1}{n}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j}(\ln|x_{n_{j-1}+1}|-\overline{\ln|x_{n_{j-1}+1}|})(\ln|x_t|-\overline{\ln|x_t|})}\xrightarrow{p}0.$$

It is equivalent to show that the numerator converges to zero in probability because the denominator converges in distribution to an almost surely positive random variable by Theorem 2. To show that the numerator converges to zero in probability, note that

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j}(\ln|x_{n_{j-1}+1}|-\overline{\ln|x_{n_{j-1}+1}|})[\Omega_{21}'\Omega_{22}^{-1}-\hat{\Omega}_{21}'\hat{\Omega}_{22}^{-1}]\Delta x_t$$

$$=[\Omega_{21}'\Omega_{22}^{-1}-\hat{\Omega}_{21}'\hat{\Omega}_{22}^{-1}]\frac{1}{\sqrt{n}}\sum_{j=1}^{k_n}\sum_{t=n_{j-1}+1}^{n_j}(\ln|x_{n_{j-1}+1}|-\overline{\ln|x_{n_{j-1}+1}|})\Delta x_t$$

$$=[\Omega_{21}'\Omega_{22}^{-1}-\hat{\Omega}_{21}'\hat{\Omega}_{22}^{-1}]\cdot C_n\xrightarrow{p}0\text{ as }n\to\infty$$

because $\widehat{\Omega}$ is a consistent estimator for Ω by Theorem 4, and C_n is $O_p(1)$ by Theorem 1. $\blacksquare$

References

Anderson, Richard G., Rasche, Robert H. 2001. The remarkable stability of monetary base velocity in the united states, 1919-1999. *Federal Reserve Bank of St. Louis Working Paper 2001-008A.*

Andrews, D.W.K 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59**(3): 817–858.

Ball, Laurence 1998. Another look at long-run money demand. *NBER Working Paper 6597.*

Chang, Yoosoon, Park, Joon Y., Phillips, Peter C.B. 2001. Nonlinear econometric models with cointegrated and deterministically trending regressors. *Econometrics Journal* **4**(1): 1–36.

Davidson, James 1994 *Stochastic Limit Theory.* Oxford University Press.

de Jong, Robert M. 2002. Nonlinear estimator with integrated regressors but without exogeneity. *Michigan State University, mimeo.*

——— 2004. Addendum to "asymptotics for nonlinear transformations of integrated time series". *Econometric Theory* **20**: 627–635.

Dutkowsky, Donald H., Cynamon, Barry Z. 2003. Sweep program: The fall of m1 and rebirth of the medium of exchange. *Journal of Money, Credit, and Banking* **35**(2): 263–279.

Hoffman, Dennis L., Rasche, Robert H. 1991. Long-run income and interest elasticities of money demand in the united states. *The Review of Economics and Statistics* **73**(4): 665–674.

Hu, Ling, Phillips, Peter C.B. 2002. Dynamics of federal funds target rate: a nonstationary discrete choice approach. *Cowles Foundation Discussion Paper No. 1365.*

Jansson, Michael 2002. Consistent covariance matrix estimation for linear processes. *Econometric Theory* **18**: 1449–1459.

Krugman, Paul R. 1998. It's baaack: Japan's slump and the return of the liquidity trap. *Brookings Papers on Economic Activity* **2**: 137–187.

Kwiatkowski, D., Phillips, Peter C.B., Schmidt, P., Shin, Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* **54**: 159–178.

Lucas Jr., Robert E. 1988. Money demand in the united states: A quantitative review. *Carnegie-Rochester Conference Series on Public Policy* **29**: 1061–1079.

Ng, S, Perron, P 1996. The exact error in estimating the spectral density at the origin. *Journal of Time Series Analysis* **17**: 379–408.

Nishiyama, Shinichi 2003. Inflation target as a buffer against liquidity trap. *IMES Discussion Paper No. 2003-E-8.*

Park, Joon Y., Phillips, Peter C.B. 1999. Asymptotics for nonlinear transformations of integrated time series. *Econometric Theory* **15**(3): 269–298.

⸺ 2001. Nonlinear regressions with integrated time series. *Econometrica* **69**(1): 117–161.

Phillips, Peter C.B., Hansen, Bruce E. 1990. Statistical inference in instrumental variables regression with i(1) processes. *Review of Economics Studies* **57**: 99–125.

Pötscher, Benedikt M. 2004. Nonlinear functions and convergence to brownian motion: Beyond the continuous mapping theorem. *Econometric Theory* **20**: 1–22.

Stock, James H., Watson, Mark W. 1993. A simple estimator of cointegrating vectors in higher order integrated system. *Econometrica* **61**(4): 783–820.