Robust Nonnested Testing and the Demand for Money

Hwan-sik Choi* Nicholas M. Kiefer†

Cornell University Cornell University

October, 2006

Abstract

Non-nested hypothesis testing procedures have been recently extended to dynamic non-nested models. We propose robust tests that generalize the J test (Davidson and MacKinnon (1981)) and the F-test (Deaton (1982) and Dastoor (1983)) for non-nested dynamic models with unknown serial correlation and conditional heteroskedasticity in errors. We investigate the finite sample properties of our test statistics and propose to use the bootstrap methods or the fixed-b asymptotics developed in Kiefer and Vogelsang (2002a,b, 2005) to improve the asymptotic approximation to the sampling distribution of the test statistics. The semiparametric bootstrap and the fixed-b approaches are compared with the standard normal or chi-square approximations using Monte Carlo simulations, and are found to give markedly superior approximations. We also present an application to U. S. money demand models, where consumption seems to be a better variable than income as a "scale" variable in the money demand function.

JEL classification: C12, C14, C15, C52

Keywords: Non-nested hypothesis, J test, F test, robust, bootstrap, fixed-b, money demand.

*hc269@cornell.edu. 404 Uris Hall, Department of Economics, Cornell University, Ithaca, NY, 14850, USA.

†nmk1@cornell.edu. 490 Uris Hall, Department of Economics and Department of Statistical Science, Cornell University, Ithaca, NY, 14850, USA.

[0]We are grateful to James MacKinnon and Tim Vogelsang for discussions and comments, and to the referees for useful comments.

1 Introduction

Distinguishing non-nested or separate families of hypotheses for model selection has been an important and active area of formal research since Cox (1961, 1962). Cox used centered log likelihood ratios between two non-nested models. Goldfeld and Quandt (1972) noted the importance of model selection tests for choosing multiplicative or additive errors, and Quandt (1974) considered a non-nested test of $\lambda = 0$ or $\lambda = 1$ in an artificial compound model $\lambda p_1 + (1 - \lambda)p_2$ from a mixture of two competing models with distributions p_1 and p_2. Pesaran (1974) introduced a test for non-nested linear regression models. Davidson and MacKinnon (1981) proposed the popular J test. Deaton (1982) and Dastoor (1983) proposed an F test. These approaches require nesting the competing models in a more general model.

McAleer (1995) compared 9 different non-nested testing procedures and found that the J test (especially a paired comparison) is most popular in empirical papers in journals he considered. The J test is computationally straightforward and easy to interpret. But he also noted that the J test is based on i.i.d. errors, and a diagnostic test to validate this assumption was not performed often.

The finite sample properties of the J test are known to be poor in some cases even with considerably large samples (Godfrey and Pesaran (1983), McAleer (1995)). In general, the J test is known to reject a correct null hypothesis more often than a specified level of the test. Fisher and McAleer (1981) and Godfrey and Pesaran (1983) suggested the J_A test using a bias correction of the numerator of the statistic. But the J_A test often has much lower power than the J test. Fan and Li (1995), Godfrey (1998) and Davidson and MacKinnon (2002) used bootstrap methods to approximate the sampling distribution of the J test statistic.

Dynamic models have not been extensively considered in the J test literature. Davidson and MacKinnon (2002) mentioned the possibility of the use of the bootstrap method for the J test in dynamic models. This paper relaxes the i.i.d. error assumption of the J test and considers a generalized J test with serially dependent observations or dynamic models. We propose a robust version of J test, namely the J_D test, using a Heteroskedasticity/Autocorrelation Consistent (HAC) estimator. A HAC version of the non-nested F test (F_D test) of Deaton (1982) and Dastoor (1983) is also proposed, and its size and power properties are compared with the proposed J_D test.

Since HAC estimators require choosing a bandwidth parameter M, the finite sample performance crucially depends on M. We use Kiefer-Vogelsang-Bunzel (KVB) fixed-b asymptotics (Kiefer et al. (2000), Kiefer and Vogelsang (2002a,b), Kiefer and Vogelsang (2005)) to approximate the sampling distribution of our test statistics. We also consider the semiparametric i.i.d. and block bootstrap methods, and compare the performance with the fixed-b asymptotics approximations as well as the standard normal approximation.

We present an empirical application to a money demand function. The question of whether the relevant scale variable in money demand is income or consumption was raised by Mankiw and Summers (1986) and subsequently examined by Elyasiani and Nasseh (1994). We revisit this question using our new technique and data through 2005. Our results support consumption as the better measure of economic activity as far as money demand is concerned.

2 KVB fixed-b asymptotics

Kiefer et al. (2000), Kiefer and Vogelsang (2002a,b, 2005) proposed a new asymptotic approximation to the sampling distribution of a HAC test statistic. We illustrate the KVB fixed-b approach with a linear model with a single regressor.

Let $\{y_t\}$ be generated by

$$y_t = \alpha + x_t\beta + u_t,\ (t = 1, 2, ..., T), \tag{2.1}$$

where $\{x_t\}$ is a regressor with $\text{plim}_{T\to\infty} \sum x_t^2/T > 0$, $E(u_t|x_t) = 0$ for all t, and $\{u_t\}$ is a weakly stationary process with autocovariance function $\gamma(j)$ $(j = 0, \pm1, \pm2, \ldots)$ with possible conditional heteroskedasticity. Let $\hat{\alpha}$ and $\hat{\beta}$ be OLS estimators of α and β respectively. For testing the null hypothesis $\beta = 0$, we consider the following statistic

$$Y_T = \frac{\sum_{t=1}^{T} \tilde{x}_t y_t/\sqrt{T}}{\sqrt{\widehat{V}_T}}, \tag{2.2}$$

where $\tilde{x}_t = x_t - \bar{x}$, $\bar{x} = \sum_{t=1}^{T} x_t/T$, and $\widehat{V}_T$ is given by

$$\widehat{V}_T = \sum_{j=1-T}^{T-1} K\left(\frac{j}{M}\right)\hat{\gamma}(j), \tag{2.3}$$

where $K(x)$ is the kernel of the non-parametric estimator $\widehat{V}_T$, $M \leq T$ is the bandwidth used in the kernel estimator, and

$$\hat{\gamma}(j) = \frac{1}{T}\sum_{t=|j|+1}^{T}(\hat{v}_t - \bar{v})(\hat{v}_{t-|j|} - \bar{v}), \tag{2.4}$$

where $\hat{v}_t = \tilde{x}_t\hat{u}_t$, $\bar{v} = \sum_{t=1}^{T}\hat{v}_t/T = 0$, and $\hat{u}_t = y_t - \hat{\alpha} - x_t\hat{\beta}$. Under the conventional asymptotics, $M/T \to 0$ as $T \to \infty$, the HAC variance estimator $\widehat{V}_T$ is consistent to the long run variance of the numerator of the test statistic, and we have asymptotic normality of the test statistic Y_T. The KVB approach models $M/T \to b$ in the $T \to \infty$ conceptual experiment. We assume that $\{v_t\} = \{\tilde{x}_t u_t\}$ satisfies the following.

Assumption 2.1 (Functional Central Limit Theorem (FCLT))

$$T^{-1/2}\sum_{t=1}^{[rT]} v_t \Rightarrow \lambda W(r), \tag{2.5}$$

where $W(r)$ is a standard Brownian motion defined on $C[0,1]$ and $\lambda^2 = \sum_{j=-\infty}^{\infty}\gamma(j) < \infty$.

As shown in Phillips and Durlauf (1986), this assumption holds under slightly weaker conditions than the assumptions for a consistent estimation in HAC variance estimation literature such as Andrews (1991) (though, weaker assumptions than Andrews (1991) can be found in Hansen (1992) and de Jong and Davidson (2000)). It allows conditional heteroskedasticity but excludes unconditional heteroskedasticity. See Kiefer and Vogelsang (2005, p.1135) for further discussion.

Under Assumption 2.1, if $\hat{\beta} \xrightarrow{p} \beta$ as in the OLS example above, and if we use the Bartlett kernel for example, Kiefer and Vogelsang (2005) showed that we have the following limiting distribution of the test statistic Y_T with $b = \lim_{T\to\infty} M/T > 0$,

$$Y_T \Rightarrow \frac{W(1)}{\sqrt{\frac{2}{b}\left[\int_0^1 \widetilde{W}(r)^2\, dr - \int_0^{1-b} \widetilde{W}(r+b)\widetilde{W}(r)\, dr\right]}}. \tag{2.6}$$

where $\widetilde{W}(r) = W(r) - rW(1)$.

See Kiefer and Vogelsang (2005, p.1137) for another example with Monte Carlo experiments and Kiefer and Vogelsang (2005, Theorem 3) for other kernel functions. Critical values of this non-standard limiting distribution must be simulated in practice, and they are tabulated in Kiefer and Vogelsang (2005, p.1146) for some popular kernels. We work with the Bartlett kernel throughout.

3 Dynamic J test and limiting distributions

We present the main idea with a non-nested testing of a pair of linear regression models. Let H_1 and H_2 be two competing non-nested linear models given, for $t = 1, ..., T$, by

$$H_1 : y_t = x_t'\beta_1 + u_{1t}, \tag{3.1}$$

$$H_2 : y_t = z_t'\beta_2 + u_{2t}, \tag{3.2}$$

where x_t, z_t are k_1 and k_2 dimensional (exogenous) regressors respectively and u_{1t}, u_{2t} are i.i.d. errors with mean zero and variance σ^2. For convenience we sweep out common regressors and intercepts in H_1 and H_2 by projection leading to eq. (3.1) and (3.2). Under the hypotheses that H_1 is the true model, the J test uses the artificial model

$$H_0 : y_t = x_t'\beta + \hat{y}_t\theta + u_t, \tag{3.3}$$

where $\hat{y} = \{\hat{y}_1, ..., \hat{y}_T\}'$ is the fitted value of the dependent variable $y = \{y_1, ..., y_T\}'$ from the regression H_2. The J test statistic is the t-statistic for θ,

$$J = \frac{\hat{\theta}}{\sqrt{\hat{\sigma}^2 \left(\hat{y}' M_X \hat{y}\right)^{-1}}}, \tag{3.4}$$

where $M_X = I - P_X$, P_X is the projection matrix onto the space spanned by the $(T \times k_1)$ regressor matrix X in the model H_1, $\hat{\theta} = \left(\hat{y}' M_X \hat{y}\right)^{-1} \hat{y}' M_X y$, $\hat{\sigma}^2 = \sum_{t=1}^{T} \hat{u}_t^2 / T$ and $\hat{u}_t$ is the residual from the regression equation (3.1) or (3.3). The sampling distribution of the test statistic J is approximated by the standard normal distribution. This test is widely used because of its simplicity and intuitive appeal (see McAleer (1995)).

This paper generalizes the J test by relaxing the assumptions on the errors in the true model. First, we formally assume the non-orthogonality of regressors in H_1 and H_2 as in Davidson and MacKinnon (1981).

Assumption 3.1 *The regressors x_t in H_1 and z_t in H_2 satisfy*

$$\operatorname*{plim}_{T \to \infty} T^{-1} \sum_{t=1}^{T} x_t z_t' \neq \mathbf{0}. \tag{3.5}$$

When the correlation between x_t and z_t is weak, it is known that the J test statistic shows over-rejection (Godfrey and Pesaran (1983), Godfrey (1998)). Michelis (1999) proposed a new asymptotic approximation under near population orthogonality (NPO) in which $\operatorname{plim}_{T \to \infty} T^{-1/2} \sum x_t z_t' = \Delta$, where Δ is a matrix of constants. When the regressors are orthogonal, we have $\hat{\beta}_2 \xrightarrow{p} 0$ and the estimated H_2 converges to a nested model in H_1. We can consider in this case an F test such as testing $\beta_2 = \mathbf{0}$ in

$$y_t = x_t' \beta + z_t' \beta_2 + u_t. \tag{3.6}$$

Under Assumption 3.1, we have non-degenerating $\hat{y}$.

We generalize the model by allowing serial correlation in the error process $\{u_{1t}\}$ of the true model H_1 with possible conditional heteroskedasticity.

Assumption 3.2 *The error process $\{u_{1t}\}$ of the true model H_1 is weakly stationary with unknown serial correlation and conditional heteroskedasticity.*

Assumption 3.2 does not specify a form of serial correlation or conditional heteroskedasticity. The test statistic proposed in this paper is robust to unknown serial correlation or heteroskedasticity, thus an empirical researcher need not test the existence of serial correlation. We exclude unconditional heteroskedasticity in the errors. The following assumption restricts the serial correlation under certain situation.

Assumption 3.3 *The error process $\{u_{it}\}$ for $i = 1, 2$ is assumed to satisfy*

$$E(u_{it}|x_t, z_t) = 0. \tag{3.7}$$

Assumption 3.3 excludes, for example, having both lagged dependent variables and serial correlation in the errors. It also excludes endogenous variables. But conditional heteroskedasticity in the errors is still allowed.

Further, we assume the FCLT holds for the partial sum of a product process $v_t = u_{1t}w_t$, where $\{w_t\}_{t=1}^T$ is asymptotically equivalent to the vector $M_X\hat{y} = M_X P_Z y$, where P_Z is the projection matrix onto the space spanned by the $(T \times k_2)$ regressor matrix Z in the model H_2. Specifically, as $T \to \infty$, w_t is the limit of the t-th observation $(M_X\hat{y})_t$ of $M_X\hat{y}$ given by

$$w_t = \operatorname*{plim}_{T\to\infty} (M_X\hat{y})_t \tag{3.8}$$

$$= (z_t' - x_t' Q_{xx}^{-1} Q_{zx}) Q_{zz}^{-1} Q_{zx} \beta_1, \tag{3.9}$$

where $Q_{xx} = \operatorname{plim}_{T\to\infty} X'X/T$, $Q_{zz} = \operatorname{plim}_{T\to\infty} Z'Z/T$ and $Q_{zx} = \operatorname{plim}_{T\to\infty} Z'X/T$. The vector $\beta_2^* = Q_{zz}^{-1} Q_{zx} \beta_1$ in eq. (3.9) is called a pseudo-true value of β_2.

Assumption 3.4 *The process* $\{v_t\} = \{u_{1t} w_t\}$ *satisfies*

$$T^{-1/2} \sum_{t=1}^{[rT]} v_t \Rightarrow \lambda W(r), \tag{3.10}$$

where λ^2 *is the long run variance of* $\{v_t\}$.

We propose the dynamic J test (J_D test) using a HAC estimator. Our statistic is given by

$$J_D = \frac{\hat{y}' M_X y / \sqrt{T}}{\sqrt{\widehat{V}_T}}, \tag{3.11}$$

where

$$\widehat{V}_T = \sum_{j=1-T}^{T-1} K\left(\frac{j}{M}\right) \hat{\gamma}(j), \tag{3.12}$$

and $K(\cdot)$ is the kernel of the non-parametric estimator $\widehat{V}_T$, bandwidth $M \leq T$ is the number of lags used for $\widehat{V}_T$,

$$\hat{\gamma}(j) = \frac{1}{T} \sum_{t=|j|+1}^{T} (\hat{v}_t - \bar{v})(\hat{v}_{t-|j|} - \bar{v}), \tag{3.13}$$

where $\hat{v}_t = \hat{u}_t \left(M_X \hat{y}\right)_t$, $\bar{v}$ is the sample mean of $\{\hat{v}_t\}$, and $\{\hat{u}_t\}$ is the residual vector from the null model or the artificial model. Although using $\{\hat{u}_t\}$ from either the null model or the artificial model is asymptotically equivalent, Davidson and MacKinnon (1985) and Ligeralde and Brown (1995) demonstrated that using null model can reduce the size problem (i.e. over-rejection) in HAC tests. But using the residuals from the null model can result in lower power when the imposed null is not the truth. We use the residuals from the null model in examples in the next section.

Theorem 1 *Under Assumption 3.1-3.4, if* $M/T \to b \in (0, 1]$, *the limiting distribution of the* J_D *test statistic in eq. (3.11) is given by the KVB fixed-b asymptotics of Kiefer and Vogelsang (2005, Theorem 3),*

$$J_D \Rightarrow \frac{W(1)}{\sqrt{Q_1(b)}}, \text{ as } T \to \infty, \tag{3.14}$$

where $Q_1(b)$ *is defined in Kiefer and Vogelsang (2005, Definition 1) depending on the kernel func-*

tion.

Proof. Under Assumption 3.3, we have a consistent estimator $\hat{\beta}_1 \xrightarrow{p} \beta_1$ as $T \to \infty$, and for all t,

$$\hat{u}_t - u_{1t} = x_t'(\beta_1 - \hat{\beta}_1) \xrightarrow{p} 0, \tag{3.15}$$

as $T \to \infty$. From Assumption 3.1, eq. (3.8) and (3.15), we have

$$\operatorname*{plim}_{T\to\infty} (\hat{v}_t - v_t) = \operatorname*{plim}_{T\to\infty} \{\hat{u}_t \,(M_X \hat{y})_t - u_{1t} w_t\} = 0, \tag{3.16}$$

for all t, as $T \to \infty$. Therefore FCLT applies to $\{\hat{v}_t\}$ by Assumption 3.4. Under Assumption 3.2, we get eq. (3.14) from Theorem 3 in Kiefer and Vogelsang (2005). ■

We also consider a HAC robust version of the non-nested F test of Deaton (1982) and Dastoor (1983). Consider the artificial model

$$y_t = x_t'\beta + z_t'\beta_2 + u_t. \tag{3.17}$$

Let $(k_2 \times 1)$ vector process $\{\hat{v}_t\} = \{\hat{u}_t(Z'M_X)_t\}$, where $(Z'M_X)_t$ is t-th column of $Z'M_X$ and $\hat{u}_t$ is the residual from the null or the artificial model. The HAC robust non-nested F test statistic F_D for $\beta_2 = \mathbf{0}$ is given by

$$F_D = \frac{Ty'(M_X Z)(\widehat{V}_T)^{-1}(Z'M_X)y}{k_2} \tag{3.18}$$

where

$$\widehat{V}_T = \sum_{j=1-T}^{T-1} K\left(\frac{j}{M}\right) \widehat{\Gamma}(j), \tag{3.19}$$

and

$$\widehat{\Gamma}(j) = \frac{1}{T} \sum_{t=j+1}^{T} (\hat{v}_t - \bar{v})(\hat{v}_{t-j} - \bar{v})' \text{ for } j \geq 0, \tag{3.20}$$

$$\widehat{\Gamma}(j) = \widehat{\Gamma}'(-j) \text{ for } j < 0. \tag{3.21}$$

Noting that $Z'M_X = Z' - Z'X(X'X)^{-1}X'$, we define the limit $\tilde{w}_t$ of the t-th observation $(Z'M_X)_t$ of $Z'M_X$ as

$$\tilde{w}_t = \underset{T\to\infty}{\text{plim}}(Z'M_X)_t = z_t - Q_{zx}Q_{xx}^{-1}x_t. \tag{3.22}$$

We assume the FCLT for the product process $v_t = u_{1t}\tilde{w}_t$.

Assumption 3.5 *For $\{v_t\} = \{u_{1t}\tilde{w}_t\}$, we have*

$$T^{-1/2}\sum_{t=1}^{[rT]} v_t \Rightarrow \Lambda W_{k_2}(r), \ \ \textit{as } T \to \infty, \tag{3.23}$$

where $\Lambda\Lambda' = \sum_{j=-\infty}^{\infty} \Gamma(j)$, $\Gamma(j) = E(v_t v_{t-j}')$, and $W_{k_2}(r)$ is a $(k_2 \times 1)$ vector standard Brownian motion.

Theorem 2 *Under the Assumption 3.1-3.3 and 3.5, if $M/T \to b \in (0,1]$, the limiting distribution of the F_D test statistic in eq. (3.18) is given by the KVB fixed-b asymptotics of Kiefer and Vogelsang (2005, Theorem 3),*

$$F_D \Rightarrow W_{k_2}(1)'Q_{k_2}(b)^{-1}W_{k_2}(1)/k_2, \tag{3.24}$$

where $Q_{k_2}(b)$ is defined in Kiefer and Vogelsang (2005, Definition 1) depending on the kernel function.

Proof. We use eq. (3.22) and $\hat{u}_t - u_t \xrightarrow{p} 0$ to get $\hat{v}_t - v_t \xrightarrow{p} 0$, as $T \to \infty$ for all t. Then FCLT applies to $\{\hat{v}_t\}$ by Assumption 3.5 on $\{v_t\}$. Eq. (3.24) follows from Kiefer and Vogelsang (2005, Theorem 3). ■

In the next section, we consider the finite sample properties of the proposed J_D and F_D test. The performance of the fixed-b approach is compared to the bootstrap methods and the standard normal approximation.

4 Monte Carlo Study

4.1 Specification of simulations

We present two examples. One is with serially correlated errors with conditional heteroskedasticity, and the other is with a lagged dependent variable and conditional heteroskedasticity only in the errors.

- Case I (serially correlated errors with conditional heteroskedasticity)

$$H_1 : y_t = x_t'\beta_1 + u_{1t} = x_{1t} + 0.5x_{2t} + u_{1t}, \tag{4.1}$$

$$H_2 : y_t = z_t'\beta_2 + u_{2t} = z_{1t} + 0.5z_{2t} + 0.5z_{3t} + 0.5z_{4t} + u_{2t}, \tag{4.2}$$

- Case II (a lagged dependent variable and conditional heteroskedasticity in errors)

$$H_1 : y_t = y_{t-1}\delta + x_{1t} + 0.5x_{2t} + u_{1t}, \tag{4.3}$$

$$H_2 : y_t = y_{t-1}\delta + z_{1t} + 0.5z_{2t} + 0.5z_{3t} + 0.5z_{4t} + u_{2t}. \tag{4.4}$$

In both cases we assume H_1 is the true model. Godfrey and Pesaran (1983) noted that J test may have poorer finite sample performance when there are different numbers of regressors in the models. Thus we define x_t and z_t to be (2×1) and (4×1) regressor vectors respectively. The regressors are generated from a vector autoregressive (VAR) process

$$W_t = \Phi W_{t-1} + \zeta_t, \tag{4.5}$$

where $W_t = (x_t', z_t')'$ is (6×1) vector, and Φ is the autoregressive coefficient matrix. The autoregressive coefficient matrix is given by a symmetric Toeplitz matrix

$$\Phi = \begin{pmatrix} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 \\ v_2 & v_1 & v_2 & v_3 & v_4 & v_5 \\ v_3 & v_2 & v_1 & v_2 & v_3 & v_4 \\ v_4 & v_3 & v_2 & v_1 & v_2 & v_3 \\ v_5 & v_4 & v_3 & v_2 & v_1 & v_2 \\ v_6 & v_5 & v_4 & v_3 & v_2 & v_1 \end{pmatrix}. \tag{4.6}$$

The error $\{\zeta_t\}$ is a mean zero Gaussian i.i.d. process with variance matrix Ω. The variance Ω of the error $\{\zeta_t\} = \{\zeta_{1t} \; ... \; \zeta_{6t}\}'$ has upper triangular (j,k)-elements given by

$$E\zeta_{jt}\zeta_{kt} = \begin{cases} 1, & \text{for } j = k, \\ 0.8, & \text{for } j = 1 \text{ and } k = 2, \\ 0.7, & \text{for } j < k \text{ and } j,k = 3,...,6, \\ C_{xz}, & \text{for } j < k \text{ and } j = 1,2, \; k = 3,...,6. \end{cases} \tag{4.7}$$

We consider the following two specifications for x_t and z_t for CASE I ensuring the stationarity of $\{W_t\}$.

- Specification 1 (Strong $\{x_t, z_t\}$): $(v_1, v_2, v_3, v_4, v_5, v_6) = (-0.3, \; 0.1, \; 0.3, \; -0.2, \; 0.1, \; -0.3)$ and $C_{xz} = 0.8$,
- Specification 2 (Weak $\{x_t, z_t\}$): $(v_1, v_2, v_3, v_4, v_5, v_6) = (0.8, \; 0, \; 0, \; 0, \; 0, \; 0)$ and $C_{xz} = 0.2$.

The specification 2 makes weaker correlation between x_t and z_t processes than the specification 1. For CASE II, we consider the specification 1 (Strong $\{x_t, z_t\}$) only.

We generated the error process $\{u_{1t}\}_{t=1}^{T}$ from AR(1) process with GARCH(1,1),

$$u_{1t} = \alpha u_{1(t-1)} + \varepsilon_t, \tag{4.8}$$

$$\varepsilon_t = \sigma_t \xi_t, \tag{4.9}$$

$$\xi_t \sim i.i.d.\ N(0,1), \tag{4.10}$$

$$\sigma_t^2 = 0.04 + 0.86\sigma_{t-1}^2 + 0.1\varepsilon_{t-1}^2. \tag{4.11}$$

We initialize $\{u_{1t}\}$ with the variance $\sigma_1^2 = 1$ and $u_{11} = \varepsilon_1$. We use $\alpha = 0,\ 0.5,\ 0.9,\ -0.5$ for CASE I, and $\alpha = 0$ only for CASE II. For CASE II, $\{y_t\}$ is generated by setting $\delta = 0.5$ or 0.9.

The GARCH(1,1) coefficients satisfy the stationarity conditions and moment conditions for $\{\varepsilon_t\}$ required for our test statistics. See Ling (1999) and Ling and McAleer (2002) for necessary and sufficient conditions on the parameters for the existence of higher moments in GARCH models.

We generated 100 observations of y_t then took the last $T = 50$ observations. Total number of iterations was $5,000$ with $B = 399$ bootstrap resamplings for each iteration. In all examples, we used the Bartlett kernel. The bandwidths were $M = 1, 2, ..., 10, 15, 20, ..., 50$. Even if we don't have serial correlation ($\alpha = 0$) in CASE II, we used bandwidths $M > 1$ to capture serial correlation in finite sample.

All tests are with 5% level, and we compare the J_D test, and the F_D test. The sampling distribution of the J_D and F_D tests was approximated by the following four methods.

1. The standard normal approximation for the J_D test, and $\chi^2(k_2)/k_2$ for the F_D test ($k_2 = 4$ in our example),

2. The fixed-b asymptotic approximation in eq. (3.14) for the J_D test, and eq. (3.24) for the F_D test,

3. Semi-parametric i.i.d. bootstrap: Using the residuals $\{\hat{u}_t\}$ from the null model H_1, bootstrap $\{u_t^*\}$ with i.i.d. resampling. Normalize u_t^* with

$$\sqrt{\frac{T}{T-4}}\left(u_t^* - \bar{u}^*\right), \tag{4.12}$$

where $\bar{u}^* = \sum_{t=1}^{T} u_t^*/T$. This step makes sure the bootstrap residuals have mean zero. The factor $\sqrt{\frac{T}{T-4}}$ is used to correct the smaller variance problem in the bootstrap residuals in small sample. See Davidson and MacKinnon (2002) for more details. Then $\{y_t^*\}$ is generated by using $\{u_t^*\}$ and estimated parameters from H_1. Calculate the bootstrap J_D^* and F_D^* test statistics $B = 399$ times. Critical values are from the quantiles of the empirical distribution of J_D^* and F_D^*.

4. Semi-parametric (overlapping) block bootstrap: We use the overlapping block bootstrap with block size five to get $\{u_t^*\}$, then follow the same procedure as above.

For power comparison, we can not compare size corrected powers of different approximations to the sampling distribution of the test statistic since they use the numerically same test statistic for a given kernel and a bandwidth. But we can compare size corrected powers of J_D and F_D tests with different kernels and bandwidths. Asymptotically speaking, where all the approximations provide correct size under an appropriate conceptual experiment, either $M/T \to 0$ or $M/T \to b$, it is notable that under the conventional standard normal approximation, asymptotic local power of a HAC test statistic is exactly same no matter which kernel function or bandwidth were used, but under the fixed-b asymptotics, the local power depends on the kernel and the bandwidth. Kiefer and Vogelsang (2005) found that the Bartlett kernel has good asymptotic local power comparing to other popular kernels, and using large bandwidths decreases local power in all the kernels. Our Monte Carlo study also supported their asymptotic results.

4.2 Size properties

Table $1-4$ show the rejection rates of the 5% level (two tail for J_D) tests for four different AR(1) error correlation coefficients α for CASE I. Table $1-2$ show the size performance of different asymptotic approximations of the strong correlation in the regressors (specification 1), and Table $3-4$ are from the weak regressor correlation (specification 2).

The bootstrap tests showed best performance in both J_D and F_D tests. The block bootstrap with large bandwidths shows robust performance in all settings we considered. The block bootstrap gave relatively good performance in the worst scenario (Table 4). It's interesting that the i.i.d. and block bootstrap methods showed similar performance to the block bootstrap in many cases (Table $1-2$). See Gonçalves and Vogelsang (2006) for an explanation of the nice performance of the i.i.d. bootstraps with serially correlated errors. They show that the bootstrap methods (both block and i.i.d.) have the same limiting distribution as the fixed-b asymptotics.

The fixed-b asymptotic approach provides a clear improvement over the standard or chi-square approximations, although it overrejects in small bandwidths. When the bandwidths are small, the fixed-b limiting distributions are "close" to the standard normal (or chi-square) distribution. Therefore they perform similarly. The fixed-b asymptotics works better in F_D tests than in J_D tests. Since J test has a finite sample bias from using fitted values $\hat{y}$ from the alternative model with the same data set, J_D will also carry this problem. Although the J_D test with the fixed-b asymptotics mitigates this problem when large bandwidths were used, it does not remove the problem because the fixed-b asymptotics corrects the sampling distribution of the asymptotic variance (denominator) of the test statistic. When the regressors are weakly correlated (the specification 2, Table $3-4$), this bias problem becomes more serious (Compare Table $1-2$ and $3-4$). The reason that the bootstrap methods work better than the fixed-b approach is that they corrects both the bias and the variance.

Table 5 shows the size of J_D and F_D tests in CASE II with the AR(1) coefficients $\delta = 0.5$ and 0.9 for the lagged dependent variable y_{t-1}. For Case II, even though the serial correlation in the errors is not present, using large bandwidths will capture the serial correlation in finite sample and gives better performance. The bootstrap (especially the block bootstrap) works best, and the fixed-b asymptotics in the J_D tests overrejects in small bandwidths but is a clear improvement over the standard normal or chi-square approximations as the bandwidth increases.

4.3 Power comparison

Table 6 (CASE I) and Table 7 (CASE II) show the size corrected (at 5%) power comparison of J_D and F_D test statistics for various bandwidths. It is known that the Davidson and MacKinnon's J test has better local power than F test in a paired comparison (Dastoor and McAleer (1989)). In our simulation, the J_D test also gave better size-adjusted power than F_D test especially when high bandwidths were used. The power decreases as bandwidth increases in both J_D and F_D tests, but the power decreased less in J_D tests. The decrease in power with large bandwidths is consistent with the asymptotic local power results of Kiefer and Vogelsang (2005).

Therefore we have trade-off between size and power in choosing bandwidth. It is recommended to use J_D test rather than F_D test for better power, and the block bootstrap for good size performance on the basis of our simulation results.

5 Money Demand

We test the idea of Mankiw and Summers (1986) that consumption (or personal expenditure) rather than income (Gross National Product, GNP) is the right scale variable for money demand (for M1 or M2). Elyasiani and Nasseh (1994) used various nonnested tests indicating that the consumption measure seems to be the right scale variable. They considered different measures for consumption and income. We consider their model,

$$y_t = \beta_1 + \beta_2 r_t + \beta_3 r_{t-1} + \beta_4 r_{t-2} + \beta_5 z_t + \beta_6 z_{t-1} + \beta_7 z_{t-2} + \varepsilon_t, \tag{5.1}$$

where y_t is the difference in log of real money stock $M2$, r_t is the difference in log of the 3-month treasury bill rate, z_t is the difference in log of *real personal expenditure* (for a consumption measure) or *real GNP* (for an income measure). Elyasiani and Nasseh (1994) adjusted for serial correlation in the errors though the Cochrane-Orcutt procedure which probably helps but may not remove the serial correlation completely. They used the original J test which is valid only under no serial correlation and no heteroskedasticity in the errors. Our approach does not require this step and the J_D test is directly applicable to the data. We use quarterly data from 1959.I (Jan) $\sim$ 2005.III (July) (187 observations) from the Federal Reserve Bank. We used the GNP implicit price deflator to get real M2, the 3-month treasury rates are from the secondary market rate, and the real personal consumption expenditures and the real GNP are in year-2000-dollars. Table 8 shows the results from OLS regressions when the consumption and the income measures were used. We can see that the consumption measure gives a better fit. Figure $1-2$ shows the results from the J_D and F_D tests. The solid lines are the values of the J_D and F_D test statistics. The other lines are critical values of various asymptotic approximations. "Boot(1)" and "Boot(5)" are critical values from the i.i.d. and the block bootstrap respectively. The bootstrap critical values were calculated from $B = 3999$ resamplings. In the tests of the null hypothesis that the income measure (GNP) is the scale variable against the consumption measure, we could reject the null at 5% level for all bandwidths with J_D tests. With F_D test, we reject the null on small bandwidths but could not reject with large bandwidths. When the null hypothesis is the consumption measure, we did not reject the null at 5% level for all bandwidths with both the J_D and the F_D tests. Our test supported the idea of Mankiw and Summers (1986) that consumption is the more appropriate scale variable in the money demand function with J_D tests. With F_D tests, it's unclear whether the result for large bandwidths in Figure 1 is from the low power of F_D or not. On balance, our results support the conclusion that consumption is the better scale variable in the money demand equation.

6 Conclusion

Robust J test and F test are proposed for comparing non-nested dynamic models. We generalized the test statistics to HAC robust versions (J_D and F_D test). We have shown by Monte Carlo simulations that the bootstrap approaches and the KVB fixed-b asymptotics correct the size distortion. The usual standard normal asymptotic approximation had the worst performance. The fixed-b approach provides a great improvement on the standard normal approximation. The i.i.d. and block bootstrap showed similar size properties being better than the fixed-b asymptotic approximation. The block bootstrap method showed robust results.

When the regressors are weakly correlated, the standard normal approximation overrejects seriously and the fixed-b asymptotics also overrejects although it performs better than the standard normal approximation. This overrejection is reduced as bandwidth increases in the fixed-b approach. The block bootstrap works best in this case.

In size controlled power experiments, the J_D test showed better power than the F_D test especially when large bandwidths were used. Strong serial correlation in the errors affected both size and power, and especially decreased power significantly.

In an application to the money demand function in the US, we find that aggregate consumption provides a better scale variable than income.

The overall finding is that HAC testing of nonnested hypotheses based on the J and F test is feasible and reliable providing a sensible approximation to the sampling distribution is used. The fixed-b and the bootstrap methods are significant improvements on the normal approximation, with the semiparametric block bootstraps providing further improvement especially when the regressors are weakly correlated.

References

Andrews, D. W. K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–854.

Cox, D. (1961), "Tests of separate families of hypotheses," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 105–123.

— (1962), "Further results on tests of separate families of hypotheses," *Journal of the Royal Statistical Society, Series B*, 24, 406–424.

Dastoor, N. (1983), "Some aspects of testing non-nested hypotheses," *Journal of Econometrics*, 21, 213–228.

Dastoor, N. K. and McAleer, M. (1989), "Some Power Comparisons of Joint and Paired Tests for Non-Nested Models under Local Hypotheses," *Econometic Theory*, 5, 83–94.

Davidson, R. and MacKinnon, J. G. (1981), "Several tests for model specification in the presence of alternative hypotheses," *Econometrica*, 49, 781–793.

— (1985), "Heteroskedasticity-Robust Tests in Regression Directions," *Annales de l'INSEE*, 59/60, 183–218.

— (2002), "Bootstrap J tests of nonnested linear regression models," *Journal of Econometrics*, 109, 167–193.

de Jong, R. M. and Davidson, J. (2000), "Consistency of Kernel Estimators of Heteroscedastic and Autocorrelated Covariance Matrices," *Econometica*, 68, 407–423.

Deaton, A. S. (1982), "Model selection procedures, or, does the consumption function exist?" in *Evaluating the reliability of macro-economic models*, eds. Chow, G. and Corsi, P., New York: Wiley, pp. 43–65.

Elyasiani, E. and Nasseh, A. (1994), "The Appropriate Scale Variable in the U.S. Money Demand: An Application of Nonnested Tests of Consumption versus Income Measures," *Journal of Business & Economic Statistics*, 12, 47–55.

Fan, Y. and Li, Q. (1995), "Bootstrapping J-Type Tests for Non-Nested Regression Models," *Economics Letters*, 48, 107–112.

Fisher, G. and McAleer, M. (1981), "Alternative procedures and associated tests of significance for non-nested hypotheses," *Journal of Econometrics*, 16, 103–119.

Godfrey, L. (1998), "Tests of non-nested regression models: some results on small sample behaviour and the bootstrap," *Journal of Econometrics*, 84, 59–74.

Godfrey, L. and Pesaran, M. (1983), "Tests of non-nested regression models: small sample adjustments and Monte Carlo evidence," *Journal of Econometrics*, 21, 133–154.

Goldfeld, S. M. and Quandt, R. E. (1972), *Nonlinear Methods in Econometrics*, Amsterdam-London: North-Holland Publishing Company.

Gonçalves, S. and Vogelsang, T. J. (2006), "Block Bootstrap HAC Robust Tests: The Sophistication of the Naive Bootstrap," *Working paper, Université de Montréal and Cornell University.*

Hansen, B. E. (1992), "Consistent Covariance Matrix Estimation for Dependent Heterogeneous Processes," *Econometrica*, 60, 967–972.

Kiefer, N. M. and Vogelsang, T. J. (2002a), "Heteroskedasticity-Autocorrelation Robust Standard Errors Using the Bartlett Kernel Without Truncation," *Econometrica*, 70, 2093–2095.

— (2002b), "Heteroskedasticity-Autocorrelation Robust Testing Using Bandwidth Equal to Sample Size," *Econometric Theory*, 18, 1350–1366.

— (2005), "A New Asymptotic Theory for Heteroskedasticity-Autocorrelation Robust Tests," *Econometric Theory*, 21, 1130–1164.

Kiefer, N. M., Vogelsang, T. J., and Bunzel, H. (2000), "Simple Robust Testing of Regression Hypotheses," *Econometrica*, 68, 695–714.

Ligeralde, A. V. and Brown, B. W. (1995), "Band Covariance Matrix Estimation Using Restricted Residuals: A Monte Carlo Analysis," *International Economic Review*, 36, 751–767.

Ling, S. (1999), "On the Probabilistic Properties of a Double Threshold ARMA Conditional Heteroskedasticity Model," *Journal of Applied Probability*, 36, 688–705.

Ling, S. and McAleer, M. (2002), "Necessary and Sufficient Moment Conditions for the GARCH(r, s) and Asymmetric Power GARCH(r, s) Models," *Econometric Theory*, 18, 722–729.

Mankiw, N. G. and Summers, L. H. (1986), "Money Demand and the Effects of Fiscal Policies," *Journal of Money, Credit and Banking*, 18, 415–429.

McAleer, M. (1995), "The significance of testing empirical non-nested models," *Journal of Econometrics*, 67, 149–171.

Michelis, L. (1999), "The distributions of the J and Cox non-nested tests in regression models with weakly correlated regressors," *Journal of Econometrics*, 93, 369–401.

Pesaran, M. H. (1974), "On the general problem of model selection," *Review of Economic Studies*, 41, 153–171.

Phillips, P. C. B. and Durlauf, S. N. (1986), "Multiple Time Series Regression with Integrated Processes," *The Review of Economic Studies*, 53, 473–495.

Quandt, R. E. (1974), "A Comparison of Methods for Testing Non-Nested Hypotheses," *Review of Economics and Statistics*, 56, 92–99.

A Tables

Strong $\{x_t, z_t\}$ and $\alpha = 0$								
	J_D				F_D			
	SN	Fixed-b	Boot(1)	Boot(5)	Chi-sq	Fixed-b	Boot(1)	Boot(5)
$M = 1$	0.1250	0.1136	0.0490	0.0478	0.0892	0.0658	0.0512	0.0500
2	0.1308	0.1072	0.0476	0.0470	0.1112	0.0544	0.0512	0.0512
3	0.1396	0.1032	0.0470	0.0484	0.1364	0.0522	0.0502	0.0490
4	0.1452	0.0996	0.0488	0.0510	0.1738	0.0492	0.0490	0.0500
5	0.1538	0.0966	0.0480	0.0514	0.2080	0.0468	0.0476	0.0490
6	0.1618	0.0948	0.0490	0.0494	0.2428	0.0450	0.0494	0.0504
7	0.1706	0.0944	0.0490	0.0500	0.2776	0.0466	0.0496	0.0486
8	0.1760	0.0922	0.0490	0.0494	0.3132	0.0450	0.0478	0.0480
9	0.1840	0.0920	0.0480	0.0492	0.3484	0.0448	0.0452	0.0472
10	0.1926	0.0902	0.0480	0.0484	0.3820	0.0452	0.0484	0.0480
15	0.2288	0.0864	0.0482	0.0494	0.5350	0.0478	0.0460	0.0464
20	0.2672	0.0830	0.0466	0.0508	0.6470	0.0476	0.0468	0.0468
25	0.3022	0.0828	0.0456	0.0508	0.7260	0.0470	0.0504	0.0492
30	0.3334	0.0836	0.0470	0.0500	0.7770	0.0484	0.0488	0.0486
35	0.3664	0.0814	0.0468	0.0490	0.8160	0.0484	0.0480	0.0500
40	0.3986	0.0808	0.0480	0.0500	0.8422	0.0488	0.0486	0.0492
45	0.4266	0.0816	0.0466	0.0502	0.8666	0.0474	0.0488	0.0492
50	0.4482	0.0816	0.0466	0.0496	0.8864	0.0476	0.0482	0.0500

Strong $\{x_t, z_t\}$ and $\alpha = 0.5$								
	J_D				F_D			
	SN	Fixed-b	Boot(1)	Boot(5)	Chi-sq	Fixed-b	Boot(1)	Boot(5)
$M = 1$	0.1138	0.1024	0.0540	0.0576	0.0632	0.0454	0.0368	0.0484
2	0.1198	0.0960	0.0514	0.0526	0.0854	0.0426	0.0400	0.0468
3	0.1250	0.0912	0.0504	0.0512	0.1070	0.0406	0.0374	0.0458
4	0.1364	0.0874	0.0502	0.0514	0.1384	0.0352	0.0346	0.0436
5	0.1426	0.0864	0.0506	0.0530	0.1760	0.0348	0.0328	0.0414
6	0.1492	0.0848	0.0510	0.0528	0.2096	0.0346	0.0368	0.0414
7	0.1580	0.0834	0.0536	0.0528	0.2442	0.0376	0.0372	0.0434
8	0.1660	0.0822	0.0568	0.0550	0.2810	0.0342	0.0376	0.0432
9	0.1734	0.0824	0.0560	0.0536	0.3158	0.0338	0.0370	0.0404
10	0.1808	0.0798	0.0566	0.0532	0.3498	0.0324	0.0350	0.0398
15	0.2176	0.0802	0.0552	0.0542	0.5008	0.0372	0.0364	0.0434
20	0.2548	0.0796	0.0546	0.0516	0.6124	0.0368	0.0358	0.0410
25	0.2868	0.0766	0.0528	0.0522	0.6948	0.0370	0.0382	0.0444
30	0.3214	0.0810	0.0550	0.0546	0.7574	0.0334	0.0334	0.0384
35	0.3498	0.0794	0.0542	0.0542	0.7964	0.0352	0.0368	0.0414
40	0.3810	0.0780	0.0558	0.0538	0.8250	0.0366	0.0368	0.0430
45	0.4068	0.0780	0.0572	0.0544	0.8492	0.0366	0.0366	0.0420
50	0.4322	0.0782	0.0566	0.0552	0.8668	0.0362	0.0388	0.0422

Table 1: Size Comparison (CASE I, level=0.05). x_t and z_t are strongly correlated, and α is the AR(1) coefficient of errors

Strong $\{x_t, z_t\}$ and $\alpha = 0.9$								
	J_D				F_D			
	SN	Fixed-b	Boot(1)	Boot(5)	Chi-sq	Fixed-b	Boot(1)	Boot(5)
$M = 1$	0.1494	0.1364	0.0644	0.0548	0.0734	0.0560	0.0466	0.0572
2	0.1516	0.1278	0.0624	0.0542	0.0892	0.0454	0.0404	0.0500
3	0.1500	0.1072	0.0564	0.0500	0.0936	0.0326	0.0296	0.0416
4	0.1560	0.1020	0.0552	0.0476	0.1168	0.0270	0.0268	0.0380
5	0.1580	0.0976	0.0538	0.0458	0.1376	0.0222	0.0230	0.0346
6	0.1684	0.0936	0.0556	0.0446	0.1676	0.0198	0.0200	0.0324
7	0.1760	0.0910	0.0542	0.0444	0.1958	0.0210	0.0208	0.0340
8	0.1854	0.0880	0.0550	0.0454	0.2284	0.0202	0.0212	0.0324
9	0.1940	0.0870	0.0560	0.0464	0.2588	0.0202	0.0210	0.0324
10	0.2042	0.0882	0.0574	0.0468	0.2928	0.0188	0.0204	0.0306
15	0.2532	0.0832	0.0598	0.0472	0.4576	0.0200	0.0202	0.0286
20	0.2982	0.0866	0.0586	0.0478	0.5796	0.0220	0.0212	0.0336
25	0.3350	0.0852	0.0556	0.0476	0.6658	0.0210	0.0202	0.0316
30	0.3708	0.0820	0.0572	0.0460	0.7274	0.0206	0.0204	0.0320
35	0.4054	0.0790	0.0578	0.0478	0.7780	0.0180	0.0194	0.0314
40	0.4336	0.0814	0.0554	0.0476	0.8064	0.0196	0.0196	0.0306
45	0.4558	0.0832	0.0550	0.0480	0.8354	0.0196	0.0206	0.0328
50	0.4812	0.0844	0.0550	0.0476	0.8608	0.0202	0.0210	0.0306

Strong $\{x_t, z_t\}$ and $\alpha = -0.5$								
	J_D				F_D			
	SN	Fixed-b	Boot(1)	Boot(5)	Chi-sq	Fixed-b	Boot(1)	Boot(5)
$M = 1$	0.2404	0.2228	0.0838	0.0574	0.2050	0.1654	0.1428	0.0702
2	0.2274	0.1902	0.0684	0.0510	0.1842	0.0984	0.0878	0.0594
3	0.2220	0.1728	0.0574	0.0482	0.1984	0.0774	0.0746	0.0552
4	0.2232	0.1606	0.0540	0.0490	0.2240	0.0708	0.0686	0.0524
5	0.2300	0.1484	0.0522	0.0494	0.2538	0.0676	0.0666	0.0552
6	0.2386	0.1424	0.0496	0.0466	0.2934	0.0632	0.0654	0.0560
7	0.2454	0.1380	0.0466	0.0448	0.3294	0.0612	0.0626	0.0536
8	0.2534	0.1328	0.0470	0.0448	0.3664	0.0612	0.0610	0.0518
9	0.2636	0.1274	0.0466	0.0432	0.4032	0.0612	0.0604	0.0536
10	0.2736	0.1262	0.0456	0.0436	0.4402	0.0624	0.0608	0.0522
15	0.3084	0.1194	0.0458	0.0436	0.5918	0.0640	0.0626	0.0520
20	0.3484	0.1180	0.0462	0.0428	0.6970	0.0636	0.0612	0.0512
25	0.3912	0.1136	0.0464	0.0434	0.7708	0.0612	0.0604	0.0542
30	0.4258	0.1130	0.0450	0.0416	0.8170	0.0606	0.0602	0.0496
35	0.4558	0.1134	0.0454	0.0424	0.8468	0.0602	0.0592	0.0528
40	0.4836	0.1136	0.0454	0.0424	0.8690	0.0584	0.0584	0.0510
45	0.5126	0.1132	0.0448	0.0430	0.8910	0.0616	0.0612	0.0516
50	0.5392	0.1142	0.0448	0.0428	0.9100	0.0614	0.0598	0.0518

Table 2: Size Comparison (CASE I, level=0.05). x_t and z_t are strongly correlated, and α is the AR(1) coefficient of errors

Weak $\{x_t, z_t\}$ and $\alpha = 0$								
	J_D				F_D			
	SN	Fixed-b	Boot(1)	Boot(5)	Chi-sq	Fixed-b	Boot(1)	Boot(5)
$M = 1$	0.0924	0.0826	0.0480	0.0520	0.0886	0.0644	0.0514	0.0562
2	0.0992	0.0838	0.0474	0.0532	0.1084	0.0514	0.0470	0.0556
3	0.1092	0.0778	0.0476	0.0508	0.1392	0.0456	0.0472	0.0556
4	0.1190	0.0794	0.0438	0.0514	0.1724	0.0418	0.0434	0.0500
5	0.1272	0.0770	0.0446	0.0516	0.2106	0.0394	0.0438	0.0494
6	0.1350	0.0760	0.0450	0.0508	0.2482	0.0414	0.0442	0.0486
7	0.1436	0.0724	0.0452	0.0506	0.2848	0.0418	0.0470	0.0504
8	0.1506	0.0712	0.0438	0.0488	0.3236	0.0420	0.0446	0.0522
9	0.1586	0.0696	0.0436	0.0488	0.3582	0.0424	0.0450	0.0514
10	0.1666	0.0678	0.0432	0.0474	0.3940	0.0410	0.0442	0.0494
15	0.2038	0.0664	0.0414	0.0442	0.5470	0.0422	0.0432	0.0468
20	0.2386	0.0638	0.0416	0.0442	0.6640	0.0424	0.0430	0.0504
25	0.2668	0.0624	0.0406	0.0440	0.7410	0.0446	0.0450	0.0490
30	0.2972	0.0654	0.0426	0.0446	0.7918	0.0442	0.0462	0.0496
35	0.3300	0.0642	0.0422	0.0440	0.8354	0.0436	0.0436	0.0492
40	0.3550	0.0644	0.0432	0.0446	0.8616	0.0432	0.0434	0.0482
45	0.3828	0.0642	0.0426	0.0438	0.8838	0.0426	0.0448	0.0492
50	0.4052	0.0642	0.0426	0.0442	0.8994	0.0434	0.0446	0.0496
Weak $\{x_t, z_t\}$ and $\alpha = 0.5$								
	J_D				F_D			
	SN	Fixed-b	Boot(1)	Boot(5)	Chi-sq	Fixed-b	Boot(1)	Boot(5)
$M = 1$	0.2736	0.2560	0.1648	0.0772	0.3770	0.3200	0.2816	0.0968
2	0.2228	0.1894	0.1114	0.0664	0.2888	0.1746	0.1664	0.0806
3	0.2088	0.1634	0.0886	0.0590	0.2844	0.1270	0.1240	0.0704
4	0.2050	0.1490	0.0786	0.0550	0.3072	0.1038	0.1058	0.0660
5	0.2078	0.1386	0.0734	0.0522	0.3348	0.0970	0.1004	0.0650
6	0.2150	0.1324	0.0690	0.0506	0.3676	0.0896	0.0956	0.0624
7	0.2218	0.1268	0.0664	0.0506	0.4004	0.0866	0.0912	0.0654
8	0.2262	0.1226	0.0638	0.0498	0.4402	0.0808	0.0858	0.0588
9	0.2352	0.1184	0.0628	0.0504	0.4728	0.0806	0.0832	0.0594
10	0.2450	0.1146	0.0602	0.0482	0.5066	0.0784	0.0864	0.0574
15	0.2882	0.1126	0.0606	0.0484	0.6474	0.0862	0.0908	0.0584
20	0.3240	0.1108	0.0598	0.0500	0.7480	0.0890	0.0902	0.0580
25	0.3610	0.1108	0.0612	0.0512	0.8152	0.0900	0.0936	0.0612
30	0.3942	0.1122	0.0604	0.0526	0.8542	0.0900	0.0950	0.0664
35	0.4218	0.1126	0.0586	0.0528	0.8864	0.0898	0.0944	0.0640
40	0.4464	0.1094	0.0572	0.0512	0.9058	0.0904	0.0912	0.0626
45	0.4744	0.1092	0.0562	0.0538	0.9240	0.0906	0.0918	0.0620
50	0.4992	0.1110	0.0566	0.0532	0.9370	0.0916	0.0912	0.0612

Table 3: Size Comparison (CASE I, level=0.05). x_t and z_t are weakly correlated, and α is the AR(1) coefficient of errors

Weak $\{x_t, z_t\}$ and $\alpha = 0.9$								
	J_D				F_D			
	SN	Fixed-b	Boot(1)	Boot(5)	Chi-sq	Fixed-b	Boot(1)	Boot(5)
$M = 1$	0.6572	0.6416	0.4810	0.1554	0.7560	0.7170	0.6860	0.2318
2	0.5770	0.5382	0.3174	0.1196	0.6054	0.4696	0.4494	0.1674
3	0.5312	0.4696	0.2250	0.0970	0.5294	0.3148	0.3128	0.1302
4	0.5100	0.4162	0.1784	0.0800	0.5066	0.2314	0.2308	0.1036
5	0.4986	0.3768	0.1482	0.0714	0.5172	0.1874	0.1954	0.0912
6	0.4938	0.3478	0.1310	0.0658	0.5342	0.1640	0.1750	0.0842
7	0.4934	0.3260	0.1188	0.0632	0.5634	0.1504	0.1594	0.0774
8	0.4946	0.3068	0.1088	0.0598	0.5958	0.1438	0.1478	0.0796
9	0.4984	0.2912	0.1024	0.0568	0.6254	0.1426	0.1464	0.0782
10	0.5038	0.2814	0.1012	0.0580	0.6542	0.1410	0.1454	0.0770
15	0.5384	0.2608	0.0976	0.0576	0.7748	0.1526	0.1528	0.0780
20	0.5728	0.2524	0.1018	0.0584	0.8512	0.1528	0.1528	0.0822
25	0.6024	0.2466	0.1056	0.0626	0.8934	0.1574	0.1584	0.0806
30	0.6352	0.2426	0.1046	0.0620	0.9208	0.1534	0.1570	0.0786
35	0.6622	0.2428	0.0968	0.0614	0.9400	0.1556	0.1550	0.0828
40	0.6884	0.2400	0.0982	0.0612	0.9520	0.1522	0.1516	0.0780
45	0.7100	0.2418	0.1002	0.0582	0.9580	0.1474	0.1474	0.0776
50	0.7316	0.2430	0.1004	0.0588	0.9670	0.1484	0.1476	0.0766

Weak $\{x_t, z_t\}$ and $\alpha = -0.5$								
	J_D				F_D			
	SN	Fixed-b	Boot(1)	Boot(5)	Chi-sq	Fixed-b	Boot(1)	Boot(5)
$M = 1$	0.0236	0.0190	0.0170	0.0392	0.0130	0.0076	0.0058	0.0368
2	0.0472	0.0356	0.0296	0.0478	0.0380	0.0154	0.0152	0.0458
3	0.0610	0.0408	0.0332	0.0480	0.0616	0.0160	0.0174	0.0428
4	0.0708	0.0430	0.0342	0.0484	0.0912	0.0150	0.0176	0.0466
5	0.0802	0.0446	0.0384	0.0472	0.1230	0.0168	0.0194	0.0452
6	0.0880	0.0456	0.0398	0.0484	0.1506	0.0162	0.0196	0.0440
7	0.0958	0.0458	0.0410	0.0472	0.1838	0.0172	0.0188	0.0448
8	0.1030	0.0472	0.0394	0.0454	0.2182	0.0184	0.0220	0.0448
9	0.1098	0.0474	0.0402	0.0462	0.2508	0.0204	0.0224	0.0424
10	0.1170	0.0484	0.0388	0.0460	0.2898	0.0196	0.0216	0.0432
15	0.1552	0.0466	0.0416	0.0448	0.4396	0.0170	0.0184	0.0406
20	0.1924	0.0472	0.0394	0.0418	0.5596	0.0196	0.0200	0.0402
25	0.2252	0.0462	0.0394	0.0418	0.6486	0.0190	0.0198	0.0404
30	0.2610	0.0464	0.0396	0.0446	0.7192	0.0210	0.0218	0.0412
35	0.2922	0.0462	0.0394	0.0440	0.7638	0.0202	0.0204	0.0410
40	0.3202	0.0468	0.0392	0.0430	0.8000	0.0190	0.0194	0.0426
45	0.3466	0.0474	0.0400	0.0444	0.8320	0.0188	0.0194	0.0410
50	0.3714	0.0472	0.0394	0.0448	0.8514	0.0192	0.0212	0.0408

Table 4: Size Comparison (CASE I, level=0.05). x_t and z_t are weakly correlated, and α is the AR(1) coefficient of errors

$\delta = 0.5$								
	J_D				F_D			
	SN	Fixed-b	Boot(1)	Boot(5)	Chi-sq	Fixed-b	Boot(1)	Boot(5)
$M = 1$	0.1304	0.1182	0.0500	0.0476	0.0866	0.0622	0.0480	0.0450
2	0.1380	0.1108	0.0512	0.0478	0.1070	0.0562	0.0460	0.0456
3	0.1446	0.1080	0.0482	0.0488	0.1366	0.0502	0.0448	0.0458
4	0.1494	0.1036	0.0468	0.0480	0.1706	0.0488	0.0456	0.0444
5	0.1584	0.1024	0.0480	0.0466	0.2102	0.0470	0.0442	0.0446
6	0.1664	0.1002	0.0468	0.0472	0.2442	0.0466	0.0442	0.0458
7	0.1744	0.0972	0.0472	0.0466	0.2794	0.0456	0.0454	0.0446
8	0.1832	0.0964	0.0448	0.0450	0.3134	0.0452	0.0454	0.0450
9	0.1926	0.0938	0.0456	0.0456	0.3468	0.0438	0.0444	0.0442
10	0.1980	0.0926	0.0454	0.0460	0.3820	0.0444	0.0438	0.0438
15	0.2370	0.0896	0.0498	0.0466	0.5272	0.0458	0.0430	0.0432
20	0.2728	0.0882	0.0478	0.0466	0.6420	0.0446	0.0424	0.0450
25	0.3070	0.0904	0.0484	0.0446	0.7222	0.0434	0.0408	0.0424
30	0.3428	0.0884	0.0464	0.0442	0.7724	0.0434	0.0424	0.0418
35	0.3700	0.0860	0.0494	0.0466	0.8130	0.0460	0.0426	0.0416
40	0.3980	0.0850	0.0480	0.0456	0.8416	0.0456	0.0428	0.0416
45	0.4260	0.0854	0.0478	0.0458	0.8680	0.0446	0.0416	0.0438
50	0.4438	0.0864	0.0482	0.0456	0.8848	0.0452	0.0434	0.0416

$\delta = 0.9$								
	J_D				F_D			
	SN	Fixed-b	Boot(1)	Boot(5)	Chi-sq	Fixed-b	Boot(1)	Boot(5)
$M = 1$	0.1392	0.1260	0.0536	0.0528	0.0938	0.0686	0.0498	0.0492
2	0.1460	0.1174	0.0508	0.0520	0.1142	0.0566	0.0480	0.0482
3	0.1496	0.1138	0.0498	0.0498	0.1392	0.0536	0.0490	0.0456
4	0.1578	0.1106	0.0492	0.0492	0.1754	0.0484	0.0478	0.0466
5	0.1646	0.1046	0.0482	0.0492	0.2110	0.0482	0.0484	0.0478
6	0.1750	0.1028	0.0468	0.0490	0.2506	0.0468	0.0474	0.0470
7	0.1814	0.0998	0.0468	0.0508	0.2880	0.0486	0.0460	0.0470
8	0.1886	0.0984	0.0474	0.0490	0.3262	0.0472	0.0462	0.0468
9	0.1960	0.0964	0.0480	0.0482	0.3630	0.0454	0.0456	0.0452
10	0.2028	0.0976	0.0488	0.0480	0.3952	0.0462	0.0450	0.0458
15	0.2402	0.0930	0.0486	0.0476	0.5486	0.0474	0.0454	0.0428
20	0.2796	0.0894	0.0476	0.0454	0.6484	0.0458	0.0458	0.0434
25	0.3126	0.0862	0.0456	0.0458	0.7308	0.0478	0.0436	0.0442
30	0.3448	0.0878	0.0474	0.0458	0.7902	0.0446	0.0440	0.0448
35	0.3744	0.0880	0.0456	0.0472	0.8246	0.0470	0.0450	0.0446
40	0.4020	0.0862	0.0450	0.0472	0.8484	0.0456	0.0442	0.0428
45	0.4286	0.0872	0.0448	0.0478	0.8682	0.0480	0.0434	0.0450
50	0.4514	0.0878	0.0452	0.0488	0.8858	0.0474	0.0444	0.0438

Table 5: Size Comparison (CASE II, level=0.05). δ is the AR(1) coefficient of y_{t-1}

	Strong $\{x_t, z_t\}$				Weak $\{x_t, z_t\}$			
	$\alpha = 0$		$\alpha = 0.9$		$\alpha = 0$		$\alpha = 0.9$	
	J_D	F_D	J_D	F_D	J_D	F_D	J_D	F_D
$M = 1$	0.9962	0.9756	0.8746	0.7430	1.0000	0.9994	0.8778	0.7458
2	0.9948	0.9562	0.9172	0.8272	1.0000	0.9914	0.9208	0.8346
3	0.9918	0.9300	0.9174	0.8164	0.9990	0.9652	0.9200	0.8214
4	0.9874	0.8992	0.9204	0.8108	0.9976	0.9272	0.9260	0.8144
5	0.9820	0.8724	0.9128	0.7914	0.9916	0.8792	0.9202	0.7954
6	0.9742	0.8378	0.9006	0.7754	0.9850	0.8394	0.9104	0.7784
7	0.9664	0.7906	0.8902	0.7570	0.9786	0.8010	0.9018	0.7606
8	0.9580	0.7726	0.8858	0.7330	0.9728	0.7672	0.8962	0.7346
9	0.9496	0.7500	0.8746	0.7166	0.9656	0.7452	0.8870	0.7158
10	0.9380	0.7218	0.8686	0.7002	0.9600	0.7246	0.8780	0.6988
15	0.8812	0.6694	0.8244	0.6586	0.9116	0.6704	0.8308	0.6566
20	0.8348	0.6678	0.7692	0.6532	0.8636	0.6784	0.7744	0.6430
25	0.8050	0.6710	0.7448	0.6536	0.8314	0.6700	0.7428	0.6496
30	0.7988	0.6748	0.7240	0.6564	0.8184	0.6850	0.7276	0.6502
35	0.7904	0.6854	0.7216	0.6546	0.8122	0.6922	0.7274	0.6412
40	0.7858	0.6778	0.7240	0.6532	0.8026	0.6868	0.7292	0.6462
45	0.7800	0.6862	0.7208	0.6556	0.8022	0.6910	0.7252	0.6514
50	0.7852	0.6916	0.7210	0.6646	0.8024	0.6948	0.7240	0.6610
	$\alpha = 0.5$		$\alpha = -0.5$		$\alpha = 0.5$		$\alpha = -0.5$	
	J_D	F_D	J_D	F_D	J_D	F_D	J_D	F_D
$M = 1$	0.9918	0.9700	0.9652	0.8534	0.9934	0.9712	0.9620	0.8510
2	0.9936	0.9594	0.9540	0.8538	0.9946	0.9626	0.9506	0.8470
3	0.9924	0.9434	0.9398	0.8238	0.9924	0.9436	0.9374	0.8168
4	0.9896	0.9246	0.9220	0.7816	0.9908	0.9222	0.9212	0.7736
5	0.9874	0.8968	0.9022	0.7400	0.9876	0.8912	0.9040	0.7290
6	0.9814	0.8638	0.8924	0.6918	0.9810	0.8570	0.8904	0.6882
7	0.9748	0.8342	0.8734	0.6612	0.9758	0.8288	0.8716	0.6560
8	0.9700	0.8022	0.8540	0.6370	0.9686	0.7980	0.8562	0.6338
9	0.9630	0.7840	0.8330	0.6160	0.9632	0.7808	0.8362	0.6110
10	0.9516	0.7694	0.8168	0.5858	0.9512	0.7670	0.8190	0.5806
15	0.9066	0.7086	0.7446	0.5416	0.9096	0.7076	0.7464	0.5434
20	0.8552	0.7160	0.6748	0.5390	0.8588	0.7078	0.6744	0.5520
25	0.8252	0.7142	0.6372	0.5602	0.8268	0.7102	0.6382	0.5654
30	0.8138	0.7332	0.6384	0.5664	0.8170	0.7260	0.6362	0.5636
35	0.8052	0.7278	0.6362	0.5550	0.8056	0.7146	0.6364	0.5540
40	0.7996	0.7246	0.6358	0.5628	0.7964	0.7174	0.6388	0.5608
45	0.8038	0.7260	0.6344	0.5604	0.8008	0.7172	0.6374	0.5640
50	0.8062	0.7368	0.6376	0.5624	0.8040	0.7254	0.6418	0.5658

Table 6: Power Comparison (CASE I). Size is controlled to be 0.05. α is the AR(1) coefficient of the errors.

	$\delta = 0.5$		$\delta = 0.9$	
	J_D	F_D	J_D	F_D
$M = 1$	0.9956	0.9746	0.9956	0.9772
2	0.9918	0.9450	0.9930	0.9566
3	0.9876	0.9188	0.9904	0.9300
4	0.9806	0.8846	0.9844	0.9038
5	0.9732	0.8478	0.9786	0.8662
6	0.9612	0.8072	0.9680	0.8254
7	0.9476	0.7732	0.9612	0.7860
8	0.9388	0.7398	0.9468	0.7674
9	0.9270	0.7248	0.9378	0.7478
10	0.9142	0.6952	0.9296	0.7274
15	0.8542	0.6544	0.8826	0.6728
20	0.7964	0.6522	0.8256	0.6616
25	0.7714	0.6740	0.7848	0.6758
30	0.7446	0.6654	0.7612	0.6862
35	0.7384	0.6622	0.7596	0.6792
40	0.7430	0.6634	0.7654	0.6798
45	0.7488	0.6704	0.7648	0.6802
50	0.7532	0.6712	0.7656	0.6816

Table 7: Power Comparison (CASE II). Size is controlled to be 0.05. δ is the AR(1) coefficient of y_{t-1}

	z	
Regressor	GNP	Consumption (Expenditure)
constant	0.0024 (2.2225**)	−0.0016 (−1.3040)
r_t	−0.0164 (−3.4405**)	−0.0135 (−3.0514**)
r_{t-1}	−0.0226 (−4.6843**)	−0.0206 (−4.5532**)
r_{t-2}	−0.0092 (−1.9734*)	−0.0071 (−1.6712*)
z_t	0.3041 (4.0340**)	0.4363 (5.0398**)
z_{t-1}	0.2067 (2.6866**)	0.3836 (4.2471**)
z_{t-2}	0.1268 (1.6493)	0.2239 (2.4932**)
R^2	0.29239	0.40581
$\bar{R}^2$	0.2684	0.38567
DW	1.0896	1.2254

** significant at the 5% level, * significant at the 10% level.

Table 8: Money (M2) demand function estimation with a consumption measure and an income measure. (Quarterly data from 1959.I to 2005.III)

B Figures

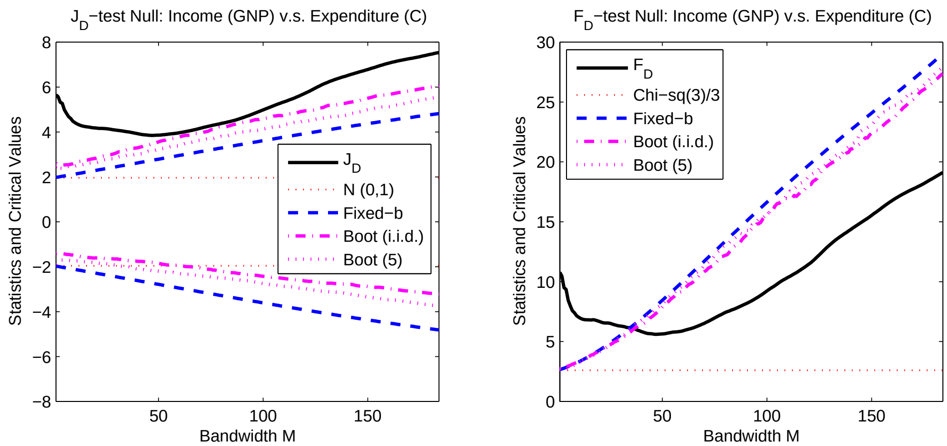

Figure 1: Testing the null of income for the scale variable for U.S. money demand. The solid lines are J_D and F_D statistics for different bandwidths and the other lines show the critical values from different asymptotic approximations.

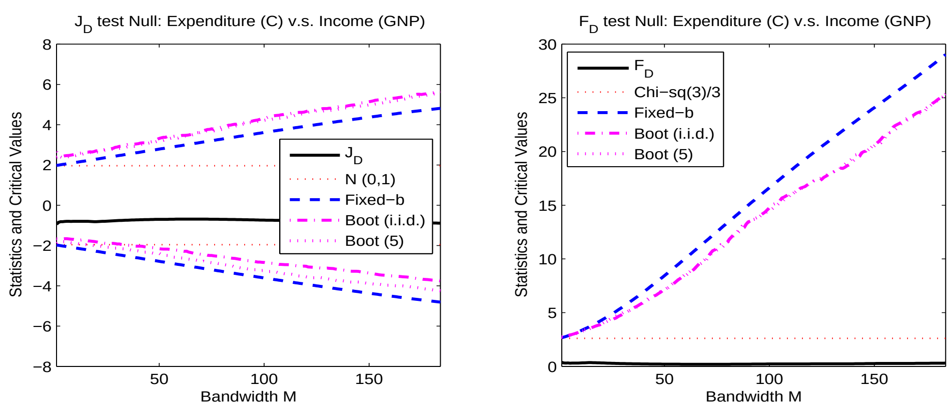

Figure 2: Testing the null of consumption for the scale variable for U.S. money demand. The solid lines are J_D and F_D statistics for different bandwidths and the other lines show the critical values from different asymptotic approximations.