The Long Memory Autoregressive Distributed Lag Model and Its Application on Congressional Approval

Wen-Jen Tsay*

March 24, 2009

The Institute of Economics, Academia Sinica, Taiwan

Abstract

This paper considers the instrumental variables (IV) estimation of the autoregressive distributed lag (ADL) model consisting of fractionally integrated regressors and errors, while allowing for part of the regressors to be endogenous. The idea of Liviatan (1963) and that of Tsay (2007) are combined to construct consistent and asymptotically normally distributed multiple-differenced two-stage-least-squares (MD-TSLS) and MD generalized method of moments (MD-GMM) estimators for the long memory ADL model. The simulations show that the performance of the MD-GMM estimator is especially excellent even though the sample size is 100. The IV estimators are applied to the data of Durr, Gilmour, and Wolbrecht (1997) on Congressional approval. As compared to the 0.08 estimate of the long-run effect of presidential approval on Congressional approval based on the scalar ADL model of De Boef and Keele (2008), a stronger support for the divided party government hypothesis is found for a class of the vector ADL model which generates a corresponding long-run impact equal to 0.26 or higher.

Key words: Autoregressive distributed lag model; Stochastic linear difference equation; Long memory; Lagged dependent variable; Instrumental variables

*Wen-Jen Tsay: The Institute of Economics, Academia Sinica. 128 Academia Road, Sec. 2, Taipei, Taiwan, R.O.C. Tel.: (886-2) 2782-2791 ext. 296. Fax: (886-2) 2785-3946. Correspondence: wtsay@econ.sinica.edu.tw

1 Introduction

It is well documented in Hendry (1995) that a dynamic analysis should start with a general model. This point has also been clearly spelled out in De Boef and Keele (2008) who emphasize that restrictive specifications might be adopted without evidence that restrictions are valid if the analysts do not start with a general model. In the context of stationary data and weakly exogenous regressors, De Boef and Keele (2008) highly recommend that the following autoregressive distributed lag or the ADL(p, k, l) model fits the criteria of Hendry (1995):

$$y_t = \alpha_0 + \sum_{i=1}^{p} \rho_i y_{t-i} + \sum_{j=1}^{k} \sum_{i=0}^{l} \beta_{ji} x_{j,t-i} + \varepsilon_t = \alpha_0 + \sum_{i=1}^{p} \rho_i y_{t-i} + f_t^\top \beta + \varepsilon_t, \tag{1}$$

where $x_{j,t}$ are weakly exogenous such that $\mathrm{E}(x_{j,s}\varepsilon_t) = 0$ for all j, s, and t. The parameter $\gamma^\top = (\rho^\top, \beta^\top)$, where $\rho = (\rho_1, \ldots, \rho_p)^\top$, is the focus of analysis, because it can be used to calculate the short-run or impact multipliers, the long-run effect or total multipliers, and other important quantities of interest to political scientists. This is the second reason why De Boef and Keele (2008) recommend the ADL model for empirical applications, because it is flexible in relating the short-run and long-run influences of the regressors on dependent variable y_t. When searching for articles published in *The American Political Science Review*, *The American Journal of Political Science*, and *The Journal of Politics* between 1995 and 2005, De Boef and Keele (2008) document that 73 articles use time series regression in the context of stationary data, indicating that the ADL model helps political scientists obtain a better understanding of political dynamics.

The ADL model has also been extensively considered in the economic literature. One example noted in Liviatan (1963) is the distributed lag model used by Koyck (1954) in his empirical analysis of investment:

$$I_t = aQ_t + b \sum_{i=0}^{\infty} \lambda^i Q_{t-1-i} + c_t, \tag{2}$$

where I_t denotes the aggregate investment and Q_t is the aggregate output. Defining

$r_t = c_t - \lambda c_{t-1}$, the model in (2) can be rewritten as:

$$I_t = aQ_t + (b - \lambda a)\, Q_{t-1} + \lambda I_{t-1} + r_t. \tag{3}$$

This happens to be a special case of the model in (1).

When $l = 0$, the ADL model reduces to the stochastic linear difference equation:

$$y_t = \alpha_0 + \sum_{i=1}^{p} \rho_i y_{t-i} + \sum_{j=1}^{k} \beta_j x_{j,t} + \varepsilon_t = \alpha_0 + \sum_{i=1}^{p} \rho_i y_{t-i} + x_t^\top \beta + \varepsilon_t, \tag{4}$$

where x_t is a $(k \times 1)$ vector of the stochastic regressors. The stochastic linear difference equation arises as a consequence of a process of partial adjustment or adaptive expectation as explained in Harvey (1990). Although the stochastic linear difference equation is a special case of the ADL model, the great potential in the stochastic linear difference equation for empirical applications and the theoretical interest in its own right stimulate a large number of estimation methods, including Durbin (1959), Walker (1964), Amemiya and Fuller (1967), Box and Jenkins (1970), Hannan and Nicholls (1972), and Hatanaka (1974, 1976), to name a few. Nicholls et al. (1975) and Reinsel (1979) provide overviews about the estimation of this model. Durr, Gilmour, and Wolbrecht (DGW hereafter, 1997) apply the stochastic linear difference equation to explain Congressional approval.

The condition $\sum_{i=1}^{p} \rho_i < 1$ is regularly imposed in the ADL literature to ensure that the dependent variable y_t is stationary. The other important necessary condition for the stationarity of y_t often neglected in the literature is that the stochastic regressor $x_{j,t}$ must be stationary for all $j = 1, \ldots, k$ as well. The model satisfying these two conditions is named the stationary ADL model. If ε_t is further assumed to be serially uncorrelated, then the ordinary least squares (OLS) estimator for γ is consistent under the stationary ADL model assumption and suitable regularity conditions.

In the political science literature, Keele and Kelly (2005) provide several practical suggestions on when and how to use lagged dependent variables (LDV) on the right-hand side (RHS) of the ADL model. Essentially, the aforementioned three restrictions are implicitly or explicitly assumed in the political science literature when conducting empirical studies. One drawback of these conditions is that they rule out the possibility that some of the variables in x_t might be endogenous. These conditions also restrict the applicability of the ADL model to political science analysis, because they do not allow the errors to be serially correlated as found in many time series analysis. In addition, Keele and Keely (2005, p. 203) document that whatever the strengths are of LDV, they are inappropriate with nonstationary data that have not been differenced. Nevertheless, they do not provide any method to deal with this issue.

This paper adopts the suggestion of Keele and Keely (2005) to deal with the nonstationary ADL model based on the multiple-differencing principle of Tsay (2007). Indeed, the coverage of this paper is much broader than that of the nonstationary ADL model considered in Keele and Keely (2005). It is shown that the ADL model can be consistently estimated and effectively tested even though f_t and ε_t in (1) are stationary or nonstationary fractionally integrated, or $I(d)$, processes introduced in Granger (1980), Granger and Joyeux (1980), and Hosking (1981).

The distinguishing feature of the $I(d)$ process is that its autocovariance function declines at a slower hyperbolic rate, rather than the geometric rate found in stationary autoregressive and invertible moving-average (ARMA) processes. When $d > 0$, the $I(d)$ process is often called the long memory process. The $I(d)$ process can display nonstationary phenomenon when $d \geq 1/2$, otherwise, it is covariance stationary. With a continuum of possible values of d, both $I(0)$ and $I(1)$ processes can be viewed as special cases of the $I(d)$ process, where the differencing parameter d is 0 and 1, respectively. The $I(d)$ process thus provides a unified framework for empirical applications. Without a doubt, combining the $I(d)$ processes and the ADL literature greatly enhances the applicability of the ADL model considered in Hendry (1995).

The rationale for considering the long memory ADL model is not only from theoretical interest, but it also closely corresponds to the empirical observations in Lebo et al. (2000, p. 40) who claim that *fractional integration is extremely common among political time series*. In particular, Box-Steffensmeier and Smith (1996, 1998) use the $I(d)$

process to explain the dynamics of aggregate partisanship and macroideology. Byers et al. (2000) observe long memory phenomena in many aggregate political popularity. Dolado et al. (2003) find long memory in Spanish opinion polls, while Lai and Reiter (2005) apply the autoregressive fractionally-integrated moving-average (ARFIMA) model to study foreign policy. Clarke and Lebo (2003) employ the concept of fractional cointegration to investigate political party support in Britain. Likewise, Davidson et al. (2006) apply a fractionally cointegrating vector error correction model to describe the poll data on the performance approval of prime ministers and governments in the UK.

There are two major technical problems associated with the long memory ADL model as found in the long memory time-series-cross-section (TSCS) data of Tsay (2009). First, the presence of stationary long memory regressors and errors can induce spurious regression as discussed in Tsay and Chung (2000). Second, the incorporation of a dynamic adjustment mechanism through LDV further complicates the estimation and inference issues concerning the long memory ADL model. One remedy to resolve these problems is to combine the multiple-difference (MD) method of Tsay (2007) and the intrumental variables (IV) methodology of Liviatan (1963) to construct an MD two-stage-least-squares (MD-TSLS) estimator and an MD generalized method of moments (MD-GMM) counterpart for the long memory ADL model. In particular, the idea of Liviatan (1963) is to employ the lagged stochastic exogenous variables as the IV for the LDV. The methodology developed in this paper can be further generalized to correct the endogenous bias caused by the presence of endogenous variables at the RHS of regression by using the remaining lagged weakly exogenous variables as the IV for the endogenous variables

The MD estimator of Tsay (2007) essentially first-differences both the dependent variable and the regressors M (a non-negative integer) times before running the subsequent IV estimator in order to ensure that the resulting IV estimator is asymptotically normally distributed under suitable regularity conditions. Most importantly, the use of multiple-differencing does not always result in the efficiency loss of the MD estimator as compared to the OLS counterpart as clearly investigated in Tsay (2007, p. 835 and p. 836), even though it can always eliminate the problem of spurious regression associated with long memory processes.

The methodology developed in this paper can be applied to many studies concerning the short-run and long-run influences of regressors on the dependent variable, even though the data can be stationary or nonstationary long memory processes. In this paper we apply the proposed IV estimators to the data of DGW (1997) about Congressional approval where DGW (1997) do not support the divided party government hypothesis as shown in their Table 2. The data of DGW have also been considered in De Boef and Keele (2008). This dataset is chosen, because the $I(d)$ processes have been found useful for modeling Congressional approval and economic expectations in Box-Steffensmeier and Tomlinson (2000). They also check whether these two variables are fractionally cointegrated, implying that Box-Steffensmeier and Tomlinson (2000) believe these two variables are endogenously determined with each other. However, their testing results are not satisfactory as a large standard error of the fractional differencing parameter of the residuals from the fractional cointegration analysis between these two variables is found. Box-Steffensmeier and Tomlinson (2000) explain this large standard error as not unexpected since there are only 80 observations.

The second reason why the dataset of DGW (1997) is chosen is that the analysis in DGW (1997) and that in De Boef and Keele (2008) do not recognize the impacts induced by the potential endogeneity of presidential approval and that of economic expectations on Congressional approval. Based on the findings in De Boef and Kellstedt (2004, p. 648), whereby "consumer confidence should not solely be used as a right-hand-side variable in analysis of the political economy. Politics, in a variety of forms, affects economics," along with the implications derived from the divided party government hypothesis, the RHS variables, economic expectations, and presidential approval in the regression of DGW (1997) should be treated as endogenous. This point of view and the estimation results from the IV estimators suggest that a class of the vector ADL model might be suitable to describe the time series properties of Congressional approval. Interestingly, the results from the vector ADL model support the divided party government hypothesis if the near multicollinearity resulting from the strong similarity between economic expectations and presidential approval is carefully taken into account.

The remaining part of this paper is arranged as follows: Section 2 and Section 3 present the MD-TSLS and MD-GMM estimators and their asymptotic properties, respectively. Section 4 considers the generalization of the IV methods under a more general model specification. The important issue concerning the choice of IV for the general specifications is also taken up in this section. Section 5 verifies the theoretical findings generated from this paper through a Monte Carlo experiment. The data of DGW (1997) about Congressional approval are carefully analyzed in Section 6. Section 7 provides a conclusion.

2 Instrumental Variables Estimator

For expositional purposes, we focus on the stationary long memory ADL model with $p = 1$ and $l = 0$:

$$y_t = \alpha_0 + \rho_1 y_{t-1} + x_t^\top \beta + \varepsilon_t, \quad |\rho_1| < 1, \quad t = 1, \dots, T. \tag{5}$$

The analytical results derived from this model can be easily extended to the ADL model in (1) and the situation that some of the regressors are endogenous.

Define a short memory process $\{\eta_t, t = 0, \pm 1, \dots\}$ as a zero-mean covariance stationary process with a spectral density $f(\lambda)$, which is bounded away from zero. Here, η_t includes the stationary and invertible ARMA process as one of its special cases. With η_t and $d < 1/2$, a stationary $I(d)$ process ξ_t is defined as:

$$\xi_t = \triangle^{-d} \eta_t, \quad t = 0, \pm 1, \dots, \tag{6}$$

where $\triangle = 1 - L$, and L is the usual lag operator. Since the cases where ε_t is an $I(d_\varepsilon)$ process with $d_\varepsilon \geq 0$ are the subject of interest to empirical researchers, the focus of this paper is on the models where $d_\varepsilon \geq 0$. As a consequence, the model in (5) can be displayed as:

$$y_t = \alpha_0 + \rho_1 y_{t-1} + x_t^\top \beta + \varepsilon_t, \quad x_{h,t} = I(d_{x,h}),\ \ \varepsilon_t = I(d_\varepsilon), \quad 0 \leq d_{x,h}, d_\varepsilon < 1/2, \quad |\rho_1| < 1, \tag{7}$$

where $x_{h,t}$ is the h-th element of x_t.

The MD-TSLS estimator proposed in this paper contains three steps. Following Tsay (2007), we multiple difference the data, y_t and x_t, before running the IV estimation. The resulting multiple-differenced transformation is:

$$\triangle^M y_t = \rho_1 \triangle^M y_{t-1} + \triangle^M x_t^\top \beta + \triangle^M \varepsilon_t, \quad t = M+1, M+2, \ldots, T, \tag{8}$$

where M is defined as:

$$M = \begin{cases} g, & \text{if } d_{x,h} + d_\varepsilon < 1/2, \ \text{ for all } h = 1, 2, \ldots, k; \\ g+1, & \text{if } d_{x,h} + d_\varepsilon \geq 1/2, \ \text{ for some } h \in \{1, 2, \ldots, k\}, \end{cases} \tag{9}$$

and g is a non-negative integer. By the Wold decomposition theorem, $\triangle^M x_t$ and $\triangle^M \varepsilon_t$ are surely covariance stationary and represented with the following MA(∞) processes:

$$\triangle^M \varepsilon_t = \sum_{i=0}^{\infty} \psi_i a_{t-i}, \qquad \triangle^M x_t = \sum_{i=0}^{\infty} \varphi_i b_{t-i}, \qquad t = M+1, \ldots, T, \tag{10}$$

where a_t and b_t are assumed to satisfy the conditions in Assumption A of Robinson (1998). This indicates that x_t is weakly exogenous as regularly imposed in the ADL literature.

As will be argued more clearly in the Mathematical Appendix concerning the theoretical foundation of the choice of M in (9), the idea of choosing M for the stationary long memory ADL model is identical to that in equation (5) of Tsay (2007) for the static time series regression model. Essentially, if the regressors and the errors are 'known' to be $I(0)$ processes, then both $M = 0$ and $M = 1$ can be used. Note that the use of $M = 1$ instead of $M = 0$ does not necessarily generate a higher variance of the 'over-differenced' estimator as shown in Tsay (2007). Moreover, if $d_{x,h}, d_\varepsilon \geq 0$, and the sum of $d_{x,h}$ and d_ε is less than 1/2 for each h, then $M = 0$ can be employed. However, this does not imply that $M = 1$ or $M = 2$ cannot be used under this circumstance. On the other hand, if the sum of $d_{x,h}$ and d_ε is greater than 1/2 for some h, then $M \geq 1$ must be imposed to avoid the occurrence of the spurious regression problem discussed in Tsay and Chung (2000).

The second step of the MD-TSLS method derives from the idea of Liviatan (1963) in selecting IV to deal with the joint presence of LDV and the serially correlated disturbance term. Particularly, q lagged values of $\triangle^M x_t$ are employed as the instruments for $\triangle^M y_{t-1}$. By the weakly exogenous regressor assumption that x_t and ε_t are uncorrelated at all leads and lags, $\triangle^M x_{t-1}$, $\triangle^M x_{t-2}, \ldots$, and $\triangle^M x_{t-q}$ are all uncorrelated with $\triangle^M \varepsilon_t$, but are correlated with $\triangle^M y_{t-1}$, thus satisfying the fundamental requirement as the legitimate IV. For clarity of exposition, the regressors and the corresponding IV are denoted here:

$$w_t = \left(\triangle^M y_{t-1}, \triangle^M x_t^\top\right)^\top, \quad z_t = \left(\triangle^M x_{t-1}^\top, \triangle^M x_{t-2}^\top, \ldots, \triangle^M x_{t-q}^\top, \triangle^M x_t^\top\right)^\top. \tag{11}$$

Given that x_t is weakly exogenous, the following orthogonally condition holds:

$$\mathrm{E}(z_t \triangle^M \varepsilon_t) = \mathrm{E}\left[z_t(\triangle^M y_t - \rho_1 \triangle^M y_{t-1} - \triangle^M x_t^\top \beta)\right] = 0. \tag{12}$$

It is important to document here that, for identification purpose, only one of the qk lagged differenced exogenous regressors is needed to deliver a consistent IV estimator when the number of LDV in (5) is only 1. Nevertheless, more IV basically improves the efficiency of the IV estimator. For the general ADL model in (1), the legitimate IV for the MD-TSLS depends on the values of p, k, and l. This important issue will be discussed in Section 5 more clearly.

We now present the formulae for IV estimation and illustrate the data arrangement. Denote the stack of differenced dependent variable and that of differenced stochastic exogenous variables as:

$$\triangle^M Y(q) = \begin{bmatrix} \triangle^M y_{M+q+1} \\ \triangle^M y_{M+q+2} \\ \vdots \\ \triangle^M y_T \end{bmatrix}, \quad \triangle^M X(q) = \begin{bmatrix} \triangle^M x_{M+q+1}^\top \\ \triangle^M x_{M+q+2}^\top \\ \vdots \\ \triangle^M x_T^\top \end{bmatrix}. \tag{13}$$

The corresponding sample averages are denoted as $\overline{\triangle^M Y(q)}$ and $\overline{\triangle^M X(q)}$, respectively. The notations clarify that the observations used for estimation depend on the number of lagged differenced exogenenous variables, q, and the number of differencing, M. Similarly, the observations for w_t and z_t in (11) can be stacked as:

$$W(M,q) = \begin{bmatrix} w_{M+q+1} \\ w_{M+q+2} \\ \vdots \\ w_T \end{bmatrix}, \quad Z(M,q) = \begin{bmatrix} z_{M+q+1} \\ z_{M+q+2} \\ \vdots \\ z_T \end{bmatrix}. \tag{14}$$

The corresponding sample averages from the data in (14) are denoted as $\overline{W(M,q)}$ and $\overline{Z(M,q)}$, respectively.

Like the OLS-based MD estimator proposed in Tsay (2007), $\triangle^M y_t$, $\triangle^M y_{t-1}$, and $\triangle^M x_t$ in (8) are demeaned before running the estimation. One possible demean procedure is:

$$\widetilde{y}_t = \triangle^M y_t - \overline{\triangle^M Y(0)}, \qquad \widetilde{w}_t = w_t - \left(\overline{\triangle^M Y(0)}, \overline{\triangle^M X(0)}^\top\right)^\top, \tag{15}$$

i.e., all the sample means are calculated with the sample observations after multiple differencing, i.e., $t = M+1, \ldots, T$. Under this circumstance, the $(q+1)k$ instrumental variables for the $(k+1)$ demeaned regressors $\widetilde{w}_t$ in (15) are:

$$\widetilde{z}_t = z_t - i_{q+1} \otimes \overline{\triangle^M X(0)}, \tag{16}$$

where i_{q+1} is a $(q+1)\times 1$ vector of ones, and $\otimes$ is the Kronecker product. Another theoretically feasible demean procedure is:

$$\widetilde{y}_t = \triangle^M y_t - \overline{\triangle^M Y(q)}, \quad \widetilde{w}_t = w_t - \overline{W(M,q)}, \quad \widetilde{z}_t = z_t - \overline{Z(M,q)}. \tag{17}$$

In other words, y_t, w_t, and z_t are demeaned with their respective sample mean calculated from the observations truly used for estimation.

The observations $\widetilde{y}_t$, $\widetilde{w}_t$, and $\widetilde{z}_t$ defined in (15) and (16), or (17) are stacked as:

$$\widetilde{Y}(M,q)=\begin{bmatrix}\widetilde{y}_{M+q+1}\\ \widetilde{y}_{M+q+2}\\ \vdots\\ \widetilde{y}_T\end{bmatrix},\quad \widetilde{W}(M,q)=\begin{bmatrix}\widetilde{w}_{M+q+1}^\top\\ \widetilde{w}_{M+q+2}^\top\\ \vdots\\ \widetilde{w}_T^\top\end{bmatrix},\quad \widetilde{Z}(M,q)=\begin{bmatrix}\widetilde{z}_{M+q+1}^\top\\ \widetilde{z}_{M+q+2}^\top\\ \vdots\\ \widetilde{z}_T^\top\end{bmatrix}. \tag{18}$$

With the data defined in (18), the MD-TSLS estimator for γ is computed as:

$$\widehat{\gamma}_{\text{TSLS}}^{\text{MD}}=\left\{\widetilde{W}(M,q)^\top P_{\widetilde{Z}(M,q)}\widetilde{W}(M,q)\right\}^{-1}\widetilde{W}(M,q)^\top P_{\widetilde{Z}(M,q)}\widetilde{Y}(M,q), \tag{19}$$

where

$$P_{\widetilde{Z}(M,q)}=\widetilde{Z}(M,q)\left\{\widetilde{Z}(M,q)^\top\widetilde{Z}(M,q)\right\}^{-1}\widetilde{Z}(M,q)^\top, \tag{20}$$

provided the above inverse matrices are well defined. The associated residuals are:

$$e_t=\widetilde{y}_t-\widetilde{w}_t^\top\widehat{\gamma}_{\text{TSLS}}^{\text{MD}},\quad t=M+q+1,\ldots,T. \tag{21}$$

For ease of comparison, the OLS-based MD estimator is also presented here:

$$\widehat{\gamma}_{\text{OLS}}^{\text{MD}}=\left\{\widetilde{W}(M,q)^\top\widetilde{W}(M,q)\right\}^{-1}\widetilde{W}(M,q)^\top\widetilde{Y}(M,q), \tag{22}$$

where $\widehat{\gamma}_{\text{OLS}}^{\text{MD}}$ is expected to be biased if $\triangle^M\varepsilon_t$ is serially correlated under the model in (8). Note that $\widehat{\gamma}_{\text{OLS}}^{\text{MD}}$ actually is the MD estimator considered in Tsay (2007) where the presence of LDV is not considered. In addition, the inference method developed in Tsay (2007) cannot be applied to the model in (1), because the presence of LDV makes the regularity conditions required for Robinson's estimator unsatisfied.

The third step of the MD-TSLS approach is to employ Robinson's (1998) long-run variance estimator to construct the covariance matrix of $\widehat{\gamma}_{\text{TSLS}}^{\text{MD}}$, i.e.:

$$\widehat{D}_{\text{TSLS}}^{\text{MD}}=(\widehat{M}_{\text{TSLS}}^{\text{MD}})\widehat{V}_{\text{TSLS}}^{\text{MD}}(\widehat{M}_{\text{TSLS}}^{\text{MD}})^\top, \tag{23}$$

where

$$\widehat{M}_{\text{TSLS}}^{\text{MD}}=\widehat{A}(M,q)\widehat{B}(M,q), \tag{24}$$

such that

$$\widehat{A}(M,q) = \left\{ \frac{1}{T-M-q} \widetilde{W}(M,q)^\top P_{\widetilde{Z}(M,q)} \widetilde{W}(M,q) \right\}^{-1}, \tag{25}$$

$$\widehat{B}(M,q) = \left\{ \widetilde{W}(M,q)^\top \widetilde{Z}(M,q) \right\} \left\{ \widetilde{Z}(M,q)^\top \widetilde{Z}(M,q) \right\}^{-1}, \tag{26}$$

with

$$\widehat{V}^{\text{MD}}_{\text{TSLS}} = \sum_{j=-(T-M-q)+1}^{(T-M-q)-1} \left(\widehat{c}_j \times \widehat{d}_j \right), \tag{27}$$

and

$$\begin{cases} \widehat{c}_j = (T-M-q)^{-1} \sum_{M+q+1 \le t, t+j \le T} \widetilde{z}_t \widetilde{z}_{t+j}^\top, \\ \widehat{d}_j = (T-M-q)^{-1} \sum_{M+q+1 \le t, t+j \le T} e_t \, e_{t+j}. \end{cases} \tag{28}$$

Clearly, the computation of $\widehat{D}^{\text{MD}}_{\text{TSLS}}$ does not involve any choice of bandwidth parameter and kernel function. It can be easily implemented with standard statistic packages.

By (9), (12), and the assumption that the spectral density function of the product $z_t \triangle^M \varepsilon_t$ is finite and positive definite, the long-run variance of $z_t \triangle^M \varepsilon_t$ can be consistently estimated with Robinson's (1998) method when x_t and ε_t satisfy the conditions in Assumption A of Robinson (1998). Accordingly, the asymptotic properties of the MD-TSLS estimator in (19) are derived by using the results in Theorem 5 of Robinson (1998) and presented in the following theorem.

Theorem 1. *Under M being defined in (9), x_t and ε_t satisfy the conditions in Assumption A of Robinson (1998), and the spectral density function of $z_t \triangle^M \varepsilon_t$ is finite and positive definite; then as $T \to \infty$, the MD-TSLS estimator defined in (19) for the model in (5) is asymptotically distributed as $(T-M-q)^{1/2}(\widehat{D}^{\text{MD}}_{\text{TSLS}})^{-1/2} \left(\widehat{\gamma}^{\text{MD}}_{\text{TSLS}} - \gamma \right) \longrightarrow N(0, I_{k+1})$, where $\widehat{D}^{\text{MD}}_{\text{TSLS}}$ is defined in (23), $\longrightarrow$ stands for convergence in distribution, and I_{k+1} is an identity matrix.*

Theorem 1 reveals that $\widehat{\gamma}^{\text{MD}}_{\text{TSLS}}$ converges at the rate of $\sqrt{T}$ as usually found in the short memory stochastic linear difference equation model. Moreover, the MD-TSLS estimator is consistent and asymptotically normally distributed, even though the regressors and errors are stationary long memory processes. Thus, the usual t-ratio statistic associated with the MD-TSLS estimators is tested with the critical values from the standard normal distribution. Another important message implicit in Theorem 1 and which has been spelled out in Tsay (2007, 2009) is that: If the range and various combinations of the differencing parameters in x_t and that of ε_t are uncertain, then a larger value of M can be picked up to ensure the differencing parameters of $\triangle^M x_t$ and that of $\triangle^M \varepsilon_t$ are all negative. This feature is particularly useful for empirical applications, because the integration order of $x_{h,t}$ and that of ε_t are simply unknown in reality, and the differencing parameters of $x_{h,t}$ and that of ε_t just cannot be accurately estimated when the sample size is relatively moderate. For example, the observations used in the following empirical studies of Congressional approval are only 80, which is not large enough to promise a satisfactory estimate of the fractional differencing parameter, no matter whether the time-domain or frequency-domain methods are employed.

3 MD-GMM Estimator

This section considers the MD-GMM estimation of the stochastic linear difference equation in (5) by extending the preceding IV estimation framework. The MD-GMM estimator is calculated as:

$$\begin{aligned} \widehat{\gamma}^{\mathrm{MD}}_{\mathrm{GMM}} &= \left\{ \widetilde{W}(M,q)^\top \widetilde{Z}(M,q) (\widehat{V}^{\mathrm{MD}}_{\mathrm{TSLS}})^{-1} \widetilde{Z}(M,q)^\top \widetilde{W}(M,q) \right\}^{-1} \times \\ & \quad \widetilde{W}(M,q)^\top \widetilde{Z}(M,q) (\widehat{V}^{\mathrm{MD}}_{\mathrm{TSLS}})^{-1} \widetilde{Z}(M,q)^\top \widetilde{Y}(M,q), \end{aligned} \tag{29}$$

where $\widehat{V}^{\mathrm{MD}}_{\mathrm{TSLS}}$ is obtained from (27). This surely indicates that the MD-GMM estimator in (29) is a two-step GMM estimator.

Like the asymptotic covariance matrix of MD-TSLS estimator, $\widehat{D}^{\mathrm{MD}}_{\mathrm{TSLS}}$, the asymptotic covariance matrix of MD-GMM estimator is simple to compute, i.e.:

$$\widehat{D}^{\mathrm{MD}}_{\mathrm{GMM}} = \left\{ \frac{1}{(T-M-q)^2} \widetilde{W}(M,q)^\top \widetilde{Z}(M,q) \left(\widehat{V}^{\mathrm{MD}}_{\mathrm{TSLS}} \right)^{-1} \widetilde{Z}(M,q)^\top \widetilde{W}(M,q) \right\}^{-1}. \tag{30}$$

The asymptotic properties of the MD-GMM estimator presented in the following Theorem 2 can be derived by extending the results in Theorem 1.

THEOREM 2. *Under M being defined in (9), x_t and ε_t satisfy the conditions in Assumption A of Robinson (1998), and the spectral density function of $z_t \triangle^M \varepsilon_t$ is finite and positive definite; then as $T \to \infty$, the MD-GMM estimator defined in (29) for the model in (5) is asymptotically normally distributed and $(T - M - q)^{1/2}(\widehat{D}^{\mathrm{MD}}_{\mathrm{GMM}})^{-1/2}\left(\widehat{\gamma}^{\mathrm{MD}}_{\mathrm{GMM}} - \gamma\right) \longrightarrow N(0, I_{k+1})$, where $\widehat{D}^{\mathrm{MD}}_{\mathrm{GMM}}$ is defined in (30).*

One particular advantage of the MD-GMM estimator over its MD-TSLS counterpart is that it possesses a higher efficiency level if the number of IV exceeds that of LDV. That partly explains why the GMM estimator of Hansen (1982) is popular in the literature. Another interesting characteristic of the MD-GMM estimator is that we can follow the literature to propose the following overidentification test statistic:

$$\Xi = (T - M - q) g(\widehat{\gamma}^{\mathrm{MD}}_{\mathrm{GMM}})^\top \left(\widehat{V}^{\mathrm{MD}}_{\mathrm{TSLS}}\right)^{-1} g(\widehat{\gamma}^{\mathrm{MD}}_{\mathrm{GMM}}), \tag{31}$$

where

$$g(\widehat{\gamma}^{\mathrm{MD}}_{\mathrm{GMM}}) = (T - M - q)^{-1} \widetilde{Z}(M, q)^\top e_{\mathrm{GMM}}, \quad \text{and} \quad e_{\mathrm{GMM}} = \widetilde{Y}(M, q) - \widetilde{W}(M, q)\widehat{\gamma}^{\mathrm{MD}}_{\mathrm{GMM}}. \tag{32}$$

The overidentification statistic is tested with the critical values from a χ^2 distribution with a degree of freedom equal to the number of overidentifying restrictions. This statistic is suggested by Hansen (1982) to test whether the instrumental variables really satisfy the orthogonality conditions required for the MD-GMM estimation. In a sense, the overidentification statistic can be viewed as a specification test to check the appropriateness of the model used for describing the data under investigation.

The lead M-th differenced exogenous variables also can be jointly employed with the lagged M-th differenced explanatory variables as IV. The results in Theorem 1 and Theorem 2 remain intact under this modification. By contrast, one should be more cautious to use the lagged or lead 'level' exogenous variables, x_t, as IV when the model is being MD-transformed with a positive value of M. The reason is that the fractional differencing parameter of x_t and that of $\triangle^M \varepsilon_t$ are highly likely to have an opposite sign which might make the $\sqrt{T}$ rate of convergence of the resulting MD-TSLS and

MD-GMM estimators fail as shown in Tsay (2000). Furthermore, the choice of IV does not include the non-linear function of x_t or that of $\triangle^M x_t$ for the model in (7), simply because the analysis in Robinson (1998) does not cover this complicated situation.

4 Generalization of Theoretical Analysis

Section 2 and Section 3 illustrate the methodology for dealing with the long memory stochastic linear difference equations. The preceding findings can be easily generalized to the long memory ADL model. The major difference is on the selection of IV. When $l = 0$, the ADL model in (1) becomes:

$$y_t = \alpha_0 + \sum_{i=1}^{p} \rho_i y_{t-i} + \sum_{j=1}^{k} \beta_j x_{j,t} + \varepsilon_t. \tag{33}$$

The associate MD-transformed form is:

$$\triangle^M y_t = \sum_{i=1}^{p} \rho_i \triangle^M y_{t-i} + \triangle^M x_t^\top \beta + \triangle^M \varepsilon_t. \tag{34}$$

Under this set-up, the lagged values of $\triangle^M x_{t-j}$ $(j > 0)$ are the potential candidates as IV . The major requirement is that the number of IV must be greater than or equal to p to fulfill the identification condition.

When $l \neq 0$, after MD-transformation the ADL model becomes:

$$\triangle^M y_t = \sum_{i=1}^{p} \rho_i \triangle^M y_{t-i} + \sum_{j=1}^{k} \sum_{i=0}^{l} \beta_{ji} \triangle^M x_{j,t-i} + \triangle^M \varepsilon_t. \tag{35}$$

Under this circumstance, $\triangle^M x_{t-j}$ $(j > l)$ can be chosen as the IV to implement the MD-TSLS and MD-GMM estimators. Again, the number of IV chosen must be greater than or equal to p so as to satisfy the identification condition.

The theoretical findings can be further generalized to the cases where x_t and ε_t are nonstationary. As long as the regressors are weakly exogenous, this nonstationary long memory ADL model can be estimated with a larger value of M. However, the criteria of choosing M changes to be the one in (5) of Tsay (2007) to ensure the resulting IV estimators are asymptotically normally distributed. This implies that the problems concerning the nonstationary ADL model of De Boef and Keele (2008) can be easily resolved with the proposed IV methods.

The above considerations are based on the assumption that all the elements in x_t are weakly exogenous. If some of the regressors are suspected to be endogenous, then the IV estimators can be modified to deal with this more general framework. The remedy is to use the remaining lagged exogenous variables as the IV for the endogenous variables. Specifically, if the regressors are divided into two groups, $x_t^\top = (x_{a,t}^\top, x_{b,t}^\top)$, where $x_{a,t}$ are $(k_1 \times 1)$ endogenous variables and $x_{b,t}$ are $(k_2 \times 1)$ weakly exogenous variables, then the lagged values of $x_{b,t}$ can be used as the IV for $x_{a,t}$ if $M = 0$. When $M = 1$, the lagged values of $\triangle x_{b,t}$ can be employed as the IV for $\triangle x_{a,t}$.

The generalization of the theoretical findings does not stop here. The ADL model can even be generalized as the following vector ADL model:

$$Y_t = A_0 + \sum_{i=1}^{p} R_i Y_{t-i} + \sum_{j=1}^{k} \sum_{i=0}^{l} B_{ji} X_{j,t-i} + V_t, \tag{36}$$

where Y_t and X_t are $(S \times 1)$ vectors of observations, while R_i and B_{ji} are the $(S \times S)$ matrix of parameters to be estimated. When B_{ji} are zero, the vector ADL model reduces to the well-known vector autoregressive (VAR) model. Given that the value of S is finite, the IV methodology developed in this paper and the modified long-run variance estimator in (13) of Tsay (2009) can be merged to deal with the long memory vector ADL model. Indeed, this paper is the first to apply the vector ADL model to the data of DGW (1997). As compared to the vector ADL model in (36), the model used in De Boef and Keele (2008) is named as the scalar ADL model.

5 Numerical Illustrations

This section assesses the finite sample performance of the MD-OLS, MD-TSLS, and MD-GMM estimators under the long memory ADL model. Without loss of generality, there is only one stochastic regressor considered, i.e., $k = 1$, and the simulations center on the stochastic linear difference equation. Furthermore, both x_t and ε_t are generated as the following ARFIMA$(0, d, 0)$ processes:

$$x_t = \triangle^{-d_X} v_t, \quad \varepsilon_t = \triangle^{-d_\varepsilon} w_t, \quad d_X \in (0, 1/2), \quad d_\varepsilon \in (0, 1/2), \tag{37}$$

where v_t and w_t both are generated as zero-mean normally, independently identically distributed white noise (i.i.d.) with:

$$\mathrm{E}(v_t^2) = \sigma_v^2, \quad \mathrm{E}(w_t^2) = \sigma_w^2, \quad \mathrm{E}(v_t w_s) = 0, \quad \text{for all } t, s. \tag{38}$$

Terms σ_v^2 and σ_w^2 in (38) are specified to ensure the resulting variance of x_t and that of ε_t are both equal to 1. In addition, all the programs are written in GAUSS. In the website, http://idv.sinica.edu.tw/wjtsay/htm/jen02a.htm, there exists a GAUSS code for estimating the scalar ADL model. The associated estimation results are presented in the following Table 8.

In the context of a stochastic regressor framework, 2000 replications of y_t are generated with the following model:

$$y_t^n = \rho_1 y_{t-1}^n + \beta_1 x_{1,t}^n + \varepsilon_t^n, \quad t = 1, 2, \ldots, T, \;\; n = 1, 2, \ldots, 2000, \tag{39}$$

where n denotes the n-th replication of the data, and β_1 is set to be 1 throughout this section. Two hundred additional values are generated in order to obtain random starting values. Moreover, $M = 1$ is chosen for the entire simulation studies, even though $M = 0$ can be used for the cases where the sum of d_X and d_ε is less than 1/2. The flexibility of choosing M is clearly demonstrated with this experimental design. Moreover, the demean procedures in (15) and (16) are adopted to carry out the Monte Carlo experiment.

Like the MD-TSLS and MD-GMM estimators, the MD-OLS estimator is calculated with the the same number of MD-transformed observations and demean procedure. The number of observations used for these three estimators under various choices of q is the same so as to create a fair comparison scheme. Note that the case $q = 1$ is not considered in the experimental design, because the MD-TSLS estimator is equivalent to the MD-GMM counterpart under the model in (39) with a number of just-identified IV.

Table 1 contains the bias from estimating ρ_1 when $\rho_1 = 0.5$ and $q = 2$ are selected for the model in (39). The robustness of the results in Table 1 will be checked and discussed later. Table 1 reveals that the MD-OLS estimator is seriously biased in estimating ρ_1 and is expected, because $\triangle\varepsilon_t$ is serially correlated under the experimental design. On the other hand, the performance of both MD-TSLS and MD-GMM estimators in estimating ρ_1 is satisfactory although T is only 100.

Both MD-TSLS and MD-GMM estimators are found to perform well in estimating β_1. The bias and RMSE from estimating β_1 are not explicitly presented to shorten the length of this paper. Instead, the performance about ρ_1 is the focal point of this section, because the coefficient of LDV is crucial for describing the short-run and long-run dynamics of the model under consideration.

Tables 2 illustrates the RMSE of the MD-TSLS estimators in estimating ρ_1 and that of the MD-GMM counterpart when $\rho_1 = 0.5$, $q = 2$, and $T = 100$. For expositional purposes, RMSE_ξ is defined as the RMSE of the estimator ξ in estimating ρ_1 of the model in (39). Thus, the finite sample relative efficiency of the MD-GMM estimator to its MD-TSLS counterpart is measured as:

$$\text{Relative efficiency of MD-GMM to MD-TSLS estimator} \; = \frac{\text{RMSE}^{\text{MD}}_{\text{TSLS}}}{\text{RMSE}^{\text{MD}}_{\text{GMM}}}. \tag{40}$$

It follows that the MD-GMM estimator is more efficient than the MD-TSLS counterpart in estimating ρ_1 if the ratio in (40) is greater than 1. As shown in Hansen (1982), the MD-GMM estimator is more efficient than the MD-TSLS counterpart, and the estimation of the covariance matrix in (27) does not affect the asymptotic variance of the MD-GMM estimator under this circumstance. Accordingly, the relative ratio in (40) is expected to be greater than 1 as $T \to \infty$. This phenomenon is clearly borne out in Table 3 for all configurations therein when $T \geq 200$. This implies that the MD-GMM estimator should be recommended for empirical studies, and the following discussions focus on the performance of the MD-GMM estimator.

The sampling properties of the t-ratio statistic generated from the MD-GMM estimator are important, because inference plays an inevitable role in empirical applications. Figure 1 and Figure 2 display the empirical size of the MD-GMM t ratio statistic in a two-tailed test at the 5% level of significance and show that the size control of the t statistic for testing the value of ρ_1 and that of β_1 in (39) is very promising, respectively, even though the sample size is 100. This also illustrates the power of Robinson's long-run variance estimator in controlling the effects of nuisance parameters on the inference performance of the MD-GMM estimator.

The above findings are generated by experimental design, $q = 2$. Whether more IV really improves the efficiency of the IV estimators is another important focal point of this section. Table 4 presents the simulation results when the number of q in Table 2 increases to be 5. It is clear that more IV really improves the performance of both the MD-TSLS and MD-GMM estimators in that the RMSE found in Table 4 is lower than its corresponding value in Table 2. Moreover, the MD-GMM estimator benefits more from the increase of IV than does the MD-TSLS counterpart, because the relative efficiency ratio in Table 5 is bigger than the corresponding ratio in Table 3.

Can the increase of over-identified IV be cost-free? In terms of the RMSE found in Table 6, the answer is 'no'. The simulations indicate that the RMSE of the MD-TSLS estimator in estimating ρ_1 deteriorates when $q = 10$ is used as compared to that observed in Table 4 with $q = 5$. Although the RMSE of the corresponding MD-GMM estimator with $q = 10$ is smaller than that of Table 2 with $q = 2$ in 61 out of 64 configurations considered in the experiment, the performance of the MD-GMM estimator with $q = 10$ is inferior to that found in Table 4 with $q = 5$. The information indicates that the number of IV cannot exceed that of endogenous regressors too much when the sample size is not large enough. Put differently, the value $(q+1)\times k-(k+1) = qk - 1$ should be moderate when T is relatively small.

The first 6 tables and Figures 1 and 2 deal with the cases where $\rho_1 = 0.5$. The robustness of the findings generated from these tables and figures is verified by replicating the simulations in Tables 1-6 and Figures 1-2 by setting ρ_1 to be 0.8. The simulation results are qualitatively identical to those generated from Tables 1-6 and Figures 1-2. This confirms that the simulation results from Tables 1-6 and Figures 1-2 are not sensitive to the value of the short-run dynamic parameter, ρ_1. These results can be provided upon request.

6 Re-explaining Congressional Approval

The methodology developed in this paper is motivated by combining the long memory and the ADL literature into a unified framework. The estimators proposed in this paper can be applied to many studies concerning the short-run and long-run influences of regressors on the dependent variable, even though such data can be stationary or nonstationary long memory processes. Due to the increasing importance of the $I(d)$ process in political science, the IV estimators are applied to re-examine the data of DGW (1997) about Congressional approval for the period 1974:1 to 1993:4 where they employ the stochastic linear difference equation with $p = 1$. As discussed previously, this dataset is chosen, because the $I(d)$ processes have been found useful for modeling Congressional approval and economic expectations in Box-Steffensmeier and Tomlinson (2000). However, their testing results are not satisfactory as a large standard error of the fractional differencing parameter of the residuals from the fractional cointegration analysis between these two variables is found. Box-Steffensmeier and Tomlinson (2000) explain this large standard error as not unexpected since there are only 80 observations. Nevertheless, as will be shown later, the IV estimators still provide a reasonable estimate of the coefficient of LDV when the data-generating processes (DGP) are $I(d)$ processes.

In the political science literature, Lebo (2008) also applies the multivariate ARFIMA model and finds a strong and positive relationship between Congressional approval and lagged presidential approval with the monthly data spanning from 1995 to 2005. Nevertheless, the analysis of Lebo (2008) requires estimating the differencing parameter of each data under consideration. Moreover, the original data used in Lebo (2008) need to be differenced by their respective estimated values of differencing parameter. Thus, the inference method of Lebo (2008) is a kind of two-step procedure and might be subject to the bias induced from estimation. On the other hand, our IV-based method can be viewed as a one-step procedure, because it does not involve the estimation of the differencing parameter and can test the short-run and long-run influence of regressors on dependent variable in one step.

Although the coefficient of the variable, Presidential Approval$_t$, is crucial to testing the divided party government hypothesis, this variable is not included in the final model of Table 2 of DGW (1997), because it is insignificant based on the usual OLS estimation procedure. The OLS estimates in Table 2 of DGW (1997) are replicated in Table 7 for ease of comparison.

The focus of this section is to show that recognizing the variables, Presidential Approval$_t$ and Economic Expectations$_t$, at the RHS of the regression model of DGW (1997) are endogenous is the first step to test the divided party government hypothesis. By the spirit of the divided party government hypothesis, competing parties certainly employ various strategies to countermeasure the action of the opposite party if this hypothesis is truly in the mind of political actors. Presidential approval and Congressional approval by nature are the functions of political actions, e.g., vetoes or overriding vetoes, undertaken by different parties and other measures beyond the control of the president and Congress. Accordingly, the two approvals should not have a clear causal relationship as specified in DGW (1997) and De Boef and Keele (2008).

DGW (1997, p. 186) also explain that "because citizens hold the president accountable for the state of the economy, economic evaluations will affect the president's standing among the public. We believe the same holds for Congress." The statement induces them to employ economic expectations as one of the exogenous variables to explain Congressional approval. Nevertheless, this argument does not rule out the possibility that citizens might hold higher economic expectations if they show a better presidential or Congressional approval, i.e., the variable concerning economic expectations could be endogenous in explaining both presidential and Congressional approval. These political and economic variables are complexly intertwined with each other. The standing point of this paper resembles closely to the finding in De Boef and Kellstedt (2004) in which consumer confidence should not solely be treated as a RHS variable in analyzing the political economy. Politics also affects economics. As a consequence, both the OLS estimate from the stochastic linear difference equation of DGW (1997) and that from the scalar ADL model of De Boef and Keele (2008) are biased no matter whether the disturbance term is serially correlated or not, if the economic expectations and presidential approval are endogenous.

The proposed IV methods by contrast deal with the presence of endogenous variables in the ADL model easily. Specifically, when the RHS variables of Presidential Approval_t and Economic Expectations_t in DGW (1997) are modeled as endogenous, the lagged values of *NY Times* Coverage_t, Presidential Vetoes_t, Veto Override_t, Intra-Congress Conflict_t, and Major Bills_t are potential candidates as IV. Fortunately, the overidentification statistic defined in (31) can be used to check whether these five variables are the legitimate IV for the aforementioned two endogenous variables.

Following DGW (1997), the impacts of the two dummy variables, Koreagate and House Bank shown in Table 2 of DGW (1997), on Congressional approval are netted out for all the following IV estimations. The detailed procedure is to individually regress the dependent variable, LDV, two endogenous variables, and five exogenous variables on a constant and the two dummies. The residuals from these regressions are collected to implement the IV methods. Note that, if the target is the scalar ADL model, then in a similar vein the procedure is to regress the dependent variable, LDV, endogenous variables, and exogenous variables and the one-period lagged values of the endogenous and those of exogenous variables, respectively, on a constant, the aforementioned two dummies, and the one-period lagged values of these dummies. When $M \geq 1$ is adopted, all the preceding procedures remain intact except the data are differenced before running these regressions.

Table 7 presents the MD-GMM estimate when $M = 0$ and the IV are *NY Times* Coverage$_{t-1}$, Presidential Vetoes$_{t-1}$, Veto Override$_{t-1}$, Intra-Congress Conflict$_{t-1}$, and Major Bills$_{t-1}$. The associated overidentification statistic is tested with the critical values from a χ^2 distribution in a one-tailed test with 2 degrees of freedom. The statistic, 0.87, shown in Table 7 indicates that the population moment condition in (12) cannot be rejected, indicating that these five variables are the legitimate IV for implementing the MD-GMM estimator. This finding provides theoretical justification for using these exogenous variables as the IV for the following analysis concerning the stochastic linear difference equation and the scalar ADL model.

Table 7 also shows that the MD-GMM estimate of ρ_1 is 0.91 and the LDV is highly significant. The remaining variables shown in Table 7 are qualitatively identical to those generated from Table 2 of DGW (1997), except that the variable Veto Override$_t$ is insignificant. Nevertheless, the estimate, -0.87, is still very close to the one, -0.99, generated from DGW (1997). Another interesting finding in Table 7 is that, similar to the finding in DGW (1997), the important variable, Presidential Approval$_t$, is insignificant based on the MD-GMM estimator with $M = 0$. There are two explanations for this phenomenon. First, the sample observation is not large enough as argued in Box-Steffensmeier and Tomlinson (2000). Second, beyond correcting the problem of endogenous bias, we need to pay more attention to the time series properties of the endogenous variables. The second explanation is shown to be important to the final conclusion of the empirical analysis.

The robustness of the findings from the MD-GMM estimator with $M = 0$ in Table 7 is checked when M increases to be 1. Under this circumstance, the LDV is $\triangle$Congressional Approval$_{t-1}$, and the endogenous variables are $\triangle$Presidential Approval$_t$, and $\triangle$Economic Expectations$_t$. We employ the first-differenced of the IV used in Table 7 as the instruments when $M = 1$, i.e., $\triangle$*NY Times* Coverage$_{t-1}$, $\triangle$Presidential Vetoes$_{t-1}$, $\triangle$Veto Override$_{t-1}$, $\triangle$Intra-Congress Conflict$_{t-1}$, and $\triangle$Major Bills$_{t-1}$. The resulting estimate of ρ_1 is 0.73 and the associated value of the overidentification statistic is 1.42 which is not significant based on the critical values from a χ^2 distribution in a one-tailed test with 2 degrees of freedom, either. The testing result from the overidentification test further demonstrates that these five exogenous variables are the legitimate IV for the endogenous variables in explaining Congressional approval under the stochastic linear difference equation framework. Accordingly, these five variables are employed as the IV for the following discussions about the scalar ADL model.

The preceding analysis hinges on the stochastic linear difference equation framework. This section also considers issues about the scalar ADL model of Boef and Keele (2008) with $M = 0$. Under this set-up, the RHS endogenous variables are Presidential Approval$_t$, Presidential Approval$_{t-1}$, Economic Expectations$_t$, and Economic Expectations$_{t-1}$. Therefore, *NY Times* Coverage$_{t-2}$, Presidential Vetoes$_{t-2}$, Veto Override$_{t-2}$, Intra-Congress Conflict$_{t-2}$, and Major Bills$_{t-2}$ can be chosen as the corresponding IV. The sum of the number of LDV and that of RHS endogenous variables are the same as the number of IV, MD-TSLS, and MD-GMM estimators that generate identical results. Table 8 shows that the new estimate for the coefficient of LDV is 0.78 which is again very close to the one, 0.80, found in Table 2 of DGW (1997).

Like the analysis for the stochastic linear difference equation, the robustness of the findings about the scalar ADL model in Table 8 is checked by changing the value of M to be 1. Under this situation, the LDV becomes $\triangle$Congressional Approval$_{t-1}$, the endogenous variables are $\triangle$Presidential Approval$_t$, $\triangle$Presidential Approval$_{t-1}$, $\triangle$Economic Expectations$_t$, and $\triangle$Economic Expectations$_{t-1}$, and the IV employed are $\triangle$*NY Times* Coverage$_{t-2}$, $\triangle$Presidential Vetoes$_{t-2}$, $\triangle$Veto Override$_{t-2}$, $\triangle$Intra-Congress Conflict$_{t-2}$, and $\triangle$Major Bills$_{t-2}$. The resulting MD-GMM estimate of ρ_1 is 0.89.

In summary, the estimates of ρ_1 are 0.91 and 0.73 from the stochastic linear difference model when $M = 0$ and $M = 1$ are used, respectively. The corresponding estimates become 0.78 and 0.89 under the scalar ADL framework. The common feature among these four IV estimates is that they are all close to 0.825, which is the mean of these four estimates, no matter whether $M = 0$ or $M = 1$ is used. This indicates that the IV estimators can deal with the endogenous bias effectively and also implies that the population value of ρ_1 should be around 0.825. Interestingly, the mean estimate 0.825 is very close to the estimate 0.80 found in DGW (1997) using the level data. This additional information reveals that the choice of $M = 0$ adopted by DGW (1997) in carrying out their OLS estimation is reasonable for describing the time series properties of Congressional approval, except that they do not consider the endogenous bias in estimating the coefficient of economic expectations and that of presidential approval. Accordingly, the vector ADL model in (36) based on the level data of DGW (1997) is a good candidate to capture the time-varying feature of Congressional approval, because it can encompass the idea of DGW (1997) and the standing point of this paper that the three endogenous variables, economic expectations, and Congressional approval and presidential approval should be modeled simultaneously.

Table 8 displays the OLS estimates of the vector ADL model. The OLS estimates are found to be qualitatively similar to those presented in Table 2 of DGW (1997). For example, the estimate of ρ_1 is 0.81 and the variable Presidential Approval$_{t-1}$ is insignificant. Why do both presidential approval and economic expectations fail to explain Congressional approval given that DGW (1997, p. 186) argue that "economic expectations drive both presidential and Congressional approval, and presidential approval affects Congressional approval"? One explanation is that there might exist a strong co-movement between presidential approval and economic expectations as the previous arguments predict.

Figure 3 illustrates the time path of presidential approval and that of economic expectations. The data in Figure 3 have been netted out the effects of a constant, two dummies, and the one-period lagged values of these two dummies as mentioned previously, and the resulting values have been further demeaned and standardized to create a clear comparison among these two variables. As expected, the time paths of these two variables are very close to each other, supporting the argument that citizens hold the president accountable for the state of the economy. This also reveals that there might exist some sort of near multicollinearity if these two variables are both included in the regression. In fact, Figure 3 of De Boef and Kellstedt (2004) demonstrates an obvious similarity between presidential approval and economic expectations based on a different dataset compiled by them.

We replicate the regression based on the vector ADL model in Table 8, but delete one of the endogenous variables at a time. The results shown in Table 9 illustrate a new picture about the impacts of presidential approval on Congressional approval. The divided party government hypotheis is strongly supported if Economic Expectations$_{t-1}$ is deleted from the regression since the corresponding long-run effects are $\frac{0.09}{1-0.85} = 0.6$ and much higher than the value of 0.08 found in De Boef and Keele (2008). On the other hand, the results of the vector ADL model in Table 9 are almost qualitatively identical to those presented in Table 2 of DGW (1997) if Presidential Approval$_{t-1}$ is discarded, while keeping Economic Expectations$_{t-1}$ in the regression.

Since the variable of economic expectations is omitted in the regression, the 0.6 estimate of the long-run effects of presidential approval on Congressional approval found in Table 9 is biased. Fortunately, the two regressions in Table 9 provide an opportunity to compute the real impacts of presidential approval on Congressional approval. Based on the explanations in Greene (2000) about the impacts of omitted variables, X_2, on the remaining variables, X_1, there is a formula to link the expected value of the OLS estimator for the coefficient of X_1 to the corresponding population parameter β_1 and the slopes in the least squares regression of the corresponding column of X_2 on X_1. Plugging the OLS estimates in Table 9 for the expected value of the OLS estimator in (8.16) of Greene (2000), two set of equations are found to solve two unobserved parameters pertaining to economic expectations and presidential approval. With some calculations, a new estimate of 0.0532 for economic expectations and an estimate of 0.0570 for presidential approval are obtained. As expected, these estimates are very similar to those contained in Table 8 about the vector ADL model, supporting that the OLS estimator is consistent under the near multicollinearity framework as documented in the econometric literature. Combining the estimate, 0.057, with the two coefficient estimates of LDV in Table 9, two estimates of the long-run effects of presidential approval on Congressional approval are observed , i.e., $\frac{0.057}{1-0.85} = 0.38$ and $\frac{0.057}{1-0.78} = 0.26$. If the 0.81 estimate of the coefficient of LDV from the vector ADL in Table 8 is employed, then the corresponding long-run effects are $\frac{0.057}{1-0.81} = 0.30$. All three estimates are several times larger than the value, 0.08, found in De Boef and Keele (2008).

We point out here that we do not construct a theoretically justified confidence interval to test the significance of the preceding three estimates of the long-run effect of presidential approval on Congressional approval under the near multicollinearity scenario. That is out of this paper's scope and might be left for future research. Nevertheless, one of the objectives of this section is to explore the reason why the variable Presidential Approval$_t$ being not included in Table 2 of DGW (1997) might be due to the near multicollinearity problem induced by the joint presence of presidential approval and economic expectations. This also explains why the decomposition method along with the vector ADL model proposed in this paper provides a much stronger support of the divided party government hypothesis than does the scalar ADL model adopted in De Boef and Keele (2008), because the aforementioned near multicollinearity problem is well taken into account in this paper.

7 Conclusions

A class of IV-based estimators built on the idea of Liviatan (1963) and the MD estimator of Tsay (2007) is shown to be useful to deal with the estimation and inference issues associated with a general class of the long memory ADL models. These IV estimators are consistent and asymptotically normally distributed under suitable regularity conditions. The Monte Carlo simulation reveals that the bias control of the IV estimators are excellent under various long memory regressors and errors considered in this paper even though the sample size is 100. In addition, the size control of the t test generated from the MD-GMM estimator is satisfactory. However, the finite sample performance of the MD-GMM estimator in estimating the coefficient of LDV does not improve monotonically with the number of over-identified IV, indicating that the number of over-identified IV is crucial to the performance of the MD-GMM estimation and should be cautiously selected. Although choosing an optimal number of IV for the long memory ADL model is out of this paper's scope, the lessons from the Monte Carlo experiment do reveal that, when the sample size is relatively small, the number of IV should not exceed that of endogenous regressors too much.

The proposed IV estimators are applied to the data of DGW (1997) to test the divided party government hypothesis. The cutting-edge of the empirical study is to point out that economic expectations and presidential approval should be treated as endogenous variables in explaining Congressional approval. This point of view and the findings from the IV estimators inspire us to propose a vector ADL model to re-explain Congressional approval. The results from the vector ADL model support the argument of Patterson and Caldeira (1990) that presidential approval does play an important role in affecting Congressional approval after controlling the near multicollinearity induced by the joint presence of presidential approval and economic expectations in the vector ADL model. Most importantly, the support of the divided party government hypothesis from the vector ADL model is at least three times stronger than that generated from the scalar ADL model of De Boef and Keele (2008).

Acknowledgements

I would like to thank two anonymous referees and the editor for their valuable comments and suggestions. I also thank Peng-Hsuan Ke for her excellent research assistance.

Mathematical Appendix

Before discussing the method of choosing M in (9) for implementing the MD estimator, let us remark here that M must be big enough to ensure that the resulting transformed regressors and transformed error terms are covariance stationary in order to represent them with the following infinite-order MA processes:

$$\triangle^M \varepsilon_t = \sum_{i=0}^{\infty} \psi_i a_{t-i}, \qquad \triangle^M x_t = \sum_{i=0}^{\infty} \varphi_i b_{t-i}, \qquad t = M+1, \ldots, T,$$

where a_t and b_t are assumed to satisfy the conditions in Assumption A of Robinson (1998).

Given the above-mentioned restrictions imposed on $\triangle^M x_t$ and $\triangle^M \varepsilon_t$, Robinson (1998) shows that the long-run variance of $\triangle^M x_t \, \triangle^M \varepsilon_t$ can be consistently estimated if the fractional differencing parameter of $\triangle^M x_t$ and that of $\triangle^M \varepsilon_t$ are all greater than or equal to zero and $0 \leq (d_{x,h} - M) + (d_\varepsilon - M) < 1/2$ for all $h = 1, \ldots, k$. Theoretically, the finding concerning Robinson's (1998) long-run variance estimator is of great importance to the inference of ρ_1 and β in (8). One drawback, empirically, is that the possibility that the conditions $0 \leq (d_{x,h} - M) + (d_\varepsilon - M) < 1/2$ are satisfied for all $h = 1, \ldots, k$ could be very low, especially when the values of k are relatively large. If we further consider the difficulties that the integration orders of $\triangle^M x_t$ and that of $\triangle^M \varepsilon_t$ are simply unknown in reality, and the differencing parameters of the transformed regressors $\triangle^M x_t$ and that of transformed error term $\triangle^M \varepsilon_t$ just cannot be estimated without bias when the sample size is relatively moderate, then we quickly realize that it may be risky to apply Robinson's (1998) estimator to the long memory ADL data without modifications, not to mention that the total differencing parameters to be estimated are $k+1$.

The major advantages of using the choice of M in (9) is that: Even though we are not sure about the range and the combinations of the $k+1$ differencing parameters inherent in (8), we still can apply the methodology developed in Robinson (1998) to test the value of ρ_1 and β in the long memory ADL model, provided that we are willing to pick a larger value of M to ensure that the differencing parameters of the transformed regressors and those of the transformed error terms are all negative. Moreover, under the choice of M in (9) and suitable regularity conditions, Robinson's (1998) long-run variance estimator is useful to control the impacts of nuisance parameters on the sampling properties of the propose IV methods in this paper. As a result, both estimators are easy to implement in that the MD-TSLS and MD-GMM estimators have a closed-form expression, and the computation of the long-run variance estimator of Robinson (1998) involves no choice of a kernel function and bandwidth parameter. The cost associated with Robinson's approach is that the stochastic regressors and errors must satisfy Assumption A of Robinson (1998). The most stringent condition imposed in this Assumption A is that the stochastic regressors and disturbance term must be uncorrelated at all leads and lags. This condition in fact is not restrictive at all and is exactly the weakly exogenous regressors condition imposed in the existing ADL literature.

REFERENCES

1. Amemiya, Takeshi, and Wayne A. Fuller. 1967. "A Comparative Study of Alternative Estimators in a Distributed Lag Model." *Econometrica* 35 (July-October): 509–529.

2. Box, George E.P., and Gwilym M. Jenkins. 1970. *Time Series Analysis Forecasting and Control*. San Francisco: Holden-Day.

3. Box-Steffensmeier, Janet M., and Renée M. Smith. 1996. "The Dynamics of Aggregate Partisanship." *American Political Science Review* 90 (September): 567–580.

4. Box-Steffensmeier, Janet M., and Renée M. Smith. 1998. "Investigating Political Dynamics Using Fractional Integration Methods." *American Journal of Political Science* 42 (April): 661–689.

5. Box-Steffensmeier, Janet M., and Andrew R. Tomlinson. 2000. "Fractional Integration Methods in Political Science." *Electoral Studies* 19: 62–76.

6. Byers, David, Davidson, James, and David Peel. 2000. "The Dynamics of Aggregate Political Popularity: Evidence from Eight Countries." *Electoral Studies* 19: 49–62.

7. Clarke, Harold D., and Matthew Lebo. 2003. "Fractional (Co)integration and Government Party Support in Britain." *British Journal of Political Science* 33 (April): 283–301.

8. Davidson, James E.H., Peel, David A., and J. David Byers. 2006. "Support for Government and Leaders: Fractional Cointegration Analysis of Poll Evidence from the UK, 1960-2004." *Studies in Nonlinear Dynamics & Econometrics* 10, No. 1, Article 3.

9. De Boef, Suzanna, and Luke Keele. 2008. "Taking Time Seriously." *American Journal of Political Science* 52 (January): 184–200.

10. De Boef, Suzanna, and Paul M. Kellstedt. 2004. "The Political (and Economic) Origins of Consumer Confidence." *American Journal of Political Science* 48 (October): 633–649.

11. Dolado, Juan J., Gonzalo, Jesús, and Laura Mayoral. 2003. "Long-Range Dependence in Spanish Political Opinion Poll Series." *Journal of Applied Econometrics* 18 (January): 137–155.

12. Durbin, J. 1959. "Efficient Estimation of Parameters in Moving Average Models." *Biometrika* 46 (December): 306–316.

13. Durr, Robert H., Gilmour, John B., and Christina Wolbrecht. 1997. "Explaining Congressional Approval." *American Journal of Political Science* 41 (January): 175–207.

14. Granger, C.W.J. 1980. "Long Memory Relationships and the Aggregation of Dynamic Models." *Journal of Econometrics* 14: 227–238.

15. Granger, C.W.J., and Roselyne Joyeux. 1980. "An Introduction to Long-Memory Time Series Models and Fractional Differencing." *Journal of Time Series Analysis* 1, 15–29.

16. Greene, William H. 2000. *Econometric Analysis*, 4th Ed. New York: Prentice Hall.

17. Hannan, E.J., and D.F. Nicholls. 1972. "The Estimation of Mixed Regression, Autoregression, Moving Average, and Distributed Lag Models." *Econometrica* 40 (May): 529–547.

18. Hansen, Lars Peter. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50 (July): 1029–1054.

19. Harvey, Andrew C. 1990. *The Econometric Analysis of Time Series*, 2nd ed. New York: Phillip Allan.

20. Hatanaka, Michio. 1974. "An Efficient Two-Step Estimator for the Dynamic Adjustment Model with Autoregressive Errors." *Journal of Econometrics* 2: 199–220.

21. Hatanaka, Michio. 1976. "Several Efficient Two-Step Estimators for the Dynamic Simultaneous Equations Model with Autoregressive Disturbances." *Journal of Econometrics* 4: 189–204.

22. Hendry, David F. 1995. *Dynamic Econometrics.* Oxford: Oxford University Press.

23. Hosking, J.R.M. 1981. "Fractional Differencing." *Biometrika* 68: 165–176.

24. Keele, Luke, Nathan J. Kelly. 2005. "Dynamic Models for Dynamic Theories: The Ins and Outs of Lagged Dependent Variables." *Political Analysis* 14 (November): 186–205.

25. Koyck, L.M. 1954. *Distributed Lags and Investment Analysis.* Amsterdam: North-Holland.

26. Lai, Brian, and Dan Reiter. 2005. "Rally 'Round the Union Jack? Public Opinion and the Use of Force in the United Kingdom, 1948-2001." *International Studies Quarterly* 49: 255–272.

27. Lebo, Matthew J. 2008. "Divided Government, United Approval: The Dynamic of Congressional and Presidential Approval." *Congress & The Presidency* 35 (Autumn): 1–16

28. Lebo, Matthew J., Walker, Robert W., and Harold D. Clarke. 2000. "You Must Remember This: Dealing with Long Memory in Political Analysis." *Electoral Studies* 19: 31–48.

29. Liviatan, Nissan. 1963. "Consistent Estimation of Distributed Lags." *The International Economic Review* 4 (January): 44–52.

30. Nicholls, D.F., Pagan, A.R., and R.D. Terrell. 1975. "The Estimation and Use of Models with Moving Average Disturbance Terms: A Survey." *International Economic Review* 16 (Febuary): 113–134.

31. Patterson, Samuel C., and Gregory A. Caldeira. 1990. "Standing up for Congress: Variations in Public Esteem Since the 1960s." *Legislative Studies Quarterly* 15 (February): 25–47.

32. Reinsel, Greg. 1979. "Maximum Likelihood Estimation of Stochastic Linear Difference Equations with Autoregressive Moving Average Errors." *Econometrica* 47 (January): 129–151.

33. Robinson, Peter M. 1998. "Inference-without-Smoothing in the Presence of Nonparametric Autocorrelation." *Econometrica* 66 (September): 1163–1182.

34. Tsay, Wen-Jen. 2000. "Estimating Trending Variables in the Presence of Fractionally Integrated Errors." *Econometric Theory* 16 (June): 324–346.

35. Tsay, Wen-Jen. 2007. "Using Difference-Based Methods for Inference in Regression with Fractionally-Integrated Processes." *Journal of Time Series Analysis* 28 (November): 827–843.

36. Tsay, Wen-Jen. 2009. "Estimating Long Memory Time-Series-Cross-Section Data." *Electoral Studies* 28: 129–140.

37. Tsay, Wen-Jen, and Ching-Fan Chung. 2000. "The Spurious Regression of Fractionally Integrated Processes." *Journal of Econometrics* 96: 155–182.

38. Walker, A.M. 1964. "Asymptotic Properties of Least-Squares Estimates of Parameters of the Spectrum of a Stationary Non-deterministic Time Series." *Journal of the Australian Mathematical Society* 4: 363–384.

Table 1. Bias in Estimating ρ_1: $T = 100$, $\rho_1 = 0.5$, and $q = 2$

		d_X							
d_ε		0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
0.05	MD-OLS	−0.440	−0.442	−0.447	−0.453	−0.472	−0.483	−0.511	−0.546
	MD-TSLS	−0.009	−0.017	−0.015	−0.005	−0.020	−0.014	−0.021	−0.035
	MD-GMM	−0.003	−0.013	−0.011	0.000	−0.014	−0.006	−0.012	−0.025
0.10	MD-OLS	−0.416	−0.418	−0.425	−0.437	−0.443	−0.460	−0.485	−0.522
	MD-TSLS	−0.010	−0.008	−0.012	−0.018	−0.018	−0.018	−0.019	−0.022
	MD-GMM	−0.007	−0.004	−0.008	−0.015	−0.012	−0.012	−0.012	−0.011
0.15	MD-OLS	−0.395	−0.396	−0.401	−0.406	−0.421	−0.433	−0.458	−0.498
	MD-TSLS	−0.021	−0.007	−0.009	−0.010	−0.014	−0.017	−0.010	−0.023
	MD-GMM	−0.017	−0.004	−0.007	−0.006	−0.010	−0.011	−0.004	−0.013
0.20	MD-OLS	−0.363	−0.370	−0.371	−0.382	−0.389	−0.405	−0.431	−0.465
	MD-TSLS	−0.010	−0.006	−0.009	−0.012	−0.011	−0.010	−0.016	−0.018
	MD-GMM	−0.007	−0.004	−0.005	−0.008	−0.006	−0.005	−0.010	−0.011
0.25	MD-OLS	−0.334	−0.338	−0.340	−0.349	−0.361	−0.372	−0.398	−0.436
	MD-TSLS	−0.011	−0.009	−0.010	−0.011	−0.014	−0.010	−0.018	−0.024
	MD-GMM	−0.009	−0.005	−0.009	−0.007	−0.011	−0.007	−0.015	−0.018
0.30	MD-OLS	−0.298	−0.300	−0.307	−0.313	−0.324	−0.337	−0.360	−0.396
	MD-TSLS	−0.004	−0.007	−0.011	−0.009	−0.011	−0.010	−0.010	−0.014
	MD-GMM	−0.003	−0.005	−0.009	−0.007	−0.010	−0.008	−0.007	−0.009
0.35	MD-OLS	−0.258	−0.259	−0.262	−0.271	−0.279	−0.296	−0.319	−0.352
	MD-TSLS	−0.007	−0.008	−0.002	−0.007	−0.009	−0.008	−0.012	−0.009
	MD-GMM	−0.006	−0.007	0.000	−0.006	−0.008	−0.006	−0.009	−0.006
0.40	MD-OLS	−0.206	−0.204	−0.210	−0.217	−0.223	−0.239	−0.258	−0.295
	MD-TSLS	−0.008	−0.004	−0.005	−0.004	−0.001	−0.008	−0.006	−0.014
	MD-GMM	−0.007	−0.003	−0.004	−0.003	−0.001	−0.006	−0.004	−0.011

Notes: All the results are based on 2,000 replications of the simulated data defined in (37)-(39) with $M = 1$ chosen for the MD-OLS, MD-TSLS, and MD-GMM estimators, respectively. Bias is computed as the average estimated values minus the true parameter value.

Table 2. RMSE in Estimating ρ_1: $T = 100$, $\rho_1 = 0.5$, and $q = 2$

		d_X							
d_ε		0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
0.05	MD-TSLS	0.153	0.149	0.151	0.154	0.159	0.168	0.177	0.204
	MD-GMM	0.150	0.148	0.152	0.156	0.160	0.169	0.179	0.207
0.10	MD-TSLS	0.146	0.144	0.148	0.146	0.155	0.160	0.179	0.201
	MD-GMM	0.147	0.144	0.149	0.145	0.153	0.160	0.181	0.204
0.15	MD-TSLS	0.142	0.141	0.143	0.147	0.149	0.151	0.164	0.193
	MD-GMM	0.141	0.141	0.143	0.146	0.148	0.151	0.166	0.196
0.20	MD-TSLS	0.140	0.138	0.138	0.138	0.143	0.148	0.167	0.186
	MD-GMM	0.140	0.138	0.138	0.139	0.142	0.148	0.168	0.189
0.25	MD-TSLS	0.132	0.131	0.128	0.138	0.137	0.137	0.152	0.175
	MD-GMM	0.131	0.132	0.127	0.137	0.137	0.137	0.153	0.176
0.30	MD-TSLS	0.117	0.120	0.122	0.124	0.128	0.129	0.141	0.166
	MD-GMM	0.118	0.119	0.121	0.124	0.127	0.129	0.142	0.167
0.35	MD-TSLS	0.108	0.108	0.112	0.113	0.115	0.124	0.129	0.149
	MD-GMM	0.108	0.107	0.112	0.114	0.114	0.124	0.129	0.151
0.40	MD-TSLS	0.093	0.090	0.092	0.093	0.101	0.100	0.111	0.127
	MD-GMM	0.093	0.089	0.092	0.092	0.100	0.100	0.110	0.126

Notes: All the results are based on 2,000 replications of the simulated data defined in (37)-(39) with $M = 1$ chosen for the MD-TSLS and MD-GMM estimators, respectively.

Table 3. Relative Efficiency of the MD-GMM Estimator to the MD-TSLS Counterpart in Estimating ρ_1: $\rho_1 = 0.5$ and $q = 2$

	d_X							
d_ε	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
				$T = 100$				
0.05	1.0162	1.0043	0.9986	0.9856	0.9959	0.9929	0.9864	0.9824
0.10	0.9971	0.9999	0.9952	1.0055	1.0104	0.9996	0.9910	0.9858
0.15	1.0056	0.9969	1.0016	1.0061	1.0066	0.9970	0.9904	0.9860
0.20	0.9993	1.0004	0.9957	0.9996	1.0041	1.0043	0.9936	0.9879
0.25	1.0038	0.9971	1.0011	1.0020	1.0066	0.9965	0.9938	0.9893
0.30	0.9901	1.0092	1.0053	0.9969	1.0096	0.9978	0.9908	0.9942
0.35	0.9988	1.0091	0.9972	0.9976	1.0056	1.0034	0.9981	0.9849
0.40	1.0021	1.0120	1.0092	1.0045	1.0030	1.0031	1.0033	1.0046
				$T = 200$				
0.05	1.0320	1.0287	1.0321	1.0172	1.0313	1.0259	1.0209	1.0203
0.10	1.0184	1.0120	1.0243	1.0201	1.0213	1.0299	1.0172	1.0057
0.15	1.0199	1.0174	1.0126	1.0132	1.0186	1.0246	1.0178	1.0042
0.20	1.0201	1.0245	1.0214	1.0154	1.0156	1.0162	1.0105	1.0125
0.25	1.0117	1.0173	1.0242	1.0140	1.0144	1.0088	1.0095	1.0032
0.30	1.0168	1.0135	1.0078	1.0159	1.0140	1.0119	1.0094	1.0036
0.35	1.0198	1.0119	1.0156	1.0107	1.0203	1.0016	1.0122	1.0090
0.40	1.0182	1.0103	1.0021	1.0015	1.0170	1.0119	1.0167	1.0123
				$T = 300$				
0.05	1.0328	1.0354	1.0426	1.0225	1.0324	1.0312	1.0181	1.0259
0.10	1.0162	1.0307	1.0252	1.0260	1.0374	1.0295	1.0358	1.0242
0.15	1.0233	1.0390	1.0222	1.0272	1.0206	1.0255	1.0231	1.0261
0.20	1.0216	1.0189	1.0176	1.0187	1.0244	1.0177	1.0281	1.0280
0.25	1.0221	1.0228	1.0240	1.0258	1.0212	1.0246	1.0257	1.0194
0.30	1.0179	1.0132	1.0169	1.0242	1.0285	1.0209	1.0244	1.0154
0.35	1.0177	1.0056	1.0243	1.0180	1.0176	1.0180	1.0153	1.0202
0.40	1.0040	1.0205	1.0098	1.0236	1.0207	1.0147	1.0141	1.0175

Notes: All the results are based on 2,000 replications of the simulated data defined in (37)-(39) with $M = 1$ chosen for the MD-TSLS and MD-GMM estimators, respectively.

Table 4. **RMSE in Estimating ρ_1: $T = 100$, $\rho_1 = 0.5$, and $q = 5$**

		d_X							
d_ε		0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
0.05	MD-TSLS	0.148	0.146	0.148	0.145	0.157	0.160	0.171	0.197
	MD-GMM	0.134	0.134	0.135	0.133	0.141	0.143	0.151	0.175
0.10	MD-TSLS	0.141	0.138	0.144	0.145	0.151	0.155	0.170	0.194
	MD-GMM	0.132	0.127	0.130	0.131	0.135	0.140	0.154	0.173
0.15	MD-TSLS	0.140	0.135	0.137	0.139	0.148	0.148	0.156	0.184
	MD-GMM	0.129	0.126	0.126	0.128	0.136	0.135	0.143	0.168
0.20	MD-TSLS	0.134	0.132	0.132	0.132	0.138	0.144	0.158	0.175
	MD-GMM	0.126	0.124	0.123	0.123	0.129	0.130	0.144	0.162
0.25	MD-TSLS	0.127	0.126	0.123	0.131	0.133	0.134	0.149	0.165
	MD-GMM	0.119	0.118	0.115	0.122	0.126	0.125	0.139	0.153
0.30	MD-TSLS	0.111	0.115	0.118	0.119	0.123	0.122	0.134	0.158
	MD-GMM	0.108	0.109	0.113	0.116	0.115	0.115	0.127	0.148
0.35	MD-TSLS	0.104	0.104	0.106	0.109	0.110	0.119	0.123	0.139
	MD-GMM	0.101	0.099	0.101	0.104	0.103	0.113	0.117	0.132
0.40	MD-TSLS	0.090	0.088	0.089	0.089	0.096	0.095	0.104	0.119
	MD-GMM	0.088	0.085	0.086	0.085	0.092	0.090	0.099	0.112

Notes: All the results are based on 2,000 replications of the simulated data defined in (37)-(39) with $M = 1$ chosen for the MD-TSLS and MD-GMM estimators, respectively.

Table 5. Relative Efficiency of the MD-GMM Estimator to the MD-TSLS Counterpart in Estimating ρ_1: $\rho_1 = 0.5$ and $q = 5$

	d_X							
d_ε	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
				$T = 100$				
0.05	1.1006	1.0883	1.0932	1.0964	1.1201	1.1241	1.1352	1.1249
0.10	1.0669	1.0883	1.1082	1.1126	1.1181	1.1028	1.1017	1.1193
0.15	1.0829	1.0725	1.0846	1.0849	1.0837	1.1002	1.0952	1.0949
0.20	1.0634	1.0646	1.0738	1.0808	1.0712	1.1061	1.0940	1.0846
0.25	1.0655	1.0625	1.0696	1.0703	1.0605	1.0700	1.0688	1.0764
0.30	1.0280	1.0519	1.0496	1.0287	1.0711	1.0633	1.0549	1.0684
0.35	1.0278	1.0488	1.0417	1.0480	1.0610	1.0528	1.0465	1.0598
0.40	1.0313	1.0331	1.0366	1.0488	1.0413	1.0565	1.0590	1.0623
				$T = 200$				
0.05	1.1176	1.1417	1.1426	1.1293	1.1657	1.1569	1.1661	1.1695
0.10	1.1068	1.0996	1.1174	1.1243	1.1304	1.1563	1.1552	1.1390
0.15	1.0923	1.0961	1.0944	1.1176	1.1241	1.1362	1.1309	1.1356
0.20	1.0867	1.0841	1.0884	1.0994	1.1112	1.1099	1.1137	1.1251
0.25	1.0569	1.0762	1.0960	1.0893	1.0835	1.1019	1.1113	1.1012
0.30	1.0673	1.0662	1.0716	1.0731	1.0754	1.0782	1.0825	1.0829
0.35	1.0521	1.0570	1.0604	1.0476	1.0881	1.0644	1.0710	1.0796
0.40	1.0447	1.0443	1.0433	1.0373	1.0719	1.0695	1.0859	1.0688
				$T = 300$				
0.05	1.1152	1.1306	1.1647	1.1362	1.1513	1.1754	1.1662	1.1935
0.10	1.1129	1.1110	1.1265	1.1339	1.1495	1.1649	1.1857	1.1677
0.15	1.0922	1.1225	1.1004	1.1290	1.1209	1.1489	1.1350	1.1425
0.20	1.0780	1.0997	1.0885	1.1130	1.1264	1.1110	1.1319	1.1398
0.25	1.0734	1.0854	1.0738	1.1036	1.1060	1.1306	1.1099	1.1174
0.30	1.0619	1.0721	1.0766	1.0682	1.0974	1.0916	1.1188	1.1044
0.35	1.0662	1.0567	1.0836	1.0628	1.0694	1.0721	1.0689	1.1049
0.40	1.0351	1.0565	1.0454	1.0771	1.0737	1.0768	1.0745	1.0979

Notes: All the results are based on 2,000 replications of the simulated data defined in (37)-(39) with $M = 1$ chosen for the MD-TSLS and MD-GMM estimators, respectively.

Table 6. RMSE in Estimating ρ_1: $T = 100$, $\rho_1 = 0.5$, and $q = 10$

		d_X							
d_ε		0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
0.05	MD-TSLS	0.171	0.174	0.176	0.174	0.190	0.197	0.214	0.253
	MD-GMM	0.144	0.147	0.147	0.143	0.153	0.159	0.169	0.203
0.10	MD-TSLS	0.163	0.162	0.170	0.175	0.179	0.189	0.207	0.238
	MD-GMM	0.143	0.137	0.141	0.145	0.147	0.154	0.171	0.193
0.15	MD-TSLS	0.160	0.155	0.158	0.164	0.172	0.176	0.189	0.228
	MD-GMM	0.139	0.136	0.137	0.140	0.146	0.148	0.155	0.190
0.20	MD-TSLS	0.150	0.149	0.151	0.154	0.158	0.169	0.187	0.211
	MD-GMM	0.133	0.132	0.131	0.132	0.136	0.142	0.156	0.178
0.25	MD-TSLS	0.141	0.141	0.137	0.147	0.153	0.155	0.173	0.201
	MD-GMM	0.128	0.126	0.122	0.129	0.135	0.135	0.150	0.170
0.30	MD-TSLS	0.123	0.127	0.131	0.132	0.136	0.140	0.153	0.180
	MD-GMM	0.115	0.115	0.118	0.122	0.122	0.123	0.135	0.158
0.35	MD-TSLS	0.112	0.114	0.112	0.119	0.122	0.130	0.137	0.160
	MD-GMM	0.105	0.106	0.105	0.111	0.111	0.120	0.125	0.143
0.40	MD-TSLS	0.097	0.094	0.095	0.094	0.101	0.105	0.112	0.133
	MD-GMM	0.093	0.090	0.090	0.089	0.095	0.095	0.103	0.120

Notes: All the results are based on 2,000 replications of the simulated data defined in (37)-(39) with $M = 1$ chosen for the MD-TSLS and MD-GMM estimators, respectively.

Table 7. Estimates of the Data of Durr, Gilmour, and Wolbrecht (1997) Based on the Stochastic Linear Difference Equation

	OLS		MD-GMM with $M = 0$	
Variables	Estimate	$\lvert t\text{-ratio}\rvert$	Estimate	$\lvert t\text{-ratio}\rvert$
Congressional Approval$_{t-1}$	0.80	16.00***	0.91	5.77***
Presidential Approval$_t$			0.05	0.26
Economic Expectations$_t$	0.07	2.33**	0.05	0.41
NY Times Coverage$_t$	0.21	3.00***	0.19	2.16**
Presidential Vetoes$_t$	0.24	2.67***	0.21	2.17**
Veto Override$_t$	−0.99	1.80*	−0.87	1.44
Intra-Congress Conflict$_t$	−0.17	1.42	−0.18	1.47
Major Bills$_t$	−0.44	1.57	−0.38	1.41
Overidentification Test			0.87	
Observations	79		79	

Notes: The results are based on the data of DGW (1997) about Congressional approval for the period 1974:1 to 1993:4. Other variables from the analysis are included in the estimating equation, but omitted from the table. The results for the OLS estimation are replicated from Table 2 of DGW (1997). The endogenous variables for the MD-GMM estimation are Presidential Approval$_t$, and Economic Expectations$_t$, while the instrumental variables for the MD-GMM estimation are *NY Times* Coverage$_{t-1}$, Presidential Vetoes$_{t-1}$, Veto Override$_{t-1}$, Intra-Congress Conflict$_{t-1}$, and Major Bills$_{t-1}$. ***, **, and * denote significance at the 1%, 5%, and 10% levels in a two-tailed test, respectively.

Table 8. Estimates of the Data of Durr, Gilmour, and Wolbrecht (1997) Based on the Scalar and Vector ADL Models

	Scalar ADL		Vector ADL	
	MD-GMM with $M = 0$		OLS	
Variables	Estimate	$\lvert t\text{-ratio}\rvert$	Estimate	$\lvert t\text{-ratio}\rvert$
Congressional Approval$_{t-1}$	0.78	1.81*	0.81	12.64***
Presidential Approval$_{t}$	−0.08	0.12		
Presidential Approval$_{t-1}$	0.75	0.63	0.06	1.25
Economic Expectations$_{t}$	−0.83	0.66		
Economic Expectations$_{t-1}$	0.80	0.64	0.05	1.25
NY Times Coverage$_{t}$	0.03	0.09	0.22	2.85***
NY Times Coverage$_{t-1}$	0.32	0.65	0.09	1.16
Presidential Vetoes$_{t}$	0.16	0.70	0.24	2.59***
Presidential Vetoes$_{t-1}$	0.37	0.83	0.09	0.88
Observations	78		79	

Notes: The results are based on the data of DGW (1997) about Congressional approval for the period 1974:1 to 1993:4. Other variables from the analysis are included in the estimating equation, but omitted from the table. The endogenous variables in the MD-GMM estimation are Presidential Approval$_{t}$, Presidential Approval$_{t-1}$, Economic Expectations$_{t}$, and Economic Expectations$_{t-1}$, while the instrumental variables used for the MD-GMM estimation are *NY Times* Coverage$_{t-2}$, Presidential Vetoes$_{t-2}$, Veto Override$_{t-2}$, Intra-Congress Conflict$_{t-2}$, and Major Bills$_{t-2}$. The standard errors for testing the parameters of the vector ADL model are based on the usual OLS estimation procedure. ***, **, and * denote significance at the 1%, 5%, and 10% levels in a two-tailed test, respectively.

Table 9. Estimates of the Data of Durr, Gilmour, and Wolbrecht (1997) Based on the Vector ADL Models with 79 Observations

	OLS		OLS	
Variables	Estimate	$\lvert t\text{-ratio}\rvert$	Estimate	$\lvert t\text{-ratio}\rvert$
Congressional Approval$_{t-1}$	0.85	16.81***	0.78	12.85***
Presidential Approval$_{t-1}$	0.09	2.27**		
Economic Expectations$_{t-1}$			0.08	2.28**
NY Times Coverage$_t$	0.21	2.80***	0.23	3.05***
NY Times Coverage$_{t-1}$	0.08	1.04	0.09	1.20
Presidential Vetoes$_t$	0.22	2.39**	0.25	2.74***
Presidential Vetoes$_{t-1}$	0.08	0.77	0.09	0.84
Veto Override$_t$	−1.12	1.91*	−1.00	1.72*
Veto Override$_{t-1}$	0.23	0.38	0.23	0.37
Intra-Congress Conflict$_t$	−0.17	1.33	−0.15	1.14
Intro-Congress Conflict$_{t-1}$	−0.06	0.43	−0.04	0.31
Major Bills$_t$	−0.30	0.99	−0.49	1.61
Major Bills$_{t-1}$	−0.36	1.19	−0.33	1.10

Notes: The results are based on the level data of DGW (1997) about Congressional approval for the period 1974:1 to 1993:4 with the OLS estimation procedure. ***, **, and * denote significance at the 1%, 5%, and 10% levels in a two-tailed test, respectively.

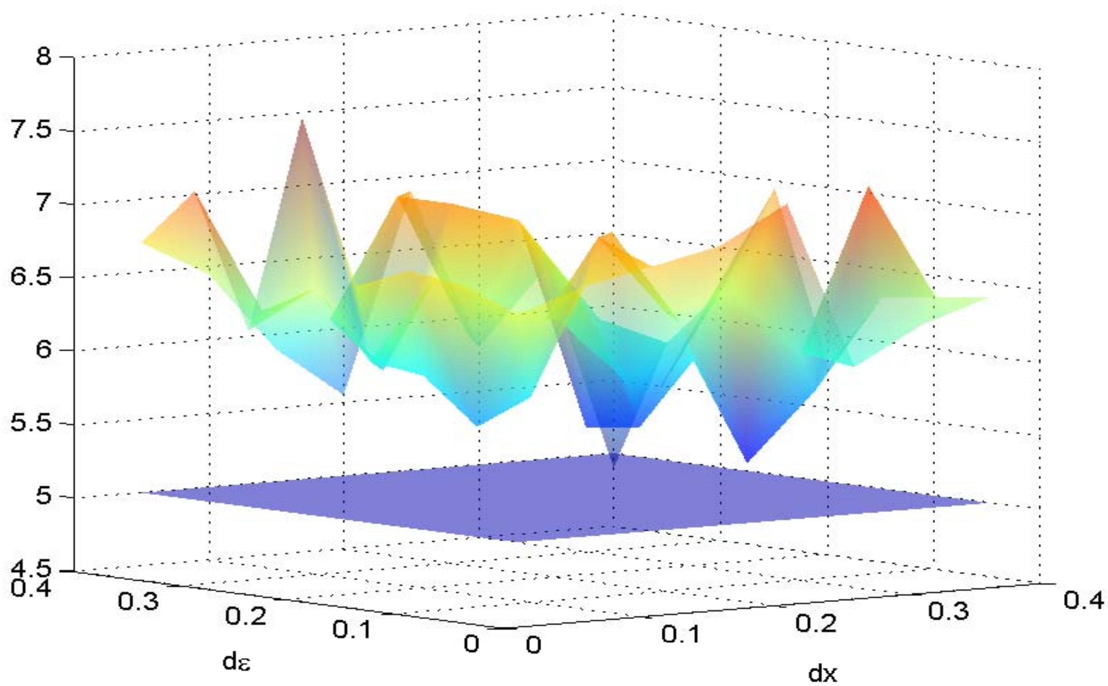

Figure 1: Rejection percentages of t test from the MD-GMM Estimator of ρ_1 at the 5% level of significance: $T = 100$, $\rho_1 = 0.5$, and $q = 2$. All the results are based on 2,000 replications of the simulated data defined in (37)-(39) with $M = 1$ chosen for the MD-GMM estimator.

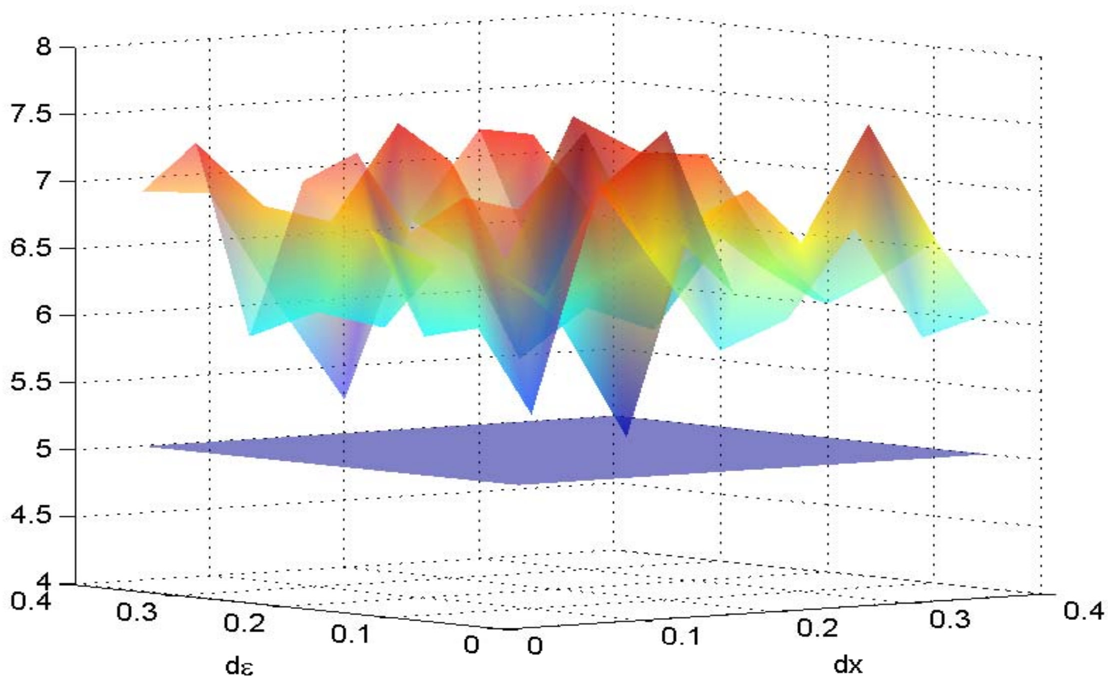

Figure 2: Rejection percentages of t test from the MD-GMM Estimator of β_1 at the 5% level of significance: $T = 100$, $\rho_1 = 0.5$, and $q = 2$. All the results are based on 2,000 replications of the simulated data defined in (37)-(39) with $M = 1$ chosen for the MD-GMM estimator.

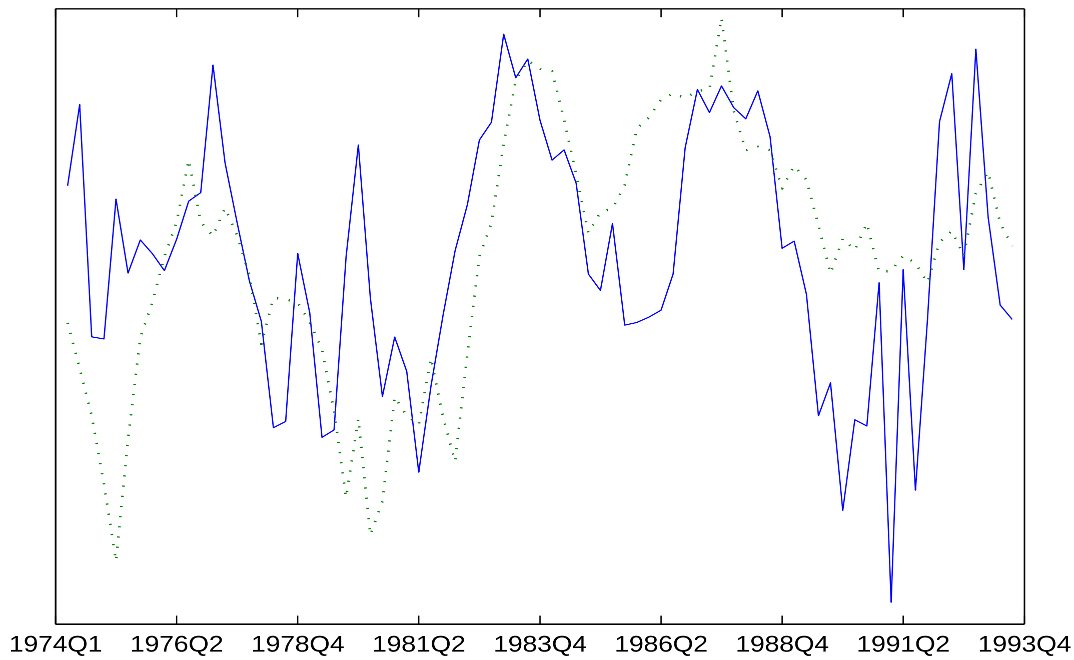

Figure 3: Solid line denotes presidential approval, while dotted line denotes economic expectations. The data have been netted out the effects of a constant, two dummies, and the one-period lagged values of these two dummies on economic expectations and presidential approval as mentioned in the text, and the resulting values have been further demeaned and standardized to make a clear comparison among these two variables.