EFFICIENT IV ESTIMATION IN NONSTATIONARY REGRESSION

An Overview and Simulation Study

YUICHI KITAMURA
University of Minnesota

PETER C.B. PHILLIPS
Cowles Foundation for Research in Economics
Yale University

A limit theory for instrumental variables (IV) estimation that allows for possibly nonstationary processes was developed in Kitamura and Phillips (1992, Fully Modified IV, GIVE, and GMM Estimation with Possibly Non-stationary Regressors and Instruments, mimeo, Yale University). This theory covers a case that is important for practitioners, where the nonstationarity of the regressors may not be of full rank, and shows that the fully modified (FM) regression procedure of Phillips and Hansen (1990) is still applicable. FM versions of the generalized method of moments (GMM) estimator and the generalized instrumental variables estimator (GIVE) were also developed, and these estimators (FM-GMM and FM-GIVE) were designed specifically to take advantage of potential stationarity in the regressors (or unknown linear combinations of them). These estimators were shown to deliver efficiency gains over FM-IV in the estimation of the stationary components of a model.

This paper provides an overview of the FM-IV, FM-GMM, and FM-GIVE procedures and investigates the small sample properties of these estimation procedures by simulations. We compare the following five estimation methods: ordinary least squares, crude (conventional) IV, FM-IV, FM-GMM, and FM-GIVE. Our findings are as follows. (i) In terms of overall performance in both stationary and nonstationary cases, FM-IV is more concentrated and better centered than OLS and crude IV, though it has a higher root mean square error than crude IV due to occasional outliers. (ii) Among FM-IV, FM-GMM, and FM-GIVE, (a) when applied to the stationary coefficients, FM-GIVE generally outperforms FM-IV and FM-GMM by a wide margin, whereas the difference between the latter two is quite small when the AR roots of the stationary processes are rather large; and (b) when applied to the nonstationary coefficients, the three estimators are numerically very close. The performance of the FM-GIVE estimator is generally very encouraging.

1095

1. INTRODUCTION

In our recent paper (Kitamura and Phillips, 1992), a general theory of fully modified instrumental variables (FM-IV) estimation with possibly nonstationary regressors and instruments was developed. The estimation of models with nonstationary endogenous regressors was discussed, allowing explicitly for cointegration in the regressors even though this possibility was excluded in the FM regression procedure that was originally introduced in Phillips and Hansen (1990). Also, efficient instrumental variables (IV) methods such as the generalized instrumental variable estimation (GIVE) method (e.g., Sargan, 1988) and the generalized method of moments (GMM) (e.g., Hansen, 1982) were merged with the FM-IV procedure. The resultant estimators, the FM-GMM and FM-GIVE procedures, were shown to be (i) asymptotically more efficient than the FM-IV procedure with respect to the stationary components and (ii) asymptotically equivalent to FM-IV estimation with respect to the nonstationary components.

Because all of the preceding results are asymptotic, a finite sample analysis is naturally of interest. The versatility of the FM-IV, FM-GMM, and FM-GIVE estimators in practice seems quite important and, thus, the behavior of these estimators in a realistic setup seems worthy of study. The main goal of this paper is to investigate the sampling characteristics of these FM estimators and compare them with conventional methods like ordinary least squares (OLS) and IV through Monte Carlo simulations.

Closely related to this paper is an earlier study by Hansen and Phillips (1990), which provides a systematic treatment of the sampling behavior of FM-IV estimators in which the regressor and the instruments are known to be full rank nonstationary a priori. Simulation studies of conventional instrumental variable estimators with stationary processes can be found in existing literature. A recent study by Nelson and Startz (1990), for example, provides some finite sample results of the conventional IV estimator focusing on the relevance of the instruments and feedback from the regression errors to the regressors. (Phillips [1989] provides some analytic evidence on the finite sample effects of the same factors.) These two factors are also relevant to our FM-IV procedures and will be investigated in this paper. In our case, however, because we are considering the estimation of models with stationary, nonstationary, and cointegrated processes, the specification of the data generating process (DGP) for our simulations can quickly become much more complicated than the aforementioned conventional cases. To avoid unnecessary complexity, a rather simple and idealized DGP is employed in this paper, but it is designed to represent some interesting features of time series that may be encountered in practice.

The plan of the paper is as follows. Section 2 reviews the limit theory developed by Kitamura and Phillips (1992). Section 3 provides details of the DGP and the estimation methods that are used in the simulations and gives

the findings from the sampling experiments. Section 4 contains some conclusions.

The following notation is used throughout the paper. The symbol $\Rightarrow$ is used to signify weak convergence. The inequality > 0 denotes positive definite (p.d.) when applied to matrices. We use $I(d)$ to denote a time series that is integrated of order d. Vector Brownian motion with the covariance matrix Ω is written $BM(\Omega)$, and $MV(0, V)$ signifies a mixed normal distribution with the conditional variance matrix V. We write integrals with respect to Lebesgue measure such as $\int_0^1 W(s)\, ds$ more simply as $\int_0^1 W$ to achieve notational economy. All limits given in this paper are taken as the sample size tends to ∞ unless otherwise stated.

2. A BRIEF REVIEW OF FM-IV, FM-GMM, AND FM-GIVE

Before reporting our simulation results, we provide a brief review of the FM-IV estimation procedures with possibly integrated processes discussed by Kitamura and Phillips (1992). The purpose of this section is to illustrate our procedures in a rather simple model. For example, we will focus in this exposition on a single equation model, though the extension to multiple equation models is rather straightforward. The reader is referred to the cited paper for further details and discussion.

Consider the time series $\{y_t\}$ and $\{x_t\}$ generated by

$$y_t = \beta' x_t + u_{0t}, \tag{2.1.1}$$

$$H_1' x_t = x_{1t} = u_{1t} : m_1 \times 1, \tag{2.1.2}$$

$$H_2' \Delta x_t = \Delta x_{2t} = u_{2t} : m_2 \times 1, \tag{2.1.3}$$

where $H = [H_1, H_2]$ is an $m \times m$ orthogonal matrix. We also define the stochastic process of instrumental variables z_t by the following equations, which separate out the I(0) and I(1) components of the instruments:

$$G_1' z_t = z_{1t} : q_1 \times 1, \tag{2.2.1}$$

$$G_2' z_t = \Delta z_{2t} : q_2 \times 1, \tag{2.2.2}$$

where $G = [G_1, G_2]$ is a $q \times q$ orthogonal matrix.

As we shall see later, the preceding notation is helpful in the development of our theory, where we want to allow (explicitly or implicitly) for cointegrated regressors and instruments. We shall also extend our model to allow for cointegration between the regressors and the instruments. For that purpose, we add the following notation:

$$F_1' \begin{pmatrix} x_{2t} \\ z_{2t} \end{pmatrix} = v_{1t} : l_1 \times 1, \tag{2.3.1}$$

$$F_2' \begin{pmatrix} \Delta x_{2t} \\ \Delta z_{2t} \end{pmatrix} = v_{2t} : l_2 \times 1, \tag{2.3.2}$$

where $F = [F_1, F_2]$ is an $l \times l$ orthogonal matrix with $l = m_2 + q_2$. For later use, define $X' = (x_1, \ldots, x_T)$ and $Z' = (z_1, \ldots, z_T)$. We now impose assumptions on the random variables $w_t = (u_{0t}, u_{1t}', z_{1t}', v_{1t}', v_{2t}')'$.

Assumption EC (Error Condition).

(a) $\{w_t\}_1^\infty$ is fourth-order stationary with absolutely summable fourth-cumulant function.

(b) The partial sum process $s_t = \sum_0^t \{w_i\}$ satisfies the invariance principle

$$T^{-1/2} \sum_{i=1}^{[Tr]} w_i \Rightarrow B(r) = BM(\Omega), \qquad 0 \le r \le 1.$$

(c) The long-run variance matrix of $\{w_t\}$, $\Omega = \sum_{-\infty}^{+\infty} E(w_{t+i} w_t')$, is p.d.

Note that Assumption EC(b) can be replaced with more specific conditions in terms of mixing processes or linear processes (see, e.g., Phillips, 1991, and the papers cited therein). In fact, the invariance principle holds under very general conditions that cover many realistic cases that are encountered in practice. In sum, parts (b) and (c) of Assumption EC and equations (2.1.1)–(2.3.2) allow for the potential of cointegration among the regressors, among the instruments, and between the regressors and the instruments. We emphasize, however, that information about the number and location of the unit roots in the system is not required in the implementation of our FM-IV procedures. Thus, the investigator is assumed to have no knowledge about, for example, the matrices H and G that appear in equations (2.2) and (2.3). In fact, the fairly modest assumption of part (a) about w_t is used to show the rather convenient property of our FM-IV procedures that it is not necessary for the theory nor the practical implementation of the estimators that the precise breakdown or dimension of this vector's components be known a priori.

For later use, we decompose the long-run covariance matrix Ω as follows:

$$\Omega = \Sigma + \Lambda + \Lambda',$$

where $\Sigma = E(w_t w_t')$ and $\Lambda = \sum_{i=1}^\infty E(w_{t+i} w_t') = \sum_{i=1}^\infty \Gamma(i)$, say. Using this notation, we define the "one-sided long-run covariance matrix" $\Delta = \Sigma + \Lambda = \sum_{i=0}^\infty E(w_{t+i} w_t') = \sum_{i=0}^\infty \Gamma(i)$. Also, we partition Ω conformably with $w_t = (u_{0t}, (u_{1t}', z_{1t}', v_{1t}'), v_{2t}')' = (u_{0t}, w_{*t}', v_{2t}')'$ as

$$\Omega = \begin{bmatrix} \omega_{00} & \Omega_{0*} & \Omega_{0v_2} \\ \Omega_{*0} & \Omega_{**} & \Omega_{*v_2} \\ \Omega_{v_2 0} & \Omega_{v_2 *} & \Omega_{v_2 v_2} \end{bmatrix}.$$

The following assumptions are imposed on the score process $\phi_{z_{1t}} = u_{0t} z_{1t}$ and on the instrumental variables to validate the asymptotics.

Assumption IV (Instrument Validity Conditions and the Central Limit Theorem).

(a) $E[\phi z_{1t}] = E[u_{0t} z_{1t}] = 0$ for all t [orthogonality condition].
(b) $E[x_{1t} z_t'] = K_{z_1}$ is of full row rank (rank m_1) [relevance condition].
(c) $E[z_t z_t'] = M_z$ is nonsingular [nonsingular second moment].
(d) $T^{-1/2} \sum_0^T \phi_{z_{1t}} \Rightarrow N(0, S_{z_1})$, where $S_{z_1} = \sum_{-\infty}^{+\infty} R_{z_1}(i)$ and $R_{z_1}(j) = E(u_{0t} u_{0t+j} z_{1t} z_{1t+j}')$ [Central Limit Theorem].
(e) $m_2 \le q_2$ [order condition on I(1) instruments].

In many respects, these conditions are rather standard, as signified by the labeling we have provided for them in brackets. The reader is referred to Kitamura and Phillips (1992) for further discussion of these conditions in the present context.

The reason instrumental variables are introduced into our setup is to deal explicitly with the potential of endogenous regressors, especially those that are stationary. When the regressors are endogenous, two problems may arise. To see this point, rewrite equation (2.1.1) as

$$y_t = \beta_1' x_{1t} + \beta_2' x_{2t} + u_{0t}, \tag{2.4}$$

where $\beta_1 = H_1' \beta$ and $\beta_2 = H_2' \beta$. (Recall that this partition of the regressors is not known to the investigator.) As shown in Park and Phillips (1989), the application of OLS yields inconsistent estimators for the stationary coefficients β_1 with conventional $T^{1/2}$ asymptotics. On the other hand, the OLS estimator for the nonstationary coefficients β_2 is shown to be superconsistent with the convergence rate T, though its asymptotic distribution is not median unbiased and depends on nuisance parameters. Of course, we can solve the first problem concerning the stationary components by the use of valid instruments, though it does not solve the second (see Phillips and Hansen, 1990). In the literature, there are several estimators available that overcome the second problem, but all of these rely on prior information (possibly from pretests) about the number and the location of unit roots (i.e., information about H_1 and H_2).

Kitamura and Phillips (1992) showed that the FM-IV procedure has the very nice asymptotic property that it is well behaved when the regressors and instruments are possibly nonstationary, that is, when the number and the location of unit roots are uncertain. FM versions of the GMM and the GIVE estimators were introduced to provide efficiency gains over FM-IV. Recall that in the original GIVE procedure (e.g., Sargan, 1988), we employ GLS-type transformations of both the regressors and the instruments using a matrix $W_T (= \{W_{ij}\})$, say, such that $(W_T' W_T)^{-1} = \text{Var}(u_0) = V_{0T}$.

The following is a catalog of the FM estimators with which we shall be concerned in the rest of the paper:

The FM-IV estimator

$$\bar{\beta} = (X' P_Z X)^{-1} X' Z (Z'Z)^{-1} [Z' \hat{y}^+ - \hat{\Delta}_{0\Delta z}^+] \tag{2.5.1}$$

The FM-GMM estimator

$$\bar{\beta}_{\mathrm{GMM}} = (X'ZS_{zT}^{-1}Z'X)^{-1}X'ZS_{zT}^{-1}[Z'\hat{y}^{+} - \hat{\Delta}_{0\Delta z}^{+}] \tag{2.5.2}$$

The FM-GIVE estimator

$$\bar{\beta}_{\mathrm{GIVE}} = (X^{*'}P_{Z^{*}}X^{*})^{-1}X^{*'}Z^{*}(Z^{*'}Z^{*})^{-1}[Z^{*'}\hat{y}^{*+} - \hat{\Delta}_{0\Delta z^{*}}^{+}], \tag{2.5.3}$$

where $\hat{y}^{+} = y - \hat{\Omega}_{0a}\hat{\Omega}_{aa}^{-1}u_{a}$, $\hat{\Delta}_{0\Delta z}^{+} = \hat{\Delta}_{0\Delta z} - \hat{\Omega}_{0a}\hat{\Omega}_{aa}^{-1}\hat{\Delta}_{a\Delta z}$, $\hat{y}^{*+} = y^{*} - \hat{\Omega}_{0a}\hat{\Omega}_{aa}^{-1}u_{a}$, and $\hat{\Delta}_{0\Delta z^{*}}^{+} = \hat{\Delta}_{0\Delta z^{*}} - \hat{\Omega}_{0a}\hat{\Omega}_{aa}^{-1}\hat{\Delta}_{a\Delta z}$. The affix $*$ indicates premultiplication by the matrix $\hat{W}_{t}$, which is a consistent estimator of W_{t}, and the subscript a signifies that elements corresponding to Δx_{t} and Δz_{t} are taken together.

To construct estimators of the long-run covariance matrices Δ and Ω, we employ kernel estimators of the general form

$$\hat{\Omega}_{ab} = \sum_{j=-T+1}^{T-1} w(j/K)\hat{\Gamma}_{ab}(j) \quad \text{and} \quad \hat{\Delta}_{ab} = \sum_{j=0}^{T-1} w(j/K)\hat{\Gamma}_{ab}(j), \tag{2.6}$$

where $\hat{\Gamma}_{ab}(j) = T^{-1}\sum \hat{u}_{at+j}\hat{u}'_{bt}$, $w(\cdot)$ is a kernel function, and K is a truncation or bandwidth parameter. Accordingly, the notation $\hat{\Delta}_{0\Delta z^{*}}$ implies the use of transformed residual estimates in the calculation of the kernel smoothing estimators.

In practice, we need residuals estimates $\hat{u}_{t}$ to calculate the sample autocovariances $\hat{\Gamma}$ that appear in (2.6), and we may use, for example, the "crude IV estimator" $\hat{\beta} = (X'P_{Z}X)^{-1}X'P_{Z}Y$ to obtain them. The metric matrix S_{zT} in the FM-GMM estimator can be estimated, for example, by the use of a kernel estimator or a smoothed periodogram estimator.

To allow for possibly nonstationary processes, we need to make the following additional assumptions.

Assumption LR (Long-Run Covariance Matrix Estimation).

(a) Any of the Parzen, quadratic spectral (QS), or Tukey–Hanning kernels[1] are used in the estimation of Ω and Δ.

(b) The truncation/bandwidth parameter K grows at the rate of T^{k} for some $k \in (\frac{1}{4}, \frac{2}{3})$.

(c) The covariance functions $\Gamma_{u_{0}v_{2}}(\cdot)$ and $\Gamma_{w_{*}v_{2}}(\cdot)$ satisfy

$$\sum_{-\infty}^{\infty} j\|\Gamma(j)\| < \infty.$$

Assumption NF (No Feedback).

$E(u_{0t+j}z_{1t}) = 0 \quad \text{for all } j \geq 1.$

Assumption NF* (NF for Transformed Processes).

$E(u_{0t+j}^{*}z_{1t}^{*}) = 0 \quad \text{for all } j \geq 1$, where $u_{0t}^{*} = \sum_{\tau=1}^{T} W_{t\tau}u_{0\tau}$ and $z_{1t}^{*} = \sum_{\tau=1}^{T} W_{t\tau}z_{1\tau}$.

In Assumption LR, only part (c) is a condition concerning the property of the data, and it is, in fact, fairly modest. For instance, this summability con-

dition allows for general finite order stationary vector ARMA as the DGP of $\{w_t\}_0^\infty$.

The limit theory for these three estimators is as follows.

THEOREM. *Suppose Assumptions* EC, IV, *and* LR *hold. Also, Assumption* NF *holds for* $\bar{\beta}$ *and* $\bar{\beta}_{GMM}$ *and Assumption* NF* *holds for* $\bar{\beta}_{GIVE}$. *Then,*

(a) $T^{1/2}H_1'(\bar{\beta} - \beta)(= T^{1/2}(\bar{\beta}_1 - \beta_1)) \Rightarrow N(0, J_{z_1}S_{z_1}J_{z_1}')$,

(b) $T^{1/2}H_1'(\bar{\beta}_{GMM} - \beta)(= T^{1/2}(\bar{\beta}_{1GMM} - \beta_1)) \Rightarrow N(0, [K_{z_1}S_{z_1}^{-1}K_{z_1}']^{-1})$,

(c) $T^{1/2}H_1'(\bar{\beta}_{GIVE} - \beta)(= T^{1/2}(\bar{\beta}_{1GIVE} - \beta_1)) \Rightarrow N(0, [K_{z_1}^V M_{z_1}^V K_{z_1}^{V'}]^{-1})$, *and*

(d) $TH_2'(\bar{\beta} - \beta), TH_2(\bar{\beta}_{GMM} - \beta), TH_2(\bar{\beta}_{GIVE} - \beta)(= T(\bar{\beta}_2 - \beta_2), T(\bar{\beta}_{2GMM} - \beta_2),$
$T(\bar{\beta}_{2GIVE} - \beta_2)) \Rightarrow MN(0, \omega_{00 \cdot v_2}\int_0^1 \tilde{B}_2\tilde{B}_2')$.

See Kitamura and Phillips (1992) for the proof of this theorem.

Remarks.

(a) In the preceding theorem, $J_{z_1} = (K_{z_1}M_{z_1}^{-1}K_{z_1}')K_{zz}M_{z_1}$, $K_{z_1}^V = \text{plim}(T^{-1}X_1^{*'}Z_1^*)$, $M_{z_1}^V = \text{plim}(T^{-1}Z_1^{*'}Z_1^*)$, $\omega_{00 \cdot v_2} = \omega_{00} - \Omega_{0v_2}\Omega_{v_2v_2}^{-1}\Omega_{v_20}$, $\tilde{B}_2 = \int_0^1 B_2B_z'(\int_0^1 B_zB_z')^{-1}B_z$, and B_2 and B_z signify Brownian motions that are defined by the limits $T^{-1/2}\sum_0^{[Tr]}u_{2t} \Rightarrow B_2(r)$ and $T^{-1/2}\sum_0^{[Tr]}\Delta z_2 \Rightarrow B_{z_2}(r)$.

(b) Under fairly general conditions, we have the inequality

$$[K_{z_1}^V M_{z_1}^V K_{z_1}^{V'}]^{-1} \le [K_{z_1}S_{z_1}^{-1}K_{z_1}']^{-1} \le J_{z_1}S_{z_1}J_{z_1}',$$

as noted by Kitamura and Phillips (1992).

We shall investigate the relative efficiency among these estimators in finite samples by simulation in the next section.

3. SIMULATION RESULTS

This section provides simulation results concerning the various FM-IV procedures. Sections 3.1 and 3.2 contain a description of the experimental design and estimation methods. In Section 3.3, we present the simulation results and summarize the findings from them.

3.1. Experimental Design

We are interested in the estimation of the coefficients $b = (b_1, b_2)'$ in the equation

$$\underset{1\times1}{y_t} = \underset{1\times2}{b'} \quad \underset{2\times1}{x_t} + \underset{1\times1}{u_{0t}}, \tag{3.1}$$

where $b = (0.2, 0.5)'$. Both the regressors and the error terms are linear combinations of the two processes shown next:

$$x_t = \alpha x_{at} + (1 - \alpha)x_{bt}, \tag{3.2}$$

$$u_{0t} = \alpha u_{at} + (1 - \alpha)u_{nt}, \tag{3.3}$$

with $0 \le \alpha \le 1$ and

$$x_{at} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} x_{at-1} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot u_{at}, \tag{3.4}$$

where u_{at} and u_{nt} are not cross-correlated and are specified later in (3.11). Nevertheless, because the term u_{at} (and, thus, x_{at}) causes feedback from u_{0t} to x_t through the parameter α, instrumental variables are necessary for consistent estimation when x_t has some stationary components. The instruments z_t are generated by

$$z_t = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} z_{t-1} + u_{zt}, \tag{3.5}$$

where the innovation u_{zt} is given by

$$u_{zt} = \varepsilon_{zt} + \Gamma\varepsilon_{zt-1} \quad \text{with } \Gamma = \begin{bmatrix} 0.6 & -0.6 & 0.0 \\ -0.6 & -0.6 & 0.4 \\ 0.4 & 0.4 & 0.4 \end{bmatrix}. \tag{3.6}$$

The regressors and the instruments were related by

$$x_{bt} = \Pi_0' z_{bt} \quad \text{with } \Pi_0' = \begin{pmatrix} 0.5 & 0.5 & 0.0 \\ 0.0 & 0.0 & 1.0 \end{pmatrix}, \tag{3.7}$$

$$z_{bt} = \beta z_t + (1 - \beta) w_t, \qquad 0 \le \beta \le 1, \tag{3.8}$$

where the DGP of u_{zt} is

$$w_t = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} w_{t-1} + u_{wt}, \qquad u_{wt} = \varepsilon_{wt} + \Gamma\varepsilon_{wt-1}. \tag{3.9}$$

Of course, β is the parameter that controls the relevance of the instruments. Formulations (3.7)–(3.9) have two advantages. First, in view of (3.7), x_{bt} has the same VAR(1) form as x_{at} in (3.4), so that x_t is also a VAR(1) process and follows

$$x_t = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} x_{t-1} + u_{xt}, \tag{3.10}$$

where

$$u_{xt} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \alpha u_{at} + \Pi_0'(1 - \alpha)[\beta u_{zt} + (1 - \beta)u_{wt}].$$

Second, not only are the instruments valid, but also it is easy to see that x_t has a "reduced form" in terms of z_t with the reduced form coefficients $(1 - \alpha)\beta\Pi_0'$. Using this fact, we can show by conventional arguments that FM-GIVE for the I(0) coefficients is asymptotically efficient over FM-GMM (and FM-IV) (see, e.g., White, 1984). To explore the effect of both feedback and relevance, we shall report simulation results for various values for α and β.

Finally, the regression errors u_{0t} in equation (3.3) are driven by

$$u_{at} = \varepsilon_{at} + \theta\varepsilon_{at-1}, \tag{3.11.1}$$

$$u_{nt} = \varepsilon_{nt} + \theta\varepsilon_{nt-1}, \tag{3.11.2}$$

with $\theta = 0.9$. Consequently, u_{0t} is also MA(1):

$$u_{0t} = \varepsilon_{0t} + \theta\varepsilon_{0t-1}, \tag{3.12}$$

with $\varepsilon_{0t} = \alpha\varepsilon_{at} + (1 - \alpha)\varepsilon_{nt}$. The roots of the AR processes were parameterized as follows:

$$(\lambda_1, \lambda_2) = \begin{cases} (0.8, 0.8) & \text{the "I(0)/I(0) model,"} \\ (0.8, 1.0) & \text{the "I(0)/I(1) model,"} \\ (1.0, 1.0) & \text{the "I(1)/I(1) model."} \end{cases}$$

In each simulation, $\varepsilon_t = (\varepsilon_a, \varepsilon_n, \varepsilon_w', \varepsilon_z')'$ were i.i.d., $N(0, I_8)$, and the initial values $u_{z0} = (0,0,0)'$, $\varepsilon_{a0} = 0$, $\varepsilon_{n0} = 0$, $x_{a0} = (0,0)'$, $z_{b0} = (0,0,0)'$, $\varepsilon_{w0} = (0,0,0)'$, and $\varepsilon_{z0} = (0,0,0)'$ were used. All the results are based on 2,000 simulations of sample size 100. The pseudonormal random number generator in GAUSS was used to generate the samples.

3.2. Estimation Methods

The simulation results reported in the next section are concerned with the performance of the following five estimators: the OLS, crude IV, FM-IV, FM-GMM, and FM-GIVE estimators. The long-run covariances Δ and Ω and the metric matrix S_{zT} were estimated by the use of the Parzen kernel

$$w(x) = \begin{cases} 1 - 6x^2 + 6|x|^3 & \text{for } 0 \le |x| \le \frac{1}{2}, \\ 2(1 - |x|)^3 & \text{for } \frac{1}{2} \le |x| \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Results with other kernels, such as the Tukey-Hanning and QS kernels are not reported, because no essential differences were found in the general qualitative features. The lag lengths of the kernels used in $\hat{\Delta}$ and $\hat{\Omega}$, which we designate KLR, were chosen to be 4 in all but the last set of simulations, where various values for KLR were used. Although issues of optimal choice of the

truncation parameter for kernel estimators of long-run covariances are discussed in the literature (e.g., Andrews, 1991), little guidance is currently available for choosing KLR in a setup as complex as ours, because the introduction of possibly nonstationary processes substantially complicates the problem. Moreover, there is no reason why "optimal" choices of KLR as far as the estimation of Δ and Ω should apply or even be relevant in the estimation of the regression parameter β.[2] This is a problem of some importance that we leave for future research. The kernel lag length used in S_{zT}, which we designate KSZ, is set to be 3 throughout all the simulations reported here (results with other values of KSZ varied very little and are not reported here). In the construction of the FM-GIVE estimators, we calculated the GLS-transformation matrix $\hat{W}_T$ by fitting AR(3) models to the residual processes. Note that the true error process of (3.12) is MA(1) and, thus, this transformation is misspecified. The effect of this misspecification is discussed later.

To implement the FM-IV method, we need to estimate Ω and Δ, and the following two-step procedure was used: (i) obtain initial estimates $\hat{u}_{0t}$ using the crude IV estimator, and (ii) use $\hat{u}_{0t}$ to calculate the long-run covariances and construct the FM-IV estimator $\bar{b}$. The calculation of the FM-GMM estimators requires a few more complicated steps: (i) obtain initial estimates $\hat{u}_{0t}$ as before, (ii) use $\hat{u}_{0t}$ in the calculation of the metric matrix S_{zT} in the (unmodified) GMM estimator $\hat{b}_{\text{GMM}} = (X'ZS_{zT}^{-1}Z'X)^{-1}X'ZS_{zT}^{-1}Z'y$ and get the GMM residuals $\hat{u}_{\text{GMM}t}$, and (iii) use $\hat{u}_{\text{GMM}t}$ to calculate S_{zT} again and the long-run covariances and then form the FM-GMM estimator $\bar{b}_{\text{GMM}}$. Similarly, the FM-GIVE procedure was implemented using the following steps: (i) obtain the initial residual $\hat{u}_{0t}$ by crude IV, (ii) fit an AR(3) model to the estimated process $\hat{u}_{0t}$; use the estimated AR coefficients to transform the data matrices to form $\hat{b}_{\text{GIVE}} = (X^{*\prime}P_{Z^*}X^*)^{-1}X^{*\prime}P_{Z^*}y^*$ and calculate the GIVE residuals $\hat{u}_{\text{GIVE}t}$, and (iii) fit an AR(3) model to $\hat{u}_{\text{GIVE}t}$ again to perform the GLS transformations as in step (ii); also, use $\hat{u}_{\text{GIVE}t}$ to form the long-run covariance matrix estimates and construct the FM-GIVE $\bar{b}_{\text{GIVE}}$. All calculations were performed by using programs written in GAUSS.

3.3. Findings from the Simulations

In this section, we present a summary of our simulation results and discuss their implications. The five estimation techniques were applied to the I(0)/I(0), I(0)/I(1), and I(1)/I(1) models for each parameterization of $(\alpha, \beta, \text{KLR})$. Recall that α is the feedback parameter, β is the relevance parameter, and KLR is the truncation parameter used in $\hat{\Omega}$ and $\hat{\Delta}$. First we focus on the simulation results for the "baseline" case, and then we turn to the analysis of the estimators' sensitivity to each of these three parameters in turn. In what follows, sometimes we use the terminology "I(0) coefficients," referring to (b_1, b_2) in the I(0)/I(0) models and b_1 in the I(0)/I(1) models, and "I(1)

coefficients," referring to b_2 in the I(0)/I(1) models and (b_1, b_2) in the I(1)/I(1) models.

As indicated by the results in Tables 1, 4, and 7, sometimes the mean absolute deviation (MAD), the average bias (BIAS$_{ave}$), and the root mean square error (RMSE) of the FM estimators are exceptionally high due to the very occasional occurrence of extremely large estimation errors. Thus, Table 1 also includes the bias and RMSE based on the results excluding 1% in both tails, to see the performance of the estimators without the effect of the occasional outliers. We will sometimes call these statistics BIAS$_{.98}$ and RMSE$_{.98}$. Tables 3 and 6 contain quantiles that are robust to outliers.

Figures 1a and b display nonparametric estimates of the probability density functions (p.d.f.'s) of estimators for the I(0) and I(1) coefficients in the I(0)/I(1) model for the "baseline" case, where $\alpha = 0.5$, $\beta = 0.8$, and
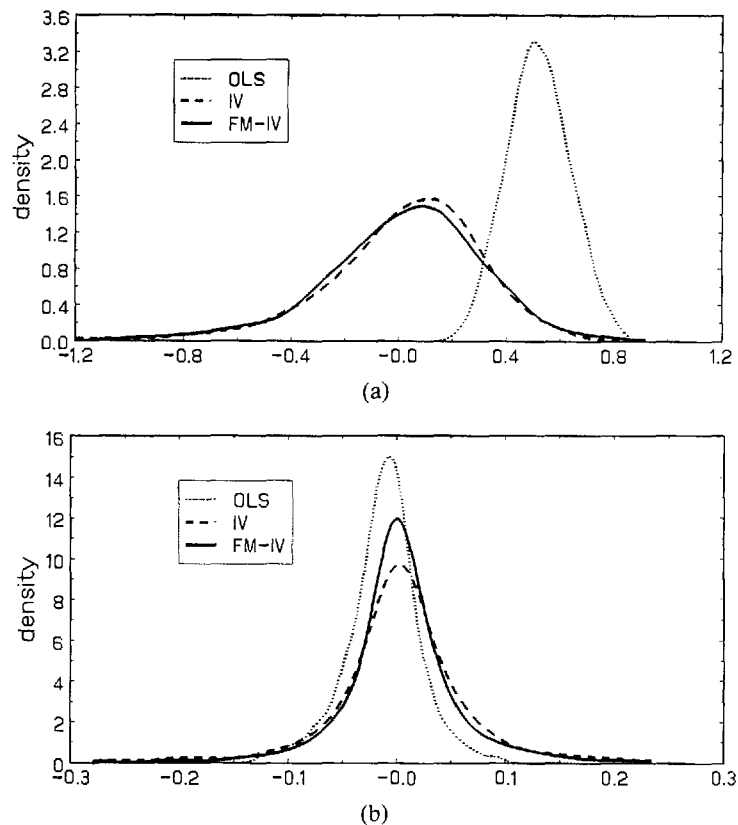

FIGURE 1. I(0)/I(1) model. (a) Coefficient b_1, regressor I(0), $\alpha = 0.5$, $\beta = 0.8$. (b) Coefficient b_2, regressor I(1), $\alpha = 0.5$, $\beta = 0.8$.

$KLR = 4$. The figure shows the basic characteristics of the OLS, crude IV, and FM-IV estimators quite clearly.

The p.d.f.'s of the estimators of the $I(0)$ coefficients displayed in Figure 1a show that OLS is evidently biased due to the feedback. On the other hand, given the valid instruments, the estimated densities of crude IV and FM-IV are well centered. These results are in accordance with what theory predicts. Figure 1b graphs the sampling densities of the three estimators applied to the $I(1)$ coefficient. Even in this case, OLS is considerably biased, although the bias is generally smaller than the $I(0)$ cases due to the superconsistency. Note the dispersion/bias trade-off between OLS and crude IV. Asymptotic theory tells us that in the estimation of the $I(1)$ coefficients, both OLS and IV are superconsistent but both also suffer from second-order bias (Phillips and Hansen, 1990). According to the sampling results displayed in Figure 1b, the introduction of instruments for the nonstationary regressor seems to reduce bias. However, that is not enough—the crude IV estimator seems deficient compared with the FM-IV estimator. These results seem to imply the desirability of FM-IV: its sampling distributions are centered and reasonably concentrated.

As seen from the theorem in Section 2, FM-IV and crude IV are asymptotically equivalent with respect to the $I(0)$ coefficients. This means that the advantage of FM-IV with respect to the $I(1)$ coefficients can be obtained at no cost asymptotically. Therefore, it is naturally interesting to ask how the additional correction terms in FM-IV affect the finite sample performance of the $I(0)$ coefficient estimators. In fact, the estimated densities in Figure 1a seem very interesting: the distribution of FM-IV is close to that of crude IV, suggesting that the price we have to pay in order to obtain the efficiency gain in the estimation of the $I(1)$ coefficients seems reasonable.

At the same time, Table 1 shows some summary statistics calculated for the baseline case. As expected, for the $I(0)$ coefficients, the OLS estimators are substantially biased compared to other estimators. On the other hand, the results concerning crude IV and FM-IV (or FM-GMM, FM-GIVE) seem somewhat less clear. First, there is nothing to choose between crude IV and FM-IV in terms of median and $BIAS_{.98}$, as both theory and Figure 1(a) imply, and this is also true for other parameterizations discussed later. However, $BIAS_{ave}$ and MAD of FM-IV are generally higher than those of crude IV. Also, we see a similar situation among the dispersion measures. In terms of $Q_{.25}$, $Q_{.75}$, and $RMSE_{.98}$ the dispersion of FM-IV applied to the $I(0)$ coefficients tends to be larger than that of crude IV; however, the difference is not large. On the other hand, in terms of $RMSE_{ave}$, the difference becomes larger. These conflicts are, of course, due to the aforementioned occasional extremely large errors.

Table 1 also shows the results on the estimation of the $I(1)$ coefficients. The relatively good performance of the FM-IV procedure observed in Figure 1b can be seen here. The bias of OLS is substantial, and FM-IV outper-

TABLE 1. Summary statistics: Baseline case ($\alpha = 0.5$, $\beta = 0.8$, KLR $= 4$)

Model		Estimator	Quantiles			MAD	Other statistics			
			$Q_{.25}$	$Q_{.50}$	$Q_{.75}$		$BIAS_{ave}$	$RMSE_{ave}$	$BIAS_{.98}$	$RMSE_{.98}$
I(0)/I(0)	b_1	OLS	0.3605	0.4764	0.5735	0.4685	0.4685	0.4933	0.4683	0.4901
		IV	-0.1630	0.0376	0.2131	0.2395	0.0013	0.3323	0.0086	0.2790
		FM-IV	-0.1900	0.0477	0.2373	0.2945	-0.0236	0.6278	0.0000	0.3302
		FM-GMM	-0.1739	0.0491	0.2236	0.2766	-0.0181	0.5529	0.0028	0.3133
		FM-GIVE	-0.1248	0.0245	0.1599	0.1847	0.0034	0.2425	0.0077	0.2181
	b_2	OLS	-0.0454	0.0282	0.1125	0.0988	0.0353	0.1265	0.0348	0.1177
		IV	-0.0937	0.0042	0.0994	0.1270	-0.0098	0.2197	-0.0021	0.1440
		FM-IV	-0.1002	0.0063	0.1055	0.1601	-0.0334	0.7765	-0.0046	0.1608
		FM-GMM	-0.1022	0.0074	0.1044	0.1561	-0.0302	0.6733	-0.0038	0.1592
		FM-GIVE	-0.0814	0.0081	0.0932	0.1117	-0.0065	0.1544	-0.0029	0.1316
I(0)/I(1)	b_1	OLS	0.4413	0.5165	0.6006	0.5208	0.5208	0.5346	0.5206	0.5326
		IV	-0.1279	0.0581	0.2145	0.2374	0.0145	0.3562	0.0248	0.2771
		FM-IV	-0.1463	0.0500	0.2152	0.2851	0.0437	1.5373	0.0238	0.2826
		FM-GMM	-0.1469	0.0460	0.2070	0.2658	0.0188	0.8263	0.0186	0.2834
		FM-GIVE	-0.1268	0.0280	0.1653	0.1904	0.0055	0.2608	0.0086	0.2238

(continued)

TABLE 1 (continued)

Model		Estimator	Quantiles			MAD	Other statistics			
			$Q_{.25}$	$Q_{.50}$	$Q_{.75}$		BIAS_{ave}	RMSE_{ave}	$\text{BIAS}_{.98}$	$\text{RMSE}_{.98}$
I(0)/I(1) continued										
	b_2	OLS	-0.0300	-0.0102	0.0062	0.0274	-0.0125	0.0380	-0.0123	0.0338
		IV	-0.0229	0.0033	0.0308	0.0582	-0.0037	0.1997	0.0021	0.0694
		FM-IV	-0.0183	0.0013	0.0246	0.0950	0.0445	1.8727	0.0029	0.0581
		FM-GMM	-0.0181	0.0013	0.0245	0.0740	0.0225	0.9064	0.0032	0.0587
		FM-GIVE	-0.0197	0.0013	0.0252	0.0580	0.0027	0.1582	0.0016	0.0718
I(1)/I(1)	b_1	OLS	0.0153	0.0701	0.1486	0.1081	0.0860	0.1446	0.0848	0.1353
		IV	-0.0625	0.0172	0.0969	0.1450	0.0133	0.2777	0.0148	0.1741
		FM-IV	-0.0637	0.0053	0.0689	0.1718	0.0267	1.4932	0.0003	0.1611
		FM-GMM	-0.0599	0.0055	0.0693	0.1751	0.0293	1.5487	0.0010	0.1614
		FM-GIVE	-0.0674	0.0077	0.0826	0.1149	0.0069	0.1845	0.0060	0.1377
	b_2	OLS	-0.0512	-0.0026	0.0469	0.0627	-0.0019	0.0841	-0.0019	0.0750
		IV	-0.0552	0.0021	0.0602	0.0911	0.0070	0.1675	0.0038	0.1061
		FM-IV	-0.0444	0.0014	0.0476	0.0986	-0.0010	0.6450	0.0041	0.0972
		FM-GMM	-0.0441	0.0019	0.0465	0.0990	-0.0024	0.6611	0.0040	0.0950
		FM-GIVE	-0.0438	-0.0001	0.0485	0.0765	0.0076	0.1497	0.0041	0.0896

forms crude IV in terms of quantiles and $RMSE_{.98}$. On the other hand, the effect of outliers results in the higher $RMSE_{ave}$ of FM-IV.

Table 2 displays the concentration probabilities of the various estimators for the $I(0)/I(1)$ model in the baseline case. The disadvantage of OLS in the estimation of the $I(0)$ coefficients is evident, whereas OLS for the $I(1)$ coefficients seems to work reasonably well. By comparing crude IV and FM-IV, we see that crude IV is only marginally better for b_1 (the $I(0)$ coefficient), whereas FM-IV beats crude IV in the estimation of b_2 (the $I(1)$ coefficient). (Check the final line in each panel below the concentration probabilities of b_1 and b_2 to see the relative difference between the concentration probabilities of crude IV and FM-IV.)

The preceding results can be summarized as follows: the estimated densities and quantiles seem to imply that FM-IV outperforms OLS and crude IV in the estimation of the $I(1)$ coefficients, whereas OLS is severely biased and the performance of FM-IV and crude IV are very close in the estimation of the $I(0)$ coefficients. Thus, in view of these results, FM-IV seems to be a desirable choice. However, also note that FM-IV has a higher RMSE than crude IV due to occasional extreme values, in fact in every case.

A closer reading of the estimation results shows that these outliers are due to poor initial estimates obtained by the use of crude IV regressions in the

TABLE 2. Concentration probabilities: $(I(0)/I(1)$ model, $\alpha = 0.5$, $\beta = 0.8$, KLR $= 4)$

				$P(|\bar{b} - b| \leq c)$		
	Estimator	(b_1) (b_2)	$c = 0.1$ $c = 0.01$	$c = 0.3$ $c = 0.03$	$c = 0.5$ $c = 0.05$	$c = 0.7$ $c = 0.07$	$c = 0.9$ $c = 0.09$
b_1	OLS		0.0000	0.0280	0.4375	0.9270	0.9980
	IV		0.2855	0.7290	0.9135	0.9650	0.9810
	FM-IV		0.2825	0.7175	0.9020	0.9590	0.9805
	FM-GMM		0.2945	0.7265	0.9025	0.9590	0.9760
	FM-GIVE		0.3495	0.8105	0.9500	0.9835	0.9925
	$P(|b_{IV} - b| \leq c)/$ $P(|b_{FM\text{-}IV} - b| \leq c)$		1.0106	1.0160	1.0127	1.0063	1.0005
b_2	OLS		0.2880	0.6675	0.8445	0.9230	0.9620
	IV		0.2320	0.5340	0.7000	0.7965	0.8590
	FM-IV		0.2835	0.6080	0.7570	0.8415	0.8840
	FM-GMM		0.2730	0.6070	0.7605	0.8360	0.8825
	FM-GIVE		0.2690	0.5845	0.7255	0.8150	0.8645
	$P(|b_{FM\text{-}IV} - b| \leq c)/$ $P(|b_{IV} - b| \leq c)$		1.2220	1.1386	1.0814	1.0565	1.0291

first stage. That is, because the long-run covariance estimators used in the FM estimators are weighted sums of sample covariances, estimation errors cumulate. Thus, when the initial IV estimates are exceptionally poor due to, for example, poor instruments, in the second stage the FM procedure can amplify the effect of poor preliminary estimates.

Next we investigate the effect of the feedback from the residuals to the regressors. Table 3 displays the quantiles of each estimator with $\alpha = 0.1$, 0.3, 0.7, $\beta = 0.8$, and KLR $= 4$. In each case, the qualitative nature is basically unchanged from the baseline case. As expected, the bias of OLS increases as the feedback becomes higher, even in the estimation of the I(1) coefficient, and the same characterization applies to crude IV (see also Figures 2a and b and 3a and b). At the same time, in terms of quantiles, FM-IV still performs
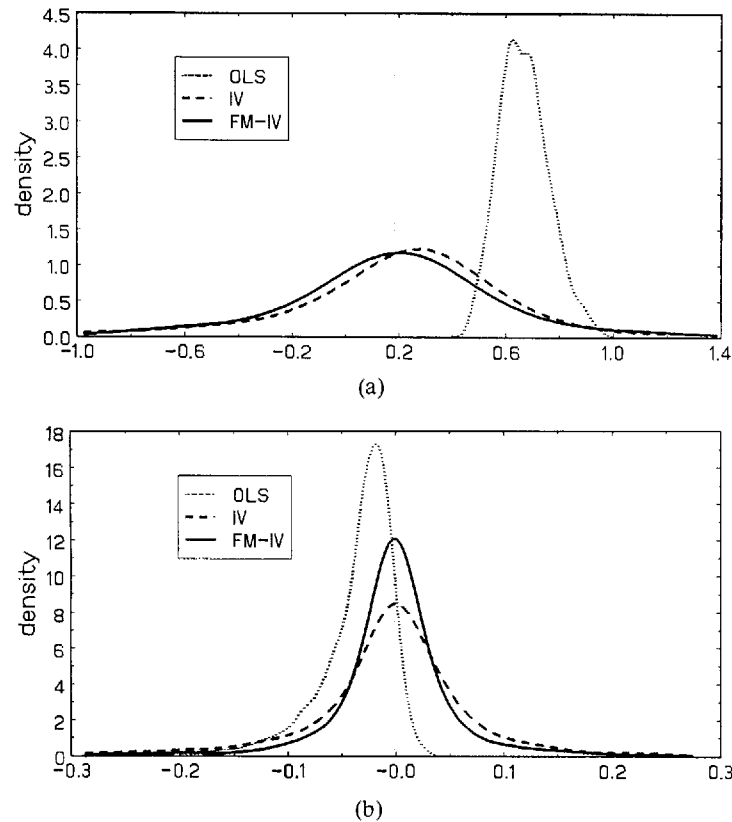

(a)

(b)

FIGURE 2. I(0)/I(1) model. (a) Coefficient b_1, regressor I(0), $\alpha = 0.7$, $\beta = 0.8$. (b) Coefficient b_2, regressor I(1), $\alpha = 0.7$, $\beta = 0.8$.

TABLE 3. Quantiles ($\beta = 0.8$, KLR $= 4$)

α	Model	Estimator	b_1			b_2		
			$Q_{.25}$	$Q_{.50}$	$Q_{.75}$	$Q_{.25}$	$Q_{.50}$	$Q_{.75}$
0.1	I(0)/I(0)	OLS	-0.1096	0.0281	0.1545	-0.0648	0.0001	0.0676
		IV	-0.1429	0.0059	0.1246	-0.0681	-0.0014	0.0687
		FM-IV	-0.1802	-0.0004	0.1597	-0.0767	0.0020	0.0815
		FM-GMM	-0.1690	0.0011	0.1475	-0.0760	0.0019	0.0777
		FM-GIVE	-0.1317	-0.0208	0.0901	-0.0850	-0.0075	0.0617
	I(0)/I(1)	OLS	-0.1049	0.0241	0.1504	-0.0177	0.0008	0.0187
		IV	-0.1457	-0.0045	0.1237	-0.0181	0.0013	0.0203
		FM-IV	-0.1757	-0.0085	0.1514	-0.0192	0.0008	0.0217
		FM-GMM	-0.1717	-0.0061	0.1415	-0.0187	0.0009	0.0217
		FM-GIVE	-0.1072	-0.0068	0.0973	-0.0180	0.0014	0.0206
	I(1)/I(1)	OLS	-0.0652	0.0068	0.0799	-0.0292	0.0010	0.0289
		IV	-0.0724	0.0011	0.0804	-0.0297	0.0000	0.0297
		FM-IV	-0.0802	0.0013	0.0835	-0.0307	0.0003	0.0340
		FM-GMM	-0.0787	-0.0005	0.0810	-0.0321	0.0004	0.0344
		FM-GIVE	-0.0727	-0.0013	0.0743	-0.0278	0.0010	0.0297
0.3	I(0)/I(0)	OLS	0.1197	0.2346	0.3550	-0.0529	0.0152	0.0852
		IV	-0.1553	0.0074	0.1399	-0.0750	0.0022	0.0743
		FM-IV	-0.1927	0.0058	0.1681	-0.0850	0.0057	0.0863
		FM-GMM	-0.1780	0.0100	0.1553	-0.0870	0.0053	0.0830
		FM-GIVE	-0.1140	-0.0004	0.1117	-0.0739	0.0011	0.0693

(continued)

TABLE 3 (continued)

| α	Model	Estimator	b_1			b_2		
			$Q_{.25}$	$Q_{.50}$	$Q_{.75}$	$Q_{.25}$	$Q_{.50}$	$Q_{.75}$
0.3 continued								
	I(0)/I(1)	OLS	0.1399	0.2449	0.3619	-0.0198	-0.0003	0.0180
		IV	-0.1466	0.0042	0.1411	-0.0204	0.0015	0.0212
		FM-IV	-0.1716	0.0037	0.1655	-0.0189	0.0013	0.0229
		FM-GMM	-0.1704	0.0074	0.1612	-0.0192	0.0015	0.0231
		FM-GIVE	-0.1134	0.0011	0.1116	-0.0181	0.0016	0.0220
	I(1)/I(1)	OLS	-0.0200	0.0399	0.1131	-0.0349	-0.0010	0.0318
		IV	-0.0721	0.0065	0.0875	-0.0352	-0.0012	0.0373
		FM-IV	-0.0736	0.0002	0.0824	-0.0360	0.0001	0.0410
		FM-GMM	-0.0788	0.0009	0.0850	-0.0359	0.0001	0.0393
		FM-GIVE	-0.0746	0.0049	0.0785	-0.0322	-0.0008	0.0361
0.7	I(0)/I(0)	OLS	0.4509	0.5795	0.6945	-0.0786	0.0426	0.1540
		IV	-0.0470	0.2231	0.4503	-0.1133	0.0378	0.1845
		FM-IV	-0.0129	0.2536	0.4379	-0.1110	0.0177	0.1535
		FM-GMM	-0.0164	0.2336	0.4247	-0.1003	0.0277	0.1664
		FM-GIVE	-0.0729	0.1618	0.3562	-0.0779	0.0450	0.1695
	I(0)/I(1)	OLS	0.6002	0.6602	0.7261	-0.0492	-0.0264	-0.0123
		IV	0.0168	0.2460	0.4429	-0.0301	-0.0000	-0.0302
		FM-IV	-0.0153	0.1973	0.4077	-0.0184	-0.0005	0.0182
		FM-GMM	-0.0312	0.1899	0.4011	-0.0192	-0.0007	0.0185
		FM-GIVE	-0.0931	0.1516	0.3472	-0.0213	-0.0003	0.0253
	I(1)/I(1)	OLS	-0.0125	0.0878	0.2003	-0.0989	-0.0036	0.0927
		IV	-0.0939	0.0240	0.1481	-0.1094	0.0007	0.1110
		FM-IV	-0.0559	0.0056	0.0747	-0.0571	0.0022	0.0617
		FM-GMM	-0.0573	0.0068	0.0726	-0.0565	0.0017	0.0619
		FM-GIVE	-0.0825	0.0117	0.1085	-0.0771	-0.0012	0.0758
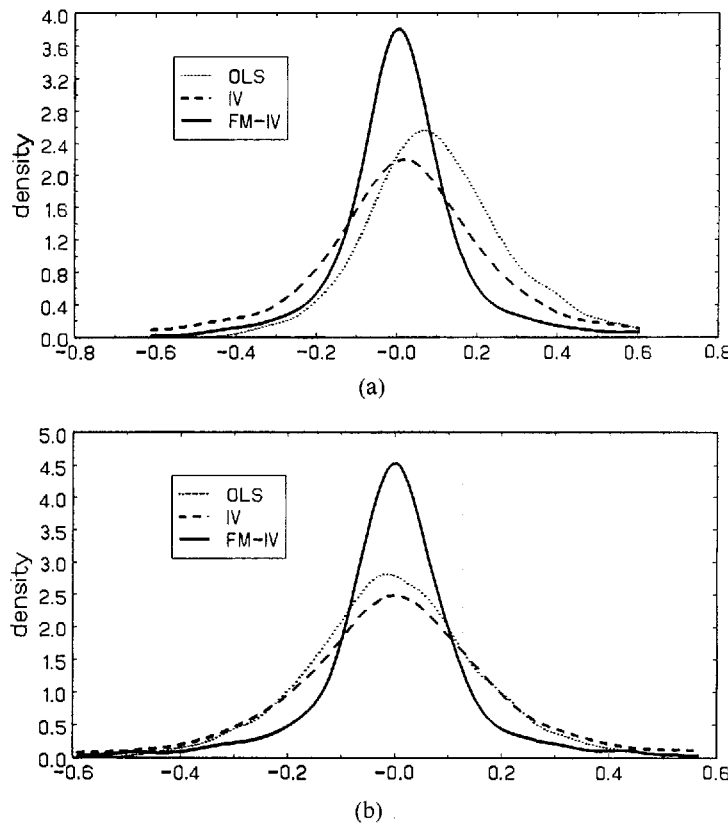
FIGURE 3. $I(1)/I(1)$ model. (a) Coefficient b_1, regressor $I(1)$, $\alpha = 0.7$, $\beta = 0.8$. (b) Coefficient b_2, regressor $I(1)$, $\alpha = 0.7$, $\beta = 0.8$.

well and is relatively stable throughout all the values for α. As seen in Figures 2b and 3a and b, which display the sampling distributions of the estimators for the $I(1)$ coefficients when $\alpha = 0.7$, FM-IV is evidently more concentrated than the others and unbiased. Interestingly, by comparing Figures 1b and 2b, we note that the shape of the estimated p.d.f. of FM-IV changes very little compared with OLS and crude IV. The concentration probabilities in these cases are reported in Table 5. When the feedback is as high as 0.7, FM-IV almost dominates crude IV even for the *stationary* components of the model. The preceding results suggest that, when the feedback is rather high, the use of FM-IV seems desirable. Note, however, as implied by the summary statistics displayed in Table 3, FM-IV yields occasional extreme deviations, as in the baseline case, so FM-IV tends to have a higher RMSE than crude IV.

TABLE 4. MAD, bias, and RMSE ($\beta = 0.8$, KLR $= 4$)

α	Model	Estimator	b_1 MAD	b_1 BIAS$_{ave}$	b_1 RMSE$_{ave}$	b_2 MAD	b_2 BIAS$_{ave}$	b_2 RMSE$_{ave}$
0.1	I(0)/I(0)	OLS	0.1557	0.0195	0.1949	0.0799	0.0012	0.1009
		IV	0.1626	-0.0085	0.2053	0.0827	-0.0004	0.1048
		FM-IV	0.2012	-0.0118	0.2530	0.0937	0.0000	0.1185
		FM-GMM	0.1863	-0.0116	0.2344	0.0926	0.0002	0.1176
		FM-GIVE	0.1289	-0.0175	0.1594	0.1048	-0.0311	0.1566
	I(0)/I(1)	OLS	0.1522	0.0185	0.1920	0.0263	0.0012	0.0372
		IV	0.1600	-0.0101	0.2028	0.0291	0.0020	0.0436
		FM-IV	0.1948	-0.0142	0.2459	0.0310	0.0029	0.0490
		FM-GMM	0.1876	-0.0147	0.2374	0.0313	0.0028	0.0492
		FM-GIVE	0.1245	-0.0061	0.1579	0.0293	0.0020	0.0454
	I(1)/I(1)	OLS	0.0937	0.0053	0.1234	0.0381	0.0012	0.0507
		IV	0.1020	-0.0004	0.1383	0.0405	0.0011	0.0559
		FM-IV	0.1152	-0.0018	0.1584	0.0448	0.0017	0.0620
		FM-GMM	0.1138	-0.0017	0.1560	0.0448	0.0014	0.0620
		FM-GIVE	0.0925	-0.0023	0.1203	0.0387	0.0019	0.0529
0.3	I(0)/I(0)	OLS	0.2486	0.2312	0.2912	0.0846	0.1073	0.1078
		IV	0.1782	-0.0100	0.2264	0.0904	-0.0012	0.1154
		FM-IV	0.2176	-0.0179	0.2772	0.1020	-0.0014	0.1303
		FM-GMM	0.2025	-0.0142	0.2581	0.1012	-0.0019	0.1294
		FM-GIVE	0.1362	-0.0054	0.1728	0.0861	-0.0026	0.1100

I(0)/I(1)	OLS	0.2578	0.2459	0.2972	0.0266	-0.0000	0.0375
	IV	0.1754	-0.0064	0.2280	0.0381	0.0014	0.1060
	FM-IV	0.2076	-0.0095	0.2802	0.0414	0.0028	0.1647
	FM-GMM	0.2013	-0.0091	0.2701	0.0418	0.0028	0.1562
	FM-GIVE	0.1374	-0.0036	0.1779	0.0375	0.0009	0.0834
I(1)/I(1)	OLS	0.0919	0.0484	0.1232	0.0434	-0.0004	0.0579
	IV	0.1237	0.0117	0.2447	0.0530	0.0025	0.0841
	FM-IV	0.1307	-0.0003	0.2473	0.0555	0.0021	0.0862
	FM-GMM	0.1387	-0.0083	0.5587	0.0570	0.0003	0.1342
	FM-GIVE	0.1028	-0.0011	0.1432	0.0485	0.0002	0.0785
0.7 I(0)/I(0)	OLS	0.5757	0.5755	0.6055	0.1446	0.0430	0.1834
	IV	0.4122	0.1818	0.6348	0.2176	0.0382	0.3514
	FM-IV	0.4334	0.1926	1.1716	0.2168	0.0137	0.4547
	FM-GMM	0.4373	0.1583	1.1344	0.2192	0.0242	0.4538
	FM-GIVE	0.3570	0.1032	0.6132	0.2069	0.0341	0.4552
I(0)/I(1)	OLS	0.6673	0.6673	0.6742	0.0371	-0.0357	0.0521
	IV	0.4498	0.2074	0.9059	0.0662	-0.0065	0.1495
	FM-IV	0.6300	0.4099	4.7789	0.0621	-0.0060	0.3144
	FM-GMM	0.5875	0.2610	4.1692	0.0617	0.0127	0.2764
	FM-GIVE	0.4112	0.1188	1.0002	0.0918	0.0060	0.3243
I(1)/I(1)	OLS	0.1575	0.1001	0.2080	0.1260	-0.0009	0.1672
	IV	0.1996	0.0298	0.3494	0.1678	0.0011	0.2686
	FM-IV	0.1378	0.0134	0.3891	0.1116	0.0041	0.2605
	FM-GMM	0.1489	0.0203	0.5706	0.1180	0.0026	0.3568
	FM-GIVE	0.1697	0.0120	0.4257	0.1406	-0.0016	0.2809

TABLE 5. Concentration probabilities: (I(0)/I(1) model, $\beta = 0.8$, KLR = 4)

			$P(|\bar{b} - b| \leq c)$				
α		Estimator	(b_1) $c = 0.1$ $c = 0.01$	$c = 0.3$ $c = 0.03$	$c = 0.5$ $c = 0.05$	$c = 0.7$ $c = 0.07$	$c = 0.9$ $c = 0.09$
0.1	b_1	OLS	0.4035	0.8875	0.9895	0.9990	1.0000
		IV	0.3935	0.8615	0.9860	0.9980	0.9995
		FM-IV	0.3265	0.7795	0.9565	0.9940	0.9985
		FM-GMM	0.3425	0.8005	0.9635	0.9950	0.9990
		FM-GIVE	0.4890	0.9370	0.9970	1.0000	1.0000
		$P(|b_{IV} - b| \leq c)/$ $P(|b_{FM\text{-}IV} - b| \leq c)$	1.2052	1.1052	1.0308	1.0040	1.0010
	b_2	OLS	0.3005	0.6845	0.8565	0.9350	0.9695
		IV	0.2895	0.6595	0.8290	0.9175	0.9650
		FM-IV	0.2825	0.6480	0.8240	0.9020	0.9435
		FM-GMM	0.2840	0.6370	0.8190	0.9010	0.9410
		FM-GIVE	0.2920	0.6630	0.8300	0.9140	0.9555
		$P(|b_{FM\text{-}IV} - b| \leq c)/$ $P(|b_{IV} - b| \leq c)$	0.9758	0.9826	0.9940	0.9831	0.9869
0.3	b_1	OLS	0.1660	0.6210	0.9440	0.9965	1.0000
		IV	0.3510	0.8385	0.9710	0.9925	0.9975
		FM-IV	0.3080	0.7750	0.9420	0.9855	0.9920
		FM-GMM	0.3165	0.7895	0.9475	0.9865	0.9935
		FM-GIVE	0.4555	0.9120	0.9915	0.9985	0.9995
		$P(|b_{IV} - b| \leq c)/$ $P(|b_{FM\text{-}IV} - b| \leq c)$	1.1396	1.0819	1.0308	1.0071	1.0055
	b_2	OLS	0.2940	0.6760	0.8515	0.9395	0.9730
		IV	0.2730	0.6220	0.7900	0.8780	0.9230
		FM-IV	0.2750	0.6215	0.8010	0.8805	0.9210
		FM-GMM	0.2655	0.6140	0.7935	0.8745	0.9210
		FM-GIVE	0.2890	0.6355	0.7975	0.8845	0.9230
		$P(|b_{FM\text{-}IV} - b| \leq c)/$ $P(|b_{IV} - b| \leq c)$	1.0073	0.9992	1.0139	1.0028	0.9978
0.7	b_1	OLS	0.0000	0.0000	0.0265	0.6565	0.9830
		IV	0.1450	0.4640	0.7310	0.8690	0.9260
		FM-IV	0.1845	0.5250	0.7550	0.8680	0.9215
		FM-GMM	0.1940	0.5240	0.7575	0.8700	0.9185
		FM-GIVE	0.2640	0.5745	0.7105	0.7810	0.8305
		$P(|b_{IV} - b| \leq c)/$ $P(|b_{FM\text{-}IV} - b| \leq c)$	0.7859	0.8838	0.9682	1.0012	1.0049
	b_2	OLS	0.1880	0.5535	0.7550	0.8645	0.9230
		IV	0.2215	0.4980	0.6530	0.7540	0.8105
		FM-IV	0.3195	0.6440	0.7860	0.8505	0.8850
		FM-GMM	0.3135	0.6410	0.7840	0.8470	0.8870
		FM-GIVE	0.2640	0.5745	0.7105	0.7810	0.8305
		$P(|b_{FM\text{-}IV} - b| \leq c)/$ $P(|b_{IV} - b| \leq c)$	1.4424	1.2932	1.2037	1.1280	1.0919

Table 6 reports the quantiles of the estimators when the relevance of the instruments is low, or the instruments are "poor." We set $\alpha = 0.5$ and KLR $= 4$ as in the baseline case, whereas the values for the relevance parameter β were chosen from the range $[0.0, 0.6]$. Notice that even if $\beta = 1$, that does not mean that the instruments are perfectly correlated with the regressors when $\alpha > 1$, and the correlation goes down as α goes up. On the other hand, when $\beta = 0$, the I(0) instruments are not valid, while the I(1) instruments are still valid due to the spurious correlation between nonstationary processes (Phillips and Hansen, 1990). Therefore, for $\beta = 0$, only the results for the I(1)/I(1) models are reported. As β becomes smaller, the sampling distributions of crude IV and FM-IV for the I(0) coefficients become more dispersed but considerably less biased than OLS. Also note that the desirability of FM-IV in the estimation of the I(1) coefficients remains even when β is low. In Figures 4b and 5a and b, again FM-IV outperforms crude IV when
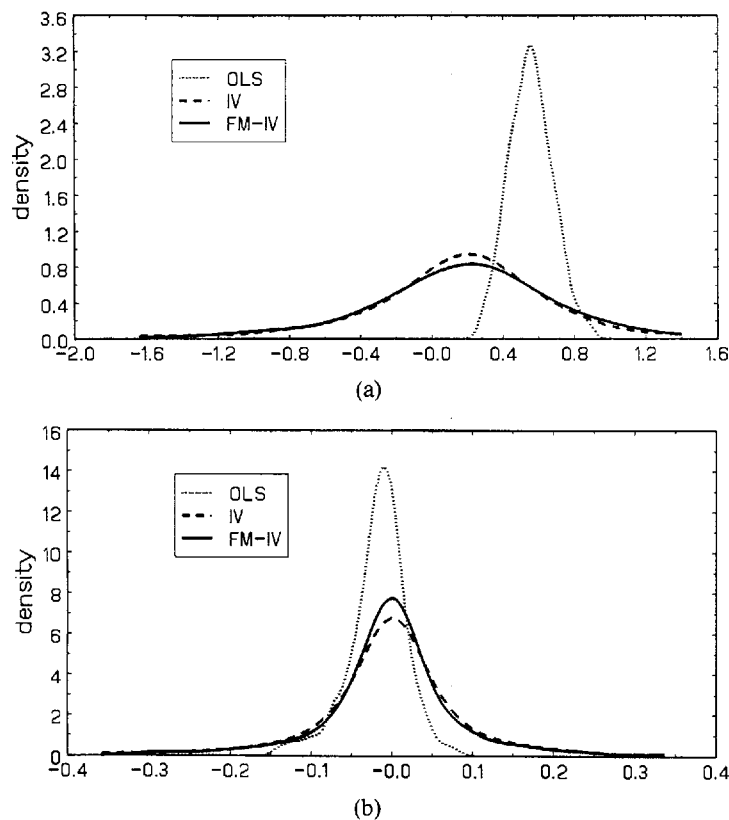

FIGURE 4. I(0)/I(1) model. (a) Coefficient b_1, regressor I(0), $\alpha = 0.5$, $\beta = 0.4$. (b) Coefficient b_2, regressor I(1), $\alpha = 0.5$, $\beta = 0.4$.

TABLE 6. Quantiles ($\alpha = 0.5$, KLR $= 4$)

β	Model	Estimator	b_1			b_2		
			$Q_{.25}$	$Q_{.50}$	$Q_{.75}$	$Q_{.25}$	$Q_{.50}$	$Q_{.75}$
0.0	I(1)/I(1)	OLS	0.0117	0.0648	0.1383	-0.0455	-0.0048	0.0341
		IV	-0.1038	0.0211	0.1465	-0.0781	-0.0012	0.0766
		FM-IV	-0.0813	0.0064	0.1058	-0.0647	-0.0003	0.0592
		FM-GMM	-0.0793	0.0091	0.1087	-0.0669	-0.0012	0.0559
		FM-GIVE	-0.1031	0.0107	0.1647	-0.0837	0.0008	0.0875
0.4	I(0)/I(0)	OLS	0.3835	0.4949	0.6093	-0.0596	0.0358	0.1209
		IV	-0.1506	0.1760	0.4659	-0.1569	0.0315	0.2207
		FM-IV	-0.1595	0.1992	0.5014	-0.1734	0.0167	0.2217
		FM-GMM	-0.1698	0.1819	0.4616	-0.1648	0.0306	0.2320
		FM-GIVE	-0.1788	0.1096	0.3646	-0.1279	0.0428	0.2044
	I(0)/I(1)	OLS	0.4686	0.5529	0.6352	-0.0343	-0.0132	0.0044
		IV	-0.0994	0.1899	0.4563	-0.0373	0.0004	0.0371
		FM-IV	-0.1390	0.1945	0.4789	-0.0340	0.0004	0.0325
		FM-GMM	-0.1487	0.1808	0.4698	-0.0327	0.0001	0.0318
		FM-GIVE	-0.1683	0.1162	0.3658	-0.0289	0.0035	0.0408

|							
I(1)/I(1)		OLS	0.0088	0.0741	0.1586	-0.0630	-0.0030	0.0514
		IV	-0.0926	0.0234	0.1483	-0.0912	0.0024	0.0871
		FM-IV	-0.0914	0.0044	0.1022	-0.0695	0.0006	0.0679
		FM-GMM	-0.0874	0.0058	0.1010	-0.0693	0.0024	0.0675
		FM-GIVE	-0.1119	0.0089	0.1294	-0.0824	-0.0017	0.0810
0.6	I(0)/I(0)	OLS	0.3840	0.4972	0.6139	-0.0601	0.0333	0.1197
		IV	-0.1697	0.0786	0.2985	-0.1156	0.0111	0.1391
		FM-IV	-0.1901	0.1067	0.3316	-0.1302	0.0084	0.1448
		FM-GMM	-0.1877	0.0945	0.3204	-0.1231	0.0132	0.1454
		FM-GIVE	-0.1514	0.0546	0.2287	-0.1033	0.0198	0.1276
	I(0)/I(1)	OLS	0.4682	0.5504	0.6361	-0.0336	-0.0133	0.0053
		IV	-0.1422	0.1068	0.3071	-0.0285	0.0034	0.0361
		FM-IV	-0.1594	0.0954	0.3175	-0.0246	0.0014	0.0308
		FM-GMM	-0.1632	0.0996	0.3119	-0.0241	0.0018	0.0303
		FM-GIVE	-0.1433	0.0583	0.2359	-0.0241	0.0033	0.0311
	I(1)/I(1)	OLS	0.0095	0.0744	0.1546	-0.0595	-0.0034	0.0525
		IV	-0.0763	0.0215	0.1219	-0.0731	0.0023	0.0741
		FM-IV	-0.0712	0.0093	0.0904	-0.0604	-0.0019	0.0585
		FM-GMM	-0.0714	0.0080	0.0889	-0.0585	-0.0004	0.0587
		FM-GIVE	-0.0813	0.0108	0.1067	-0.0601	0.0005	0.0649
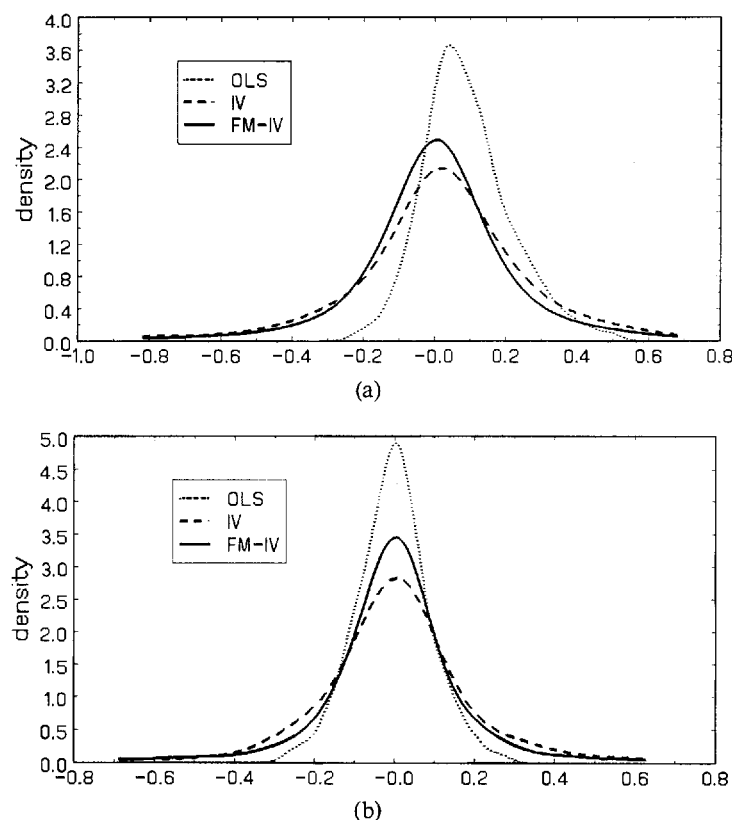
FIGURE 5. $I(1)/I(1)$ model. (a) Coefficient b_1, regressor $I(1)$, $\alpha = 0.5$, $\beta = 0.4$. (b) Coefficient b_2, regressor $I(1)$, $\alpha = 0.5$, $\beta = 0.4$.

the relevance is low and OLS beats FM-IV only occasionally. Taking the extreme case $\beta = 0$, where spurious $I(1)$ instruments are used, we find that the interquantile range $(Q_{.75}-Q_{.25})$ of FM-IV is evidently smaller than crude IV. However, not surprisingly, the dispersion of FM-IV (in terms of inter-quantile ranges, MAD and RMSE) is substantially higher than that of OLS when the instruments are spurious.

Next we turn to the efficiency issue among FM-IV, FM-GMM, and FM-GIVE in finite samples. According to our theory, the three FM estimators are asymptotically equivalent when applied to the $I(1)$ coefficients. Also, under certain conditions, FM-GIVE, FM-GMM, and FM-IV are asymptotically efficient in that descending order for the $I(0)$ coefficients, and this efficiency ordering applies under our DGP.
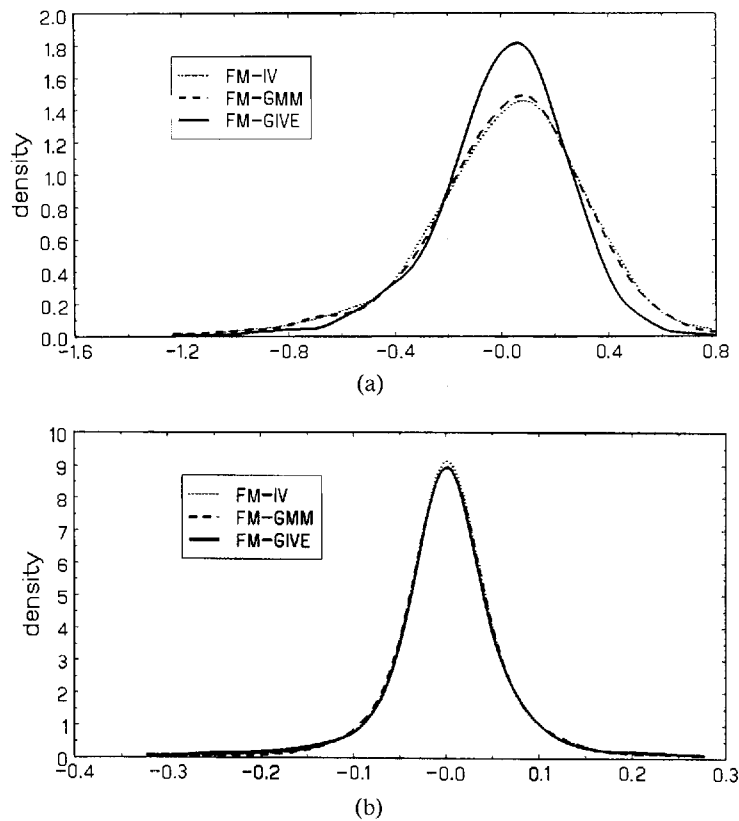
FIGURE 6. I(0)/I(1) model. (a) Coefficient b_1, regressor I(0), $\alpha = 0.5$, $\beta = 0.8$. (b) Coefficient b_2, regressor I(1), $\alpha = 0.5$, $\beta = 0.8$.

First, in returning to Table 1, we see little efficiency gain of FM-GMM over FM-IV with respect to the I(0) coefficients. Although sometimes we find evidence of a slight efficiency gain (e.g., as seen in Figure 7a), generally we do not observe any clear evidence of an efficiency gain throughout all the simulation results reported in Tables 1-8. This observation is confirmed in part (a) of Figures 6-10. One may suspect that this is due to the estimation error in the metric matrix S_{zT}. For instance, the poor initial estimates of the residuals that are used to form the matrix might cause such a problem. To check whether this explains the preceding result, the "pure" FM-GMM procedure, in which the true metric matrix S_{zT} was used in place of the estimate S_z, was applied to the I(0)/I(0) models. Although the results are not reported here, even in this case no clear evidence of a substantial efficiency gain was found.
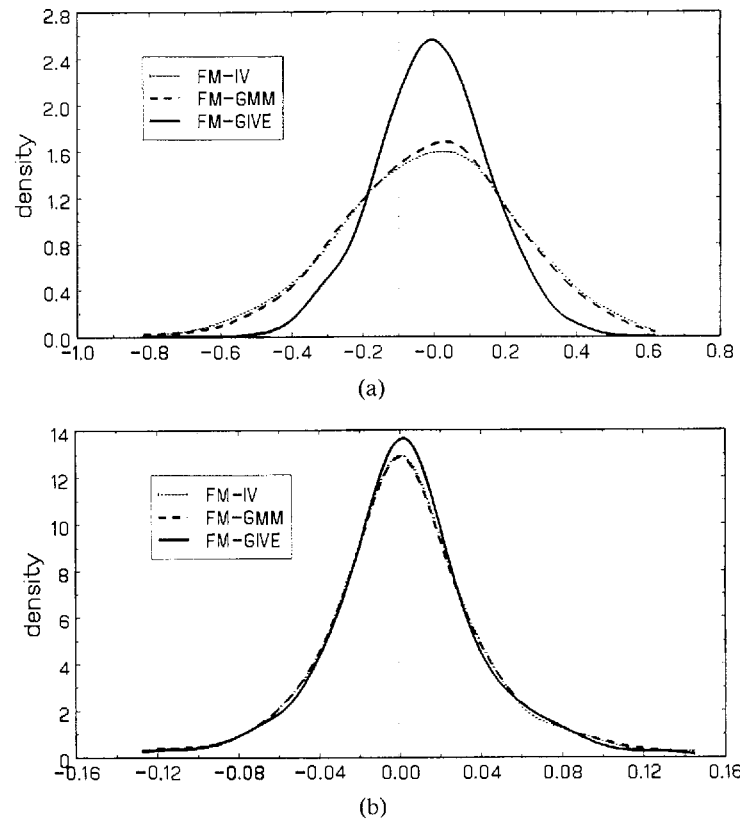
FIGURE 7. $I(0)/I(1)$ model. (a) Coefficient b_1, regressor $I(0)$, $\alpha = 0.1$, $\beta = 0.8$. (b) Coefficient b_2, regressor $I(1)$, $\alpha = 0.1$, $\beta = 0.8$.

On the other hand, Tables 1 and 2 also indicate that FM-GIVE clearly outperforms FM-GMM and FM-IV with respect to the estimation of the $I(0)$ coefficients. The results with various α reported in Tables 3-5 provide further evidence in favor of FM-GIVE. The reader should recall that the DGP for the regression errors is MA(1); therefore, the parameterization of our simulation is not directly amenable to the FM-GIVE procedure, in which AR models are utilized. Interestingly, Figures 7a, 8a, and 9a show that the relative performance of FM-GIVE is better when feedback is lower, although in all cases FM-GIVE outperforms FM-GMM and FM-IV. This is not surprising, because the efficiency gain of FM-GIVE depends on the initial stage estimation residuals, which are more accurate when the feedback is lower. On the other hand, in Figures 7b and 8b, the p.d.f.'s are quite close for the $I(1)$ coefficients, as theory predicts.
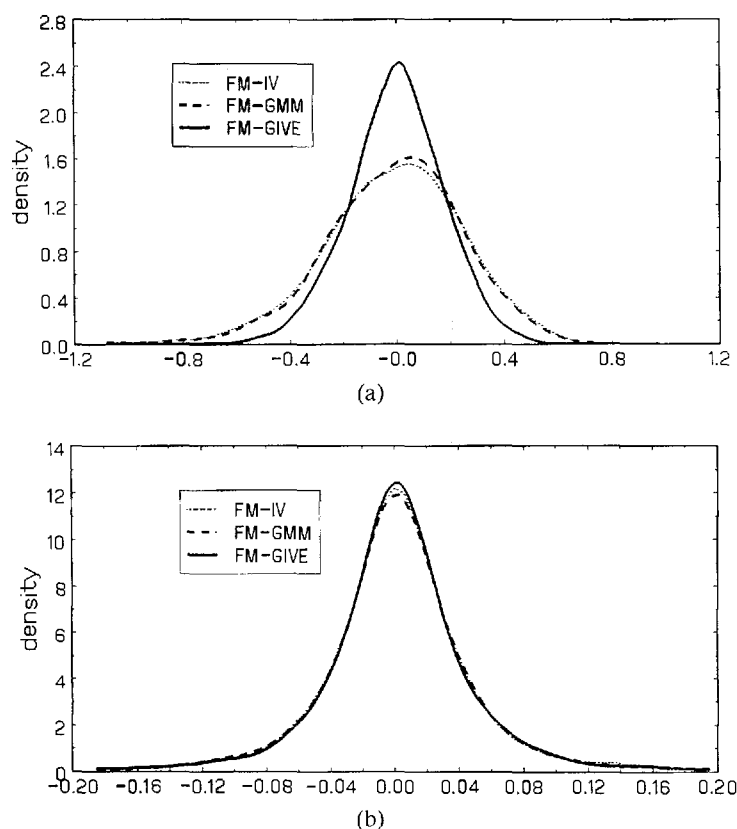
FIGURE 8. $I(0)/I(1)$ model. (a) Coefficient b_1, regressor $I(0)$, $\alpha = 0.3$, $\beta = 0.8$. (b) Coefficient b_2, regressor $I(1)$, $\alpha = 0.3$, $\beta = 0.8$.

Turning to Table 6, the good performance of FM-GIVE with respect to the $I(0)$ coefficients seems to remain for various β. In fact, as Figure 10a shows, the relative efficiency of FM-GIVE continues when the relevance of instruments is rather low. The efficiency gain becomes a little smaller as β is reduced, though, and again this is caused by the poor residual estimates obtained by the initial IV regression with poor (i.e., low-relevance) instruments.

The very limited efficiency gain obtained by the use of FM-GMM might be partly due to the fact that "nearly" nonstationary processes were used in the experiments. Consider the $I(0)/I(0)$ model, where the two roots used to generate the instrumental variables are 0.8. Given the exogeneity of $\{z_t\}$, S_z is the sum of $E(u_t u_{t+j}) \cdot E(z_t z'_{t+j})$ over $-M$ to M, say. Because the roots in the instruments are not unity but still rather large, the autocovariance $E(z_t z'_{t+j})$ is quite smooth in j. In consequence, S_t is close to $E(z_t z'_t)$ (times a scale fac-
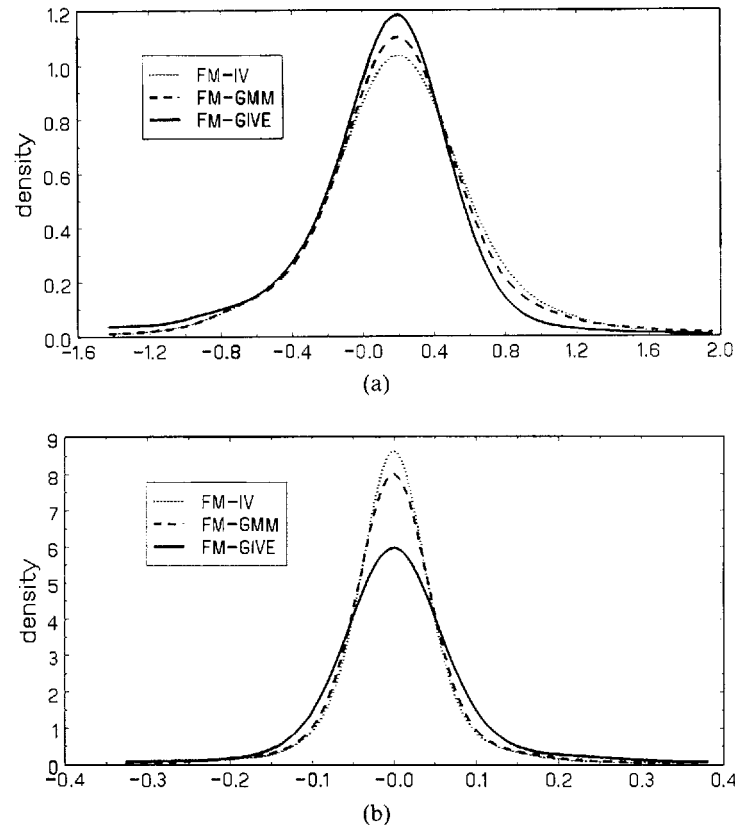
FIGURE 9. I(0)/I(1) model. (a) Coefficient b_1, regressor I(0), $\alpha = 0.7$, $\beta = 0.8$. (b) Coefficient b_2, regressor I(1), $\alpha = 0.7$, $\beta = 0.8$.

tor) and, thus, FM-IV and FM-GMM are numerically close. Note that when we have unit roots, the preceding argument shows the asymptotic equivalence of FM-IV, FM-GMM, and FM-GIVE. This is a direct consequence of the asymptotic OLS/GLS equivalence in nonstationary models that was shown by Phillips and Park (1988).

In fact, as our limit theory predicts, FM-IV, FM-GMM, and FM-GIVE are numerically very close when applied to the I(1) coefficients. This is generally true for almost all the results reported in Tables 1–8, though FM-GIVE seems relatively sensitive to the change of parameterizations; namely, it tends to work better for lower α and higher β. With strong feedback or when relevance is weak, FM-GIVE shows more dispersion than the other two FM estimators for the I(1) coefficient (see Figures 9b and 10b). As in the I(0)
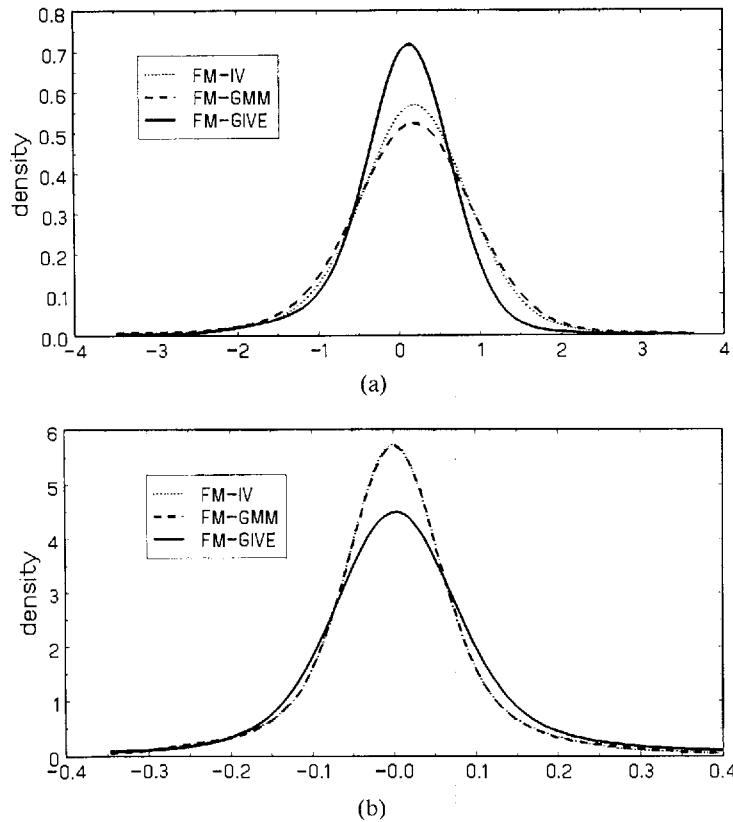
FIGURE 10. I(0)/I(1) model. (a) Coefficient b_1, regressor I(0), $\alpha = 0.5$, $\beta = 0.4$. (b) Coefficient b_2, regressor I(1), $\alpha = 0.5$, $\beta = 0.4$.

case, the latter characteristics can be attributed to the fact that the performance of the initial IV estimation is rather sensitive to the parameters (α, β). Therefore, it may be useful to add more iterative steps to the FM-GIVE procedure described in Section 3.2; this option, however, was not pursued in the present paper.

Finally, Table 8 gives the results with various values for KLR and $(\alpha, \beta) = (0.5, 0.8)$. In fact, the results for each estimator varied little for KLR $= 2, 4, 6, 8$, and 10. Thus, it seems difficult to find some decision rule for the choice of KLR based on the Monte Carlo results reported here. In other words, the estimation results seem rather robust to the choice of KLR, at least when the sample size is around 100. Thus, it seems that any moderate KLR may be used in practice.

TABLE 7. MAD, bias, and RMSE ($\alpha = 0.5$, KLR $= 4$)

β	Model	Estimator	b_1 MAD	b_1 BIAS$_{ave}$	b_1 RMSE$_{ave}$	b_2 MAD	b_2 BIAS$_{ave}$	b_2 RMSE$_{ave}$
0.0	I(0)/I(0)	OLS	0.0992	0.0800	0.1323	0.0525	−0.0048	0.0707
		IV	0.2644	0.0281	0.5724	0.1732	−0.0073	0.3934
		FM-IV	0.2288	0.0245	0.6814	0.1539	−0.0117	0.5441
		FM-GMM	0.2235	0.0240	0.7872	0.1497	−0.0057	0.5793
		FM-GIVE	0.5293	−0.0320	2.2085	0.4165	0.0203	2.5819
0.4	I(0)/I(0)	OLS	0.4968	0.4967	0.5253	0.1106	0.0347	0.1394
		IV	0.4647	0.1249	0.6853	0.2678	0.0319	0.4152
		FM-IV	0.5377	0.1234	0.9549	0.3105	0.0140	0.5640
		FM-GMM	0.5321	0.0827	1.0946	0.3254	0.0039	1.1684
		FM-GIVE	0.4576	0.0756	1.7840	0.2909	0.0303	1.0619
	I(0)/I(1)	OLS	0.5553	0.5553	0.5698	0.0296	−0.0170	0.0411
		IV	0.5344	0.1734	1.3991	0.0883	−0.0076	0.2912
		FM-IV	0.8607	−0.1028	7.9070	0.1211	0.0382	1.5363
		FM-GMM	0.8277	−0.0289	6.4698	0.1124	0.0304	1.1693
		FM-GIVE	0.4806	0.1108	1.5689	0.1249	0.0116	0.5673

I(1)/I(1)	OLS	0.1190	0.0909	0.1591	0.0732	-0.0050	0.0973
	IV	0.2262	0.0231	0.4274	0.1673	0.0058	0.3726
	FM-IV	0.2550	0.0061	1.2967	0.2261	0.0884	3.1868
	FM-GMM	0.2579	0.0112	1.3450	0.2534	0.1176	4.3932
	FM-GIVE	0.2145	0.0046	0.6672	0.1582	-0.0045	0.4253
0.6	I(0)/I(0)	OLS	0.4967	0.4966	0.5250	0.1109	0.0353	0.1409
		IV	0.3227	0.0334	0.4526	0.1738	0.0043	0.2885
		FM-IV	0.3980	0.0089	0.7777	0.2076	-0.0227	0.5320
		FM-GMM	0.3819	0.0063	0.7716	0.2053	-0.0183	0.5325
		FM-GIVE	0.2623	0.0112	0.4249	0.1635	-0.0085	0.3724
	I(0)/I(1)	OLS	0.5543	0.5543	0.5685	0.0302	-0.0159	0.0419
		IV	0.3350	0.0545	0.5019	0.0719	0.0078	0.2035
		FM-IV	0.3808	0.0472	0.7258	0.0764	0.0261	0.5910
		FM-GMM	0.3772	0.0315	0.7108	0.0765	0.0238	0.5646
		FM-GIVE	0.2611	0.0187	0.3783	0.0777	0.0067	0.1957
	I(1)/I(1)	OLS	0.1179	0.0900	0.1574	0.0737	-0.0025	0.0984
		IV	0.1793	0.0298	0.3300	0.1211	0.0023	0.1963
		FM-IV	0.1956	0.0062	1.1380	0.1145	-0.0046	0.2558
		FM-GMM	0.1982	0.0096	1.1775	0.1147	-0.0030	0.2613
		FM-GIVE	0.1453	0.0094	0.2431	0.1068	0.0047	0.2262

TABLE 8. Concentration probabilities: (I(0)/I(1) model, $\alpha = 0.5$, KLR = 4)

β		Estimator	(b_1) (b_2)	$P(\lvert \bar{b} - b \rvert \leq c)$				
				$c = 0.1$ $c = 0.01$	$c = 0.3$ $c = 0.03$	$c = 0.5$ $c = 0.05$	$c = 0.7$ $c = 0.07$	$c = 0.9$ $c = 0.09$
0.4	b_1	OLS		0.0000	0.0185	0.3345	0.8740	0.9935
		IV		0.1600	0.4550	0.6795	0.8055	0.8860
		FM-IV		0.1575	0.4380	0.6595	0.7810	0.8560
		FM-GMM		0.1605	0.4445	0.6580	0.7865	0.8580
		FM-GIVE		0.1840	0.5125	0.7255	0.8485	0.9070
		$P(\lvert b_{IV} - b \rvert \leq c)/$ $P(\lvert b_{FM\text{-}IV} - b \rvert \leq c)$		1.0159	1.0388	1.0462	1.0314	1.0350
	b_2	OLS		0.2590	0.6375	0.8275	0.9080	0.9530
		IV		0.1720	0.4380	0.5975	0.7000	0.7690
		FM-IV		0.1955	0.4710	0.6410	0.7325	0.7920
		FM-GMM		0.1955	0.4750	0.6320	0.7330	0.7885
		FM-GIVE		0.1955	0.4540	0.5970	0.7000	0.7570
		$P(\lvert b_{FM\text{-}IV} - b \rvert \leq c)/$ $P(\lvert b_{IV} - b \rvert \leq c)$		1.1366	1.0753	1.0728	1.0464	1.0299
0.6	b_1	OLS		0.0000	0.0180	0.3410	0.8710	0.9945
		IV		0.1980	0.5895	0.8125	0.9135	0.9475
		FM-IV		0.2095	0.5760	0.7915	0.8930	0.9320
		FM-GMM		0.1980	0.5825	0.7910	0.8990	0.9350
		FM-GIVE		0.2700	0.6755	0.8820	0.9505	0.9820
		$P(\lvert b_{IV} - b \rvert \leq c)/$ $P(\lvert b_{FM\text{-}IV} - b \rvert \leq c)$		0.9451	1.0234	1.0265	1.0230	1.0166
	b_2	OLS		0.2525	0.6320	0.8140	0.9040	0.9530
		IV		0.1920	0.4765	0.6320	0.7340	0.8005
		FM-IV		0.2250	0.5215	0.6860	0.7745	0.8325
		FM-GMM		0.2265	0.5255	0.6830	0.7790	0.8365
		FM-GIVE		0.2135	0.5220	0.6590	0.7360	0.8010
		$P(\lvert b_{FM\text{-}IV} - b \rvert \leq c)/$ $P(\lvert b_{IV} - b \rvert \leq c)$		1.1719	1.0944	1.0854	1.0552	1.0400

4. CONCLUSIONS

The purpose of this paper was to investigate the practical implication of the theory of FM-IV estimation with possibly nonstationary processes as developed by Kitamura and Phillips (1992). Although the DGP was rather simple, it still has some realistic features such as feedback from the regression errors to the regressors and serial dependence in the regression errors. The

two AR roots in the regressors and the instruments were set to unity or 0.8 or both, and that knowledge (about the number and location of the unit roots) was not used in the implementation of each estimation procedure. In addition to the preceding three FM estimators, the performance of OLS and crude IV was also examined for comparison. The main tasks in the simulation exercise were the assessment of the effectiveness of the FM procedures and the efficiency comparison among FM-IV, FM-GMM, and FM-GIVE.

First, the quantiles, the concentration probabilities, and the p.d.f. graphics seem to imply that among the OLS, crude IV, and FM-IV estimators, FM-IV is the most desirable choice: when applied to the I(0) coefficients, OLS is substantially biased due to the feedback; when applied to the I(1) coefficients, OLS is still biased and crude IV is much more dispersed than FM-IV. For higher feedback, the relative advantage of the FM-IV over OLS and crude IV becomes more evident. On the other hand, with poor (low-relevance) instruments, crude IV and FM-IV applied to the I(0) coefficients tend to have dispersed sampling distributions, although for the I(1) coefficients FM-IV shows very small bias and beats at least the crude IV estimators. Also, it should be noted that the effect of occasional outliers in the first-stage IV regression of the feasible FM-IV can be exaggerated in the second-stage outcome. For example, when the quality of the instruments is extremely poor, care should be taken in the use of the FM-IV procedures. As a result of these occasional extreme estimation errors, FM-IV has a higher RMSE than crude IV consistently. Thus, it seems to be important to investigate some modifications of FM-IV to adjust for this tail behavior. This could be achieved by a form of pretest estimator or decision rule, as in some reduced form estimators in simultaneous equation models. This problem is left to future research.

Second, some direct comparisons among the FM-IV, FM-GMM, and FM-GIVE were made in the simulations. When applied to the I(0) coefficients, the relative efficiency of the FM-GIVE procedure over FM-IV and FM-GMM was found to be substantial under various parameterizations. The dominance of FM-GIVE over FM-GMM is particularly important and seems likely to be relevant in many applications. At the same time, the performance of the first-stage regression in the feasible procedures also seems to be important in the performance of FM-GIVE. This suggests that iterative procedures may be useful. Also, with respect to the estimation of I(0) coefficients, the efficiency gain of FM-GMM over FM-IV observed in our Monte Carlo simulations is very limited. This characteristic is persistent for almost all the simulations. Although it would be unwise to make strong claims from the results based on the rather simple DGP used here, the use of FM-GMM may not yield as much of an efficiency gain as one might expect, especially when the DGP has stationary but rather large AR roots. When applied to the I(1) coefficients, the three estimators are generally very close even for the rather small sample size of 100.

Because we focused on estimation, the finite sample performance of statistical tests based on FM procedures was not analyzed in the paper. In fact, in Kitamura and Phillips (1992), "fully modified" instrumental validity tests were developed, and they are shown to be chi-square–distributed asymptotically. Although we need to conduct more simulation studies (e.g., on the choice of the kernels and bandwidth/truncation parameter used in the test statistics), at the present point our preliminary simulation results suggest that these statistics are biased toward rejection of the overidentifying restrictions. Therefore, in addition to more extensive sampling experiments, further theoretical research seems to be necessary on this issue and is currently being investigated by the authors.

NOTES

1. For the definition of these kernels, see, for example, Priestley (1981), Andrews (1991), or Kitamura and Phillips (1992). Note that all these kernels have the same characteristic exponent (Parzen, 1957). The Tukey–Hanning kernel estimator of a two-sided long-run variance matrix may not be positive semidefinite and, thus, is less desirable in this aspect.

2. In fact, Assumption LR(b) rules out the optimal rate $T^{1/5}$ in Andrews (1991).

REFERENCES

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

Hansen, B.E. & P.C.B. Phillips (1990) Estimation and inference in models of cointegration: A simulation study. *Advances in Econometrics* 8, 225–248.

Hansen, L.P. (1982) Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029–1054.

Kitamura, Y. & P.C.B. Phillips (1992) Fully Modified IV, GIVE, and GMM Estimation with Possibly Non-stationary Regressors and Instruments. Mimeo, Yale University.

Nelson, C.R. & R. Startz (1990) The distribution of the instrumental variables estimator and its t-ratio when the instrument is a poor one. *Journal of Business* 63, 125–140.

Park, J.Y. & P.C.B. Phillips (1989) Statistical inference in regressions with integrated processes: Part 2. *Econometric Theory* 5, 95–131.

Parzen, E. (1957) On consistent estimates of the spectrum of a stationary time series. *Annals of Mathematical Statistics* 28, 329–348.

Phillips, P.C.B. (1989) Partially identified econometric models. *Econometric Theory* 5, 181–240.

Phillips, P.C.B. (1991) Optimal inference in cointegrated systems. *Econometrica* 59, 283–306.

Phillips, P.C.B. & B.E. Hansen (1990) Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57, 99–125.

Phillips, P.C.B. & J.Y. Park (1988) Asymptotic equivalence of ordinary least squares and generalized least squares in regressions with integrated regressors. *Journal of the American Statistical Association* 83, 111–115.

Priestley, M.B. (1981) *Spectral Analysis and Time Series*. London: Academic Press.

Sargan, J.D. (1988) *Lectures on Advanced Econometrics*. New York: Basil Blackwell.

White H. (1984) *Asymptotic Theory for Econometricians*. New York: Academic Press.