8-2002

A Misspecification-Robust Impulse Response Estimator

Pao Li CHANG
Singapore Management University, plchang@smu.edu.sg

Shinichi SAKATA
University of Michigan - Ann Arbor

SMU ECONOMICS & STATISTICS
WORKING PAPER SERIES

A Misspecification-robust Impulse Response Estimator

Pao-Li Chang
Shinichi Sakata

August 2002

Paper No. 15-2002

A Misspecification-Robust Impulse Response Estimator

Pao-Li Chang

School of Economics and Social Sciences, Singapore Management University, Singapore 259756

Shinichi Sakata

Department of Economics, University of Michigan, Ann Arbor, MI 48109

August 8, 2002

Abstract

Impulse response analysis is typically conducted by fitting an autoregression model to a time series and calculating the moving average coefficients implied by the estimated autoregression model. The possible shape and persistence of the impulse response function implied by a parsimonious autoregression specification are very limited. This paper proposes an alternative approach to estimating impulse response function, which is asymptotically valid yet is less sensitive to model misspecifications in small samples. The small sample advantages of the proposed impulse response estimator over the conventional approach is demonstrated by Monte Carlo studies. The large sample validity of the proposed estimator is also established.

KEY WORDS: Nonparametric; Persistence; Two-Stage Estimation.

1. INTRODUCTION

Impulse response analysis is widely used for studying the dynamics of economic time series. In this paper, we will be concerned with impulse response analysis of a univariate time series. Suggestions for how our methodology can be extended to the multivariate case is provided at the end of the article.

A popular example of the univariate impulse response analysis is found in the purchasing power parity (PPP) literature, in which economists are concerned about how fast the real exchange rate reverts to its mean after a shock. Estimated impulse responses are often used to assess the speed of mean reversion or the degree of persistence for the real exchange rate under study. The common practice in estimating the impulse response function (IRF) is to fit a pth order autoregression (AR) model to the series, and calculate the coefficients in the MA(∞) representation of the estimated AR(p) model. For example, Murray and Papell (2002) among others applies this conventional method to analyze the persistence of the real exchange rate.

Despite its popularity, the conventional approach is subject to a potential caveat. In small samples, the order of the autoregression specification used to model a process is typically low. The possible shape and persistence of the IRF delivered by a parsimonious autoregression specification are very limited (although this limitation will disappear asymptotically if the model complexity p is allowed to diverge with the sample size to infinity). In particular, with a low order AR model, the conventional approach tends to impose a smooth shape on the estimated IRF. This restriction excludes the possibility of richer dynamics and hence might render incorrect inferences about the persistence of a process. For example, suppose that an economic series reverts to its mean suddenly after some initial periods in response to a shock, such that its underlying IRF exhibits a jump to zeros from previous high levels. With a parsimonious model, the conventional approach will typically overestimate the persistence of such process and underestimate its speed of mean reversion.

In this paper, we propose an alternative impulse response estimator that is asymptotically valid yet less sensitive to model misspecifications in small samples. The basic idea is to regress the data Y_t on the estimated innovation at lag k, $\hat{\epsilon}_{t-k}$, to estimate the impulse response at horizon k. The estimated innovations can be obtained in a prior stage by fitting a pth order AR model to the data. Because of the way the proposed estimator is constructed, it does not impose smoothness on the shape of the estimated IRF, regardless of the parametric model used at the first stage. This makes it possible for the proposed estimator to detect some interesting features of the true IRF that is excluded by the conventional estimator. Our Monte Carlo simulations demonstrate that the proposed estimator is superior to the conventional estimator in small samples, when the AR specification is incorrect for the process. In particular, in the case where the DGP exhibits sudden mean reversion in its IRF, the proposed estimator is capable of detecting such nonsmooth dynamics while the conventional estimator fails to do so, even with as complicated a model as AR(12). On the other hand, in the ideal case where a finite-order AR specification happens to be correct for the DGP, the Monte Carlo study shows that the proposed estimator performs comparably well to the conventional estimator.

The asymptotic validity of our proposed impulse response estimator is established based on the notion that the error in approximating a potential infinite-dimensional parameter space by finite-dimension parameterization will vanish provided that the dimension of parametric model is allowed to expand slowly with the sample size at an appropriate rate. The general idea of successive approximation in estimating a potentially infinite-dimensional parameter space in the context of distributed lag estimation is discussed in Sims (1971, 1972).

It is also possible to estimate univariate impulse response function using the frequency domain approach. Bhansali (1976) proposed a technique to estimate the moving average representation of a stationary process using the estimated spectrum. This nonparametric approach was evaluated against the conventional approach by Wright (1999). Although this approach presents another alternative to impulse response estimation, it is difficult to extend the technique to the multivariate case as there is no closed form expression for the moving average matrices in terms of the multivariate spectrum (see Wright, 1999, for detailed discussions).

The rest of the paper is organized as follows. In Section 2, we review the conventional practice of impulse response estimation and describe our alternative methodology. In Section 3, the asymptotic consistency of the proposed IRF estimator is established. We then apply the proposed method to the French real exchange rate and compare the results with the conventional estimates in Section 4. In Section 5, Monte Carlo simulations are conducted to examine the performance of the conventional and the proposed IRF estimators in small samples. Final remarks are given in Section 6.

2. IMPULSE RESPONSE ESTIMATION

Let C be the complex plane and D the unit disk in C. In this paper, we consider processes that satisfy the following assumption:

Assumption 1 $\{Y_t\}_{t\in\mathbb{Z}}$ is a univariate, fourth-order stationary process on a probability space $(\Omega, \mathscr{F}, P)$, with mean μ such that the Wold decomposition of $\{Z_t \equiv Y_t - \mu\}_{t\in\mathbb{Z}}$ has no deterministic component. That is, $Z_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$, $t \in \mathbb{Z}$, where $\{\epsilon_t\}_{t\in\mathbb{Z}}$ is a zero-mean white noise process, and $\psi_0, \psi_1, \psi_2, \ldots$ are real constants with $\psi_0 = 1$. The sequence $\{\psi_j : j = 0, 1, 2, \ldots\}$ is absolutely summable and satisfies the condition that $\Psi : \mathsf{D} \to \mathsf{C}$ vanishes nowhere on D, where $\Psi(z) \equiv \sum_{j=0}^{\infty} \psi_j z^j$, $z \in \mathsf{D}$.

Note that the Wold decomposition theorem (e.g., Brockwell and Davis, 1991, Thm. 5.7.1, pp. 187–189) only guarantees $\{\psi_j\}$ to be square-summable. The absolute summability imposed on $\{\psi_j\}$ in Assumption 1 is a stronger condition. The MA(∞) representation of Y_t given by the Wold decomposition,

$$Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, \quad t \in \mathbb{Z}, \tag{1}$$

is invertible under Assumption 1, because Ψ vanishes nowhere on D. By the basic properties of analytic functions, there exists an absolutely summable sequence $\{\phi_j \in \mathbb{R}\}_{j \in \mathbb{N}}$ such that $\Psi(z)^{-1} = 1 - \sum_{j=1}^{\infty} \phi_j z^j$, for each $z \in \mathsf{D}$. Using this series $\{\phi_j\}$, we can write Y_t as

$$\begin{aligned} Y_t &= \mu + \sum_{j=1}^{\infty} \phi_j (Y_{t-j} - \mu) + \epsilon_t \\ &= \alpha + \sum_{j=1}^{\infty} \phi_j Y_{t-j} + \epsilon_t, \quad t \in \mathbb{Z}, \end{aligned} \tag{2}$$

where $\alpha \equiv \mu \left(1 - \sum_{j=1}^{\infty} \phi_j\right)$.

The impulse response function of the time series $\{Y_t\}$ at horizon k is ψ_k in equation (1). It can be interpreted as the marginal effect of a unit shock at time $t-k$ on Y_t. Finding the impulse response function of $\{Y_t\}$ in (1) requires fitting an infinite number of parameters $(\psi_1, \psi_2, \ldots)$ or $(\phi_1, \phi_2, \ldots)$ to the data. With a finite number of observations on $\{Y_t\}$, this is infeasible. Instead, an AR(p_n) model is usually used to approximate the process, where the finite lag order p_n is potentially dependent on the sample size n. The coefficients of the corresponding MA(∞) representation of the AR(p_n) model are the base for the conventional impulse response estimation. Usually Least Squares (LS) method is used to estimate the AR(p_n) model, and the conventional estimator of the impulse response of $\{Y_t\}$ at horizon k is the coefficient of the corresponding MA(∞) representation of the estimated AR(p_n) model at lag k. We denote the conventional impulse response estimator ψ_{nk}^{conv}.

The conventional approach, as discussed in the introduction, is sensitive to model misspecifications in small samples. We proposed an alternative methodology, with which the impulse response at horizon k is estimated by regressing the data Y_t on the estimated innovation at lag k, $\hat{\epsilon}_{t-k}$. The estimated innovations can be obtained in a prior stage by fitting an AR(p_n) model to the data. In spite that the AR(p_n) model fitted to the process $\{Y_t\}$ in the first stage is still likely to be misspecified, the proposed IRF estimator is less sensitive to such misspecification and is able to deliver more robust description about the shape and persistence of the true IRF. These small sample properties of the conventional and proposed impulse response estimators will be demonstrated with Monte Carlo simulations in Section 5.

The proof below establishes the asymptotic validity of the proposed impulse response estimator. It shows that the proposed estimator is consistent for the true impulse response function when the complexity p_n of the fitted AR model is allowed to grow slowly with the sample size at an appropriate rate.

3. NEW IMPULSE RESPONSE ESTIMATOR

We begin by defining some notations and making notes of some mathematical results that will be used repeatedly in the following proof. First, for each vector x, let $|x|$ denote the Euclidean norm of x and for each $m \times n$ matrix A, let $|A|$ denote the norm of the linear operator $x \mapsto Ax : \mathbb{R}^n \to \mathbb{R}^m$, i.e., $|A| \equiv \sup\{|Ax| : |x| = 1, x \in \mathbb{R}^n\}$. Note that if A is an arbitrary $m \times n$ matrix and x an arbitrary $n \times 1$ vector with length one, then the ith element of the vector Ax is no greater in magnitude than the product of the length of the ith row of A and the length of x by the Cauchy-Schwartz inequality. Thus, we have that $|Ax| \leq (\mathrm{tr}(AA'))^{1/2}$. It follows that $|A| \leq (\mathrm{tr}(AA'))^{1/2}$. Note also that $|A|$ is equal to the square root of the maximum eigen value of AA'. Therefore, when A is symmetric, $|A|$ is equal to the maximum of the absolute values of A's eigen values. Finally, for each pair of random variables V_1 and V_2 in $L_2(\Omega, \mathscr{F}, P)$, let $\langle V_1, V_2 \rangle \equiv \mathrm{E}[V_1 V_2]$. Also, let $\|\cdot\|$ denote the L_2-norm on $L_2(\Omega, \mathscr{F}, P)$.

Assumption 2 (a) $\sum_{\tau=-\infty}^{\infty} |\gamma(\tau)| < \infty$ where $\gamma(\tau) \equiv \mathrm{Cov}[Y_0, Y_\tau]$ for each $\tau \in \mathbb{Z}$.

(b) $\sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} \sum_{\tau_3=-\infty}^{\infty} |\kappa_4(0, \tau_1, \tau_2, \tau_3)| < \infty$ where $\kappa_4(t_1, t_2, t_3, t_4)$ denotes the fourth-order cumulants of $(Y_{t_1}, Y_{t_2}, Y_{t_3}, Y_{t_4})$ for each $(t_1, t_2, t_3, t_4) \in \mathbb{Z}^4$.

(c) The spectral density function $f : [-\pi, \pi] \to \mathbb{R}$ of $\{Y_t\}$ satisfies that $\bar{f}_1 \equiv \inf_{v \in [-\pi,\pi]} f(v) > 0$ and $\bar{f}_2 \equiv \sup_{v \in [-\pi,\pi]} f(v) < \infty$.

In the above assumption, the absolute summability of $\{\gamma(\tau)\}_{\tau \in \mathbb{Z}}$ implies the existence of the spectral density of $\{Y_t\}$ (see Brockwell and Davis, 1991, Theorem 4.3.2, p. 120), while the absolute summability of the fourth-order cumulants can be viewed as a restriction on the memory property of the process, as discussed in Andrews (1991, pp. 823–824). It is also worth noting that all eigen values of the covariance matrix of $(Y_1, Y_2, \ldots, Y_n)$ will fall between $2\pi\bar{f}_1$ and $2\pi\bar{f}_2$ (see Brockwell and Davis, 1991, Prop. 4.5.3, pp. 137–138).

In the following assumption, the complexity p_n of the AR model used in the first stage is allowed to grow with the sample size n but at a rate slower than $n^{1/2}$.

Assumption 3 $\{p_n\}_{n=2}^{\infty}$ is a sequence of natural numbers starting with $p_2 = 1$ and diverging to ∞ such that $\{p_n/n\}_{n=2}^{\infty}$ is a nonincreasing sequence, and $p_n^2/n \to 0$ as $n \to \infty$.

We now formulate the proposed impulse response estimator and prove its consistency for the true IRF. Consider estimation of the impulse response at lag k. For each $k = 0, 1, 2, \ldots$, let $\bar{n}_k$ denote the minimum natural number $n \geq 2$ such that $n - p_n - k > 0$. The first step is to fit the process $\{Y_t\}$ with an AR(p_n) model given the available data and obtain estimated residuals. Define $X_{nt} \equiv (1, Y_{t-1}, \ldots, Y_{t-p_n})'$ and $\hat{R}_n \equiv (n - p_n)^{-1} \sum_{t=p_n+1}^{n} X_{nt} X_{nt}'$, $t \in \mathbb{Z}$, $n = 2, 3, \ldots$. Then the OLS estimator of the AR(p_n) model is $\hat{\beta}_n \equiv \hat{R}_n^{+} (n - p_n)^{-1} \sum_{t=p_n+1}^{n} X_{nt} Y_t$, $n = 2, 3, \ldots$, where $\hat{R}_n^{+}$ denotes the Moore-Penrose inverse of $\hat{R}_n$ (see Magnus and Neudecker, 1988, pp. 32–39), and the estimated residuals are $\hat{\epsilon}_{nt} \equiv Y_t - X_{nt}'\hat{\beta}_n$, $t = p_n + 1, \ldots, n$, $n = 2, 3, \ldots$. The proposed estimator of the impulse response at horizon k is obtained by regressing Y_t on the estimated residual $\hat{\epsilon}_{t-k}$ with k horizon difference. That is,

$$\hat{\psi}_{nk} \equiv \begin{cases} \left((n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} \hat{\epsilon}_{n,t-k}^2 \right)^{+} & \\ \qquad \times (n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} \hat{\epsilon}_{n,t-k} Y_t & \text{if } k \leq n - p_n - 1, \\ 0 & \text{otherwise,} \end{cases}$$

$n = 2, 3, \ldots$, $k \in \mathbb{N}$.

Let $R_n \equiv \mathrm{E}[X_{n0} X_{n0}']$, and $\beta_n \equiv R_n^{-1} \mathrm{E}[X_{n0} Y_0]$, $n = 2, 3, \ldots$. Also, define $U_{nt} \equiv Y_t - X_{nt}'\beta_n$, $t \in \mathbb{Z}$, $n = 2, 3, \ldots$, and $\bar{\psi}_{nk} \equiv \mathrm{E}[U_{n,-k} Y_0]/\mathrm{E}[U_{n0}^2]$, $n = 2, 3, \ldots$. Note that in the last equation, $\mathrm{E}[U_{n0}^2] = \mathrm{E}[U_{n,-k}^2]$, because Y_t is stationary by Assumption 1. We can show the consistency of $\{\hat{\psi}_{nk}\}_{n=2}^{\infty}$ for ψ_k for each $k \in \mathbb{N}$, by showing that both terms on the right-hand side of the following equation converges to zero prob-P:

$$\hat{\psi}_{nk} - \psi_k = (\hat{\psi}_{nk} - \bar{\psi}_{nk}) + (\bar{\psi}_{nk} - \psi_k), \quad n = 2, 3, \ldots, k \in \mathbb{N}. \tag{3}$$

That is, the estimation errors of the impulse response estimator can be decomposed into two components, the first of which is the sampling error given the AR(p_n) model and the second of which is due to the specification error of the AR(p_n) model when used as an approximation for the AR process Y_t, whose order is possibly infinite. The proof below shows that both components of estimation errors converge to zeros as the sample size tends to infinity.

First, let F_{nt} be the population forecast of Y_t that would be made if the linear process in (2) were truncated at lag p_n and F_t^* the forecast when the whole sequence of past observations are used. That is, $F_{nt} \equiv \alpha + \sum_{j=1}^{p_n} \phi_j Y_{t-j}$, $t \in \mathbb{Z}$, $n = 2, 3, \ldots$, and $F_t^* \equiv \alpha + \sum_{j=1}^{\infty} \phi_j Y_{t-j}$, $t \in \mathbb{Z}$. Moreover, let $\{\{\beta_{nj}\}_{j \in \mathbb{N}}\}_{n=2}^{\infty}$ be a double array of real numbers such that for each $n = 2, 3, \ldots$ and each $j = 1, 2, \ldots, p_n + 1$, β_{nj} is the

jth element of β_n, while for each $n = 2, 3, \ldots$ and each $j = p_n + 2, p_n + 3, \ldots$, $\beta_{nj} = 0$.

Lemma 1 *Suppose that Assumptions 1 and 3 hold. Then* $\|U_{n0}\| \to \|\epsilon_0\|$, $\|X'_{n0}\beta_n - F_0^*\| \to 0$ *and* $\|X'_{n0}\beta_n - F_{n0}\| \to 0$ *as* $n \to \infty$. *If in addition Assumption 2 holds,*

$$(\beta_{n1} - \alpha)^2 + \sum_{j=2}^{\infty} (\beta_{nj} - \phi_{j-1})^2 \to 0 \text{ as } n \to \infty, \tag{4}$$

and

$$|\beta_n| \to \left(\alpha^2 + \sum_{j=1}^{\infty} \phi_j^2 \right)^{1/2} \quad \text{as } n \to \infty. \tag{5}$$

Proof of Lemma 1. Because $X'_{n0}\beta_n$ is the MSE-best linear predictor of Y_0 in terms of X_{n0}, which contains 1, Y_{-1}, …, Y_{-p_n}, we have that $\|\epsilon_0\|^2 \leq \|U_{n0}\|^2 \leq \|Y_0 - F_{n0}\|^2$. The right-hand side of this inequality is equal to $\|\epsilon_0 - (F_{n0} - F_0^*)\|^2 = \|\epsilon_0\|^2 + \|F_{n0} - F_0^*\|^2$, which converges to $\|\epsilon_0\|^2$ as $n \to \infty$ under Assumptions 1 and 3. It follows that $\|U_{n0}\|^2 \to \|\epsilon_0\|^2$ as $n \to \infty$.

Next, note that $\|U_{n0}\|^2 = \|\epsilon_0 - (X'_{n0}\beta_n - F_0^*)\|^2 = \|\epsilon_0\|^2 + \|X'_{n0}\beta_n - F_0^*\|^2$. Because the left-hand side of this equality converges to $\|\epsilon_0\|^2$ as shown above, $\|X'_{n0}\beta_n - F_0^*\|$ converges to zero as $n \to \infty$.

From the above results, it follows immediately that $\|X'_{n0}\beta_n - F_{n0}\|$ converges to zero as $n \to \infty$, because $\|X'_{n0}\beta_n - F_{n0}\| \leq \|X'_{n0}\beta_n - F_0^*\| + \|F_{n0} - F_0^*\|$, $n = 2, 3, \ldots$.

To prove (4), note that $(\beta_{n1} - \alpha)^2 + \sum_{j=2}^{\infty} (\beta_{nj} - \phi_{j-1})^2 = |\beta_n - \zeta_n|^2 + \sum_{j=p_n+1}^{\infty} \phi_j^2$, where $\zeta_n \equiv (\alpha, \phi_1, \phi_2, \ldots, \phi_{p_n})'$. Because the second term on the right-hand side of the equality converges to zero, it suffices to show that $|\beta_n - \zeta_n|^2$ converges to zero as $n \to \infty$. First, note that $\|X'_{n0}\beta_n - F_{n0}\|^2 = (\beta_n - \zeta_n)' R_n (\beta_n - \zeta_n)$, $n = 2, 3, \ldots$.

In the equation, the right-hand side is no smaller than the product of $|\beta_n - \zeta_n|^2$ and the minimum eigen value of R_n. Because the minimum eigen value of R_n is no less than $2\pi \bar{f}_1$, which is positive by Assumption 2(c), it follows that $\|X'_{n0}\beta_n - F_{n0}\|^2 \geq 2\pi \bar{f}_1 |\beta_n - \zeta_n|^2 \geq 0$, $n = 2, 3, \ldots$. Since the left-hand side converges to zero as $n \to \infty$, so does $\{|\beta_n - \zeta_n|\}_{n \in \mathbb{N}}$.

The convergence of (4) implies that the sequence $\{\beta_{nj}\}_{n=2}^{\infty}$ converges to the sequence $(\alpha, \phi_1, \phi_2, \ldots)$ in the ℓ_2 space as $n \to \infty$. It follows that the ℓ_2-norm of $\{\beta_{nj}\}_{n=2}^{\infty}$ converges to that of $(\alpha, \phi_1, \phi_2, \ldots)$, which is $(\alpha^2 + \sum_{j=1}^{\infty} \phi_j^2)^{1/2}$. Because $|\beta_n|$ coincides with the ℓ_2-norm of $\{\beta_{nj}\}_{n=2}^{\infty}$ for each $n = 2, 3, \ldots$, equation (5) therefore follows. ■

The following lemma shows that the misspecification error, $\bar{\psi}_{nk} - \psi_k$, in equation (3) converges to zeros as sample size tends to infinity.

Lemma 2 *Suppose that Assumptions 1 and 3 hold. Then for each* $k \in \mathbb{N}$, $\bar{\psi}_{nk} - \psi_k \to 0$ *as* $n \to \infty$.

Proof of Lemma 2. Let k be an arbitrary natural number. By Lemma 1, $\{U_{n,-k}\}_{n\in\mathbb{N}}$ converges to $\epsilon_{-k} = Y_{-k} - F^*_{-k}$ in terms of the L_2-metric. It follows that $\langle U_{n,-k}, Y_0\rangle \to \langle \epsilon_{-k}, Y_0\rangle$ and $\|U_{n0}\|^2 \to \|\epsilon_0\|^2$ as $n \to \infty$. Because $\bar{\psi}_{nk} = \langle U_{n,-k}, Y_0\rangle / \|U_{n0}\|^2$, $n \in \mathbb{N}$ and $\psi_k = \langle \epsilon_{-k}, Y_0\rangle/\|\epsilon_0\|^2$, the desired result follows. ■

To prove the convergence of $\hat{\psi}_{nk} - \bar{\psi}_{nk}$ to zero, we use some lemmas given in the Appendix. The memory condition on the process Y_t imposed in Assumption 2(b) and the speed of divergence of p_n imposed in Assumption 3 are essential for the desired results to hold.

Lemma 3 *Suppose that Assumptions 1–3 hold. Then for each $k \in \mathbb{N}$,*

$$|\hat{\psi}_{nk} - \bar{\psi}_{nk}| = \mathrm{O}_P(p_n/n^{1/2}) \text{ as } n \to \infty.$$

Proof of Lemma 3. The result follows from Lemma A.9 in the Appendix, by (22) of Lemma A.10. ■

We are now ready to state the consistency of $\{\hat{\psi}_{nk}\}_{n\in\mathbb{N}}$ for ψ_k.

Theorem 4 *Suppose that Assumptions 1–3 hold. Then for each $k \in \mathbb{N}$,*

$$\hat{\psi}_{nk} - \psi_k \to 0 \text{ as } n \to \infty \text{ prob-}P.$$

Proof of Theorem 4. By (3), we have that $|\hat{\psi}_{nk} - \psi_k| \le |\hat{\psi}_{nk} - \bar{\psi}_{nk}| + |\bar{\psi}_{nk} - \psi_k|$, $n = \bar{n}_k, \bar{n}_k + 1, \ldots$. The desired result follows from this inequality by Lemmas 2 and 3. ■

In actual application, the lag order of the AR model is often chosen by using a data-based lag order selection method. Here we consider the case in which an information criterion is used to select a lag order among the lag orders that does not exceed the maximum lag order $\bar{p}_n$ preselected for each sample size n. More concretely, suppose that the lag order p among $\{1, 2, \ldots, \bar{p}_n\}$ is selected to minimize $(n - \bar{p}_n)\log S_n(p) + pC(n)$, where

$$S_n(p) \equiv \min\left\{(n-\bar{p}_n)^{-1}\sum_{t=\bar{p}_n+1}^{n}\left(Y_t - b_1 - \sum_{j=1}^{p} b_{j+1}Y_{t-j}\right)^2 : (b_1, b_2, \ldots, b_{p+1}) \in \mathbb{R}^{p+1}\right\},$$
$$n = 2, 3, \ldots, \quad p \in \{1, 2, \ldots, \bar{p}_n\},$$

$C : \mathbb{N} \to (0, \infty)$ is a known function, and $pC(n)$ is the penalty term to encourage parsimony.

For each sample size $n = 2, 3, \ldots$, let $\tilde{\epsilon}_{n,\bar{p}_n+1}, \tilde{\epsilon}_{n,\bar{p}_n+2}, \ldots, \tilde{\epsilon}_{n,n}$ be the fitted residuals in the OLS autoregression of Y_t with the selected order $\tilde{p}_n$, and $\check{\epsilon}_{n,\bar{p}_n+1}, \check{\epsilon}_{n,\bar{p}_n+2}, \ldots, \check{\epsilon}_{n,n}$ the fitted residuals in the

OLS autoregression of Y_t with the maximum lag order $\bar{p}_n$. Using the fitted residuals from the selected model, we define an estimator $\{\tilde{\psi}_{nk}\}_{n=2}^{\infty}$ of ψ_k by

$$\tilde{\psi}_{nk} \equiv \begin{cases} \left((n-\bar{p}_n-k)^{-1}\sum_{t=\bar{p}_n+k+1}^{n}\tilde{\epsilon}_{n,t-k}^2\right)^{+} & \\ \qquad \times (n-\bar{p}_n-k)^{-1}\sum_{t=\bar{p}_n+k+1}^{n}\tilde{\epsilon}_{n,t-k}Y_t & \text{if } k \leq n-\bar{p}_n-1, \\ 0 & \text{otherwise,} \end{cases}$$

$n = 2, 3, \ldots$, $k \in \mathbb{N}$.

Assumption 4 The sequence $\{\bar{p}_n \in \mathbb{N}\}_{n=2}^{\infty}$ satisfies the conditions for $\{p_n\}$ in Assumption 3.

Assumption 5 $C : \mathbb{N} \to (0, \infty)$ satisfies that $\bar{p}_n C(n)/n \to 0$.

Lemma 5 *Suppose that Assumptions 1, 2, 4, and 5 hold. Then for each $n = 2, 3, \ldots$,*

$$(n-\bar{p}_n)^{-1}\sum_{t=\bar{p}_n+1}^{n}(\tilde{\epsilon}_{nt}-\check{\epsilon}_{nt})^2 = \mathrm{O}_P(\bar{p}_n C(n)/n) \text{ as } n \to \infty. \tag{6}$$

Proof of Lemma 5. By definition, we have that $S_n(\tilde{p}_n) = (n-\bar{p}_n)^{-1}\sum_{t=\bar{p}_n+1}^{n}\tilde{\epsilon}_{nt}^2$ and $S_n(\bar{p}) = (n-\bar{p}_n)^{-1}\sum_{t=\bar{p}_n+1}^{n}\check{\epsilon}_{nt}^2$. Since the larger the lag order is, the better the model fits the data, it follows that $S_n(\tilde{p}_n) - S_n(\bar{p}_n) \geq 0$, $n = 2, 3, \ldots$.

Next, since $\tilde{p}_n$ is the selected lag order, by construction we have that

$$(n-\bar{p}_n)\log S_n(\tilde{p}_n) + \tilde{p}_n C(n) \leq (n-\bar{p}_n)\log S_n(\bar{p}_n) + \bar{p}_n C(n), \quad n = 2, 3, \ldots.$$

When $S_n(\bar{p}_n) > 0$, we can rewrite this inequality as

$$\log(S_n(\tilde{p}_n)/S_n(\bar{p}_n)) \leq (\bar{p}_n - \tilde{p}_n)C(n)/(n-\bar{p}_n), \quad n = 2, 3, \ldots.$$

Using the fact that

$$\frac{S_n(\tilde{p}_n)}{S_n(\bar{p}_n)} = \frac{S_n(\tilde{p}_n) - S_n(\bar{p}_n)}{S_n(\bar{p}_n)} + 1, \quad n = 2, 3, \ldots,$$

we have that

$$\log\left(\frac{S_n(\tilde{p}_n) - S_n(\bar{p}_n)}{S_n(\bar{p}_n)} + 1\right) \leq \frac{(\bar{p}_n - \tilde{p}_n)C(n)}{n-\bar{p}_n} \leq \frac{\bar{p}_n C(n)}{n-\bar{p}_n}, \quad n = 2, 3, \ldots,$$

or equivalently,

$$S_n(\tilde{p}_n) - S_n(\bar{p}_n) \le \left(\exp\left(\frac{\bar{p}_n C(n)}{n-\bar{p}_n}\right) - 1\right) S_n(\bar{p}_n), \quad n = 2, 3, \dots.$$

This inequality holds even when $S_n(\bar{p}_n) = 0$, because $S_n(\tilde{p}_n) = 0$ whenever $S_n(\bar{p}_n) = 0$. It follows that

$$0 \le S_n(\tilde{p}_n) - S_n(\bar{p}_n) \le \left(\exp\left(\frac{\bar{p}_n C(n)}{n-\bar{p}_n}\right) - 1\right) S_n(\bar{p}_n), \quad n = 2, 3, \dots.$$

By using the law of iterated projections (Brockwell and Davis, 1991, Prop. 2.3.2(vii), pp. 52–53) and the orthogonality condition for the OLS regression, we can easily verify that

$$S_n(\tilde{p}_n) - S_n(\bar{p}_n) = (n-\bar{p}_n)^{-1} \sum_{t=\bar{p}_n+1}^{n} (\tilde{\epsilon}_{nt} - \check{\epsilon}_{nt})^2, \quad n = 2, 3, \dots.$$

Given this, we have that

$$0 \le (n-\bar{p}_n)^{-1} \sum_{t=\bar{p}_n+1}^{n} (\tilde{\epsilon}_{nt} - \check{\epsilon}_{nt})^2 \le \left(\exp\left(\frac{\bar{p}_n C(n)}{n-\bar{p}_n}\right) - 1\right) (n-\bar{p}_n)^{-1} \sum_{t=\bar{p}_n+1}^{n} \check{\epsilon}_{nt}^2, \quad n = 2, 3, \dots.$$

Under Assumptions 4 and 5, we have that $\exp(\bar{p}_n C(n)(n-\bar{p}_n)) - 1 = O(\bar{p}_n C(n)/n)$ as $n \to \infty$. We also have that

$$(n-\bar{p}_n)^{-1} \sum_{t=\bar{p}_n+1}^{n} \check{\epsilon}_{nt}^2 = \left((n-\bar{p}_n)^{-1} \sum_{t=\bar{p}_n+1}^{n} \check{\epsilon}_{nt}^2 - \mathrm{E}[U_{n,0}^2]\right) + (\mathrm{E}[U_{n,0}^2] - \mathrm{E}[\epsilon_0^2]) + \mathrm{E}[\epsilon_0^2],$$
$$n = 2, 3, \dots.$$

Applying Lemma A.9 and (21) of Lemma A.10 in the Appendix, we can verify that the first term on the right-hand side of this equality is $\mathrm{O}_P(\bar{p}_n/n^{1/2})$ as $n \to \infty$. The second term is $o(1)$ by Lemma 1. It follows that $(n-\bar{p}_n)^{-1} \sum_{t=\bar{p}_n+1}^{n} \check{\epsilon}_{nt}^2 = \mathrm{O}_P(1)$ as $n \to \infty$. The desired result (6) therefore follows. ■

Theorem 6 *Suppose that Assumptions 1, 2, 4, and 5 hold. Then for each $k \in \mathbb{N}$,*

$$\tilde{\psi}_{nk} - \bar{\psi}_{nk} = \mathrm{O}_P(\bar{p}_n/n^{1/2}) + \mathrm{O}_P(\bar{p}_n^{1/2} C(n)^{1/2}/n^{1/2}) \text{ as } n \to \infty,$$

and

$$\tilde{\psi}_{nk} \to \psi_k \text{ as } n \to \infty \text{ prob-}P.$$

Proof of Theorem 6. When setting $p_n = \bar{p}_n$, Assumptions 1–3 hold under Assumptions 1, 2, 4, and 5. We can apply Lemma A.9 in the Appendix; then it follows from (6) that

$$\left((n-\bar{p}_n)^{-1}\sum_{t=\bar{p}_n+1}^{n}(\tilde{\epsilon}_{nt}-U_{nt})^2\right)^{1/2}$$
$$\leq \left((n-\bar{p}_n)^{-1}\sum_{t=\bar{p}_n+1}^{n}(\tilde{\epsilon}_{nt}-\check{\epsilon}_{nt})^2\right)^{1/2}+\left((n-\bar{p}_n)^{-1}\sum_{t=\bar{p}_n+1}^{n}(\check{\epsilon}_{nt}-U_{nt})^2\right)^{1/2}$$
$$= \mathrm{O}_P(\bar{p}_n^{1/2}C(n)^{1/2}/n^{1/2})+\mathrm{O}_P(\bar{p}_n/n^{1/2}) \text{ as } n\to\infty.$$

Given this, the first result follows by Lemmas A.10 in the Appendix. The second result follows from the first result and Lemma 2, since $|\tilde{\psi}_{nk}-\psi_k|\leq|\tilde{\psi}_{nk}-\bar{\psi}_{nk}|+|\bar{\psi}_{nk}-\psi_k|$, $n=2,3,\ldots$, $k\in\mathbb{N}$. ■

4. EMPIRICAL EXAMPLE

In this section, we use both the conventional and the proposed impulse response estimators to analyze the dynamics of the French real exchange rate series vis-à-vis the US dollar. The nominal monthly exchange rate and price indices used to construct the log real exchange rate come from International Financial Statistics CD-ROM. In particular, the monthly end-of-period nominal exchange rate (line "ae" in the CD-ROM) and the consumer price index (line 64) were used. The data covers the period from April 1973 to August 1998. This amounts to a sample size of 305.

We use the AIC and SIC to select the model from the AR(p) models with $p\leq 12$. Both criteria select AR(1) for the data under study, and the model estimates of (α,ϕ_1) are presented in Table 1. After the model is selected and estimated, we can calculate the conventional impulse response estimates ψ_{nk}^{conv} and the proposed impulse response estimates $\tilde{\psi}_{nk}$. To gauge the uncertainty of these two impulse response function estimates, we also present their standard errors. The proposed impulse response estimator can be viewed as a two-stage quasi-maximum likelihood estimator, where the AR($\tilde{p}_n$) model is estimated at the first stage and $\tilde{\psi}_{nk}$ is estimated at the second stage. We can apply the asymptotic normality result of Theorem 6.10 in White (1994) for two-stage estimation to obtain the covariance matrix estimates of $(\tilde{\beta}_n,\tilde{\psi}_{nk})$. The methodology in Newey and West (1994) is used to obtain positive semi-definite covariance matrix estimates. In particular, the Bartlett window is used, and the lag selection parameter, used to compute the bandwidth of the Bartlett window based on the data, is set to be $4(T/100)^{2/9}$ (where $T=n-\bar{p}_n$, in our current context). Since the conventional impulse response estimator ψ_{nk}^{conv} is a function of $\tilde{\beta}_n$, its variance can be estimated based on the covariance matrix of $\tilde{\beta}_n$ using the delta method. These estimates are presented in Table 2 and plotted in Figure A.

In Figure A, ψ_{nk}^{conv} and $\tilde{\psi}_{nk}$ describe the dynamics of the French real exchange rate quite differently.

In panel (b), $\tilde{\psi}_{nk}$ shows nonmonotonic decay in the response of the French real exchange rate to shocks, which contrasts with the smooth shape of ψ_{nk}^{conv}. There are also some initial hikes in the response function of $\tilde{\psi}_{nk}$, which ψ_{nk}^{conv} does not indicate. By $\tilde{\psi}_{nk}$, the French real exchange rate also shows a possible drop in the IRF around the twentieth month. Because negative impulse responses are not sensible in this case, the fact that French $\tilde{\psi}_{nk}$ deviates from zeros and goes into the negative zone after the twentieth month might be reflecting the estimation errors around the true value close to zero. We may consider the true IRF in this range to be negligibly small.

The confidence interval of ψ_{nk}^{conv} inflates along the horizon up to a point where it starts to shrink, which is consistent with the fact that any stationary AR model forces ψ_{nk}^{conv} to converge to zero eventually. Therefore, there is less sampling uncertainty about ψ_{nk}^{conv} once the horizon k becomes large. On the other hand, the confidence interval of $\tilde{\psi}_{nk}$ just widens along the horizon. Because $\tilde{\psi}_{nk}$ is calculated from the regression of Y_t on $\tilde{\epsilon}_{n,t-k}$. As k increases, $\tilde{\epsilon}_{n,t-k}$ has less power in predicting Y_t and that leads to higher noise relative to signals contained in $\tilde{\psi}_{nk}$.

5. MONTE CARLO SIMULATION

The IRF estimator $\tilde{\psi}_{nk}$ shows an interesting picture for the dynamics of the French real exchange rates. The estimated IRF is nonsmooth and exhibits jumps. If this estimated IRF were the true IRF, could the proposed IRF estimator capture the shape of the IRF well? How about the conventional IRF estimator? Motivated by these questions, we are going to use this estimated IRF with slight modifications as our DGP in the following Monte Carlo simulation and see whether the conventional and the proposed IRF estimator could capture these features successfully. Specifically, we take the IRF estimate $\tilde{\psi}_{nk}$ for the French real exchange rate and truncate the impulse responses at $k = 20$. The resulting MA(19) process with coefficients equal to the values of French $\tilde{\psi}_{nk}$ for $k = 1, 2, \cdots, 19$ is the DGP used in our Monte Carlo simulation.

To isolate the effect of uncertainty in model selection, we run our Monte Carlo simulation for each of the models $\{\text{AR}(p)|p \leq 12\}$ and compare the performance of ψ_{nk}^{conv} and $\tilde{\psi}_{nk}$. In each simulation, the number of replications is 1000. The i.i.d. standard Gaussian errors are used for innovations, and the simulated sample size is 305, which is the size of the empirical example in Section 4. The performance of the impulse response estimators are compared based on their biases, and root mean squared errors (RMSE). We call this Monte Carlo simulation I.

The results are reported in Figures A and A. Figure A plots the means of both estimators as opposed to the true IRF. As can be seen from the figure, $\tilde{\psi}_{nk}$ performs remarkably well in detecting the nonsmooth shape of the true IRF. It mimics the shape of the true IRF well when the impulse responses are significant and, most importantly, it can capture the sudden drop to zeros in the true IRF regardless of the estimation model. On the other hand, we see that ψ_{nk}^{conv} exhibits a smooth shape and fails to convey irregular changes of the true IRF. The above result implies that if we use the conventional, instead of the proposed, impulse response estimator in estimating this kind of true IRF, we could have overestimated the persistence of the true process or underestimated its degree of mean reversion.

Figure A shows the RMSE's of these two IRF estimators. As can be seen from the figures, $\tilde{\psi}_{nk}$ has roughly constant RMSE across middle to long horizons, while ψ_{nk}^{conv} exhibits more erratic RMSE around the mid-range. The smaller RMSE of the ψ_{nk}^{conv} at the long horizons does not mean that ψ_{nk}^{conv} outperforms $\tilde{\psi}_{nk}$, but it is an artifact resulting from the setup of the conventional method that ψ_{nk}^{conv} is forced to converge to zero eventually for any stationary AR models.

It is expected that the conventional IRF estimator ψ_{nk}^{conv} will perform well when the model is correctly specified for the underlying DGP. Meanwhile, the new IRF estimator $\tilde{\psi}_{nk}$ should also work well, given that the estimated innovations $\{\tilde{\epsilon}_t\}$ accurately approximate the true innovations $\{\epsilon_t\}$. To verify this, we run another Monte Carlo simulation. In this simulation, the DGP is set to be AR(2): $Y_t = 1.2Y_{t-1} - 0.25Y_{t-2} + \epsilon_t$, where ϵ_t is i.i.d. standard Gaussian errors. The IRF of this DGP exhibits three features. It shows a hump in the early horizon with a magnitude of 1.2, the impulse response around the mid-range ($k = 20$) is still large, but it dies out significantly in the long horizon ($k = 40$).

Again, we run our Monte Carlo simulation for each of the models $\{\text{AR}(p)|p \leq 12\}$ and compare the performance of ψ_{nk}^{conv} and $\tilde{\psi}_{nk}$ based on their biases, and RMSE's. In each simulation, the number of replications is 1000, and the simulated sample size is 305, the size of the empirical example in Section 4. We call this Monte Carlo simulation II.

The results are reported in Figures A and A. Figure A plots the means of both estimators as opposed to the true IRF. As can be seen from the figure, ψ_{nk}^{conv} performs quite well, except for the AR(1) model, which is still not adequate and misspecified for the true DGP. Meanwhile, the new IRF estimator $\tilde{\psi}_{nk}$ is not affected by such misspecification and performs equally well across all estimation models. Both estimators exhibit noticeable bias in the mid-range but the conventional estimator improves as the horizon increases, while the new IRF estimator maintains a similar magnitude of bias into the long horizon. The bias could be explained by the small sample size of the LS estimation for AR models. It should be improved as the sample size of the data increases. That the conventional estimator outperforms the new IRF estimator in the long horizon might be attributed to the fact that the former imposes more structure on the estimated IRF than the latter. Therefore, when the model is correctly specified for the data, the conventional estimator could pin down the true IRF in the long horizon more precisely.

Figure A shows the RMSE's of these two IRF estimators. Except for the AR(1) model, the pictures are quite similar across different models. The RMSE's for both estimators increase in the early horizon, and start to decrease around the mid-range for the conventional estimator but maintain a roughly constant magnitude in the case of the new IRF estimator. Combined with the observations on the bias above, we can see that the difference in the RMSE's of these two estimators basically reflect the difference in their biases, and is slightly augmented by the relatively larger sampling uncertainty of the new IRF estimator.

In sum, the Monte Carlo simulations show that when the estimation model is misspecified for the underlying DGP, the new IRF estimator is still capable of detecting interesting aspects of the true IRF, whether it is smooth or nonsmooth with sudden reversion to zeros. On the other hand, the conventional IRF estimator is sensitive to model specifications. It works well when the specified model is correct for the underlying process, but works poorly if the estimation model is misspecified. In this sense, the new IRF estimator is a more robust IRF estimator.

6. CONCLUSION

This paper proposes an alternative approach to estimating impulse response function, which is asymptotically valid yet is less sensitive to model misspecifications in small samples. Our Monte Carlo simulations demonstrate that the proposed estimator is superior to the conventional estimator in small samples when the estimation model is incorrect for the underlying process. On the other hand, in the ideal case where the estimation specification happens to be correct for the DGP, the Monte Carlo study shows that the proposed estimator performs comparably well to the conventional estimator.

Although we only investigate the performance of this new IRF estimator in the univariate case, the proposed methodology can be generalized to cover the multivariate case. For example, to obtain the new IRF estimate in a vector autoregression (VAR) framework, we can obtain the estimated vector of residuals by running OLS, and then regress the vector of variables on the vector of residuals with k horizons difference to obtain the square matrix of impulse responses of the variables to the (un-orthogonalized) residuals at horizon k. This generalization is left for future work.

ACKNOWLEDGEMENTS

This paper is a revision of the third chapter of Chang's Ph.D. dissertation at the University of Michigan. Chang gratefully acknowledges the financial support of Dissertation Fellowships from the Chiang Ching-kuo Foundation for International Scholarly Exchange (USA).

APPENDIX: LEMMAS FOR SECTION 3

Among the results given in this appendix, Lemmas A.9 and A.10 are directly used in Section 3. All other results are used to prove Lemmas A.9 and A.10.

Lemma A.1 *Suppose that Assumptions 1 and 2(a)(b) hold. Let*

$$A \equiv (1+2|\mu|)\left(\sum_{\tau=-\infty}^{\infty}|\gamma(\tau)|\right)^{1/2} + \left(2\sum_{\tau=-\infty}^{\infty}\gamma(\tau)^2 + \sum_{\tau_1=-\infty}^{\infty}\sum_{\tau_2=-\infty}^{\infty}\sum_{\tau_3=-\infty}^{\infty}|\kappa_4(0,\tau_1,\tau_2,\tau_3)|\right)^{1/2}.$$

Then for each integers t_1 and t_2 such that $t_1 < t_2$,

$$\mathrm{var}\left[(t_2-t_1)^{-1}\sum_{t=t_1+1}^{t_2} Y_t\right] \le (t_2-t_1)^{-1}A^2 \tag{7}$$

and

$$\mathrm{var}\left[(t_2-t_1)^{-1}\sum_{t=t_1+1}^{t_2} Y_tY_{t+m}\right] \le (t_2-t_1)^{-1}A^2, \quad m\in\mathbb{Z}. \tag{8}$$

Proof of Lemma A.1. Inequality (7) holds because

$$\begin{aligned}
\mathrm{var}\left[(t_2-t_1)^{-1}\sum_{t=t_1+1}^{t_2} Y_t\right] =&(t_2-t_1)^{-2}\sum_{t=t_1+1}^{t_2}\sum_{s=t_1+1}^{t_2}\mathrm{Cov}[Y_t,Y_s] \le (t_2-t_1)^{-2}\sum_{t=t_1+1}^{t_2}\sum_{s=t_1+1}^{t_2}|\gamma(t-s)|\\
=&(t_2-t_1)^{-2}\sum_{\tau=-(t_2-t_1-1)}^{(t_2-t_1-1)}(t_2-t_1-|\tau|)|\gamma(\tau)|\\
=&(t_2-t_1)^{-1}\sum_{\tau=-(t_2-t_1-1)}^{(t_2-t_1-1)}\left(1-\frac{|\tau|}{t_2-t_1}\right)|\gamma(\tau)|\\
\le&(t_2-t_1)^{-1}\sum_{\tau=-(t_2-t_1-1)}^{(t_2-t_1-1)}|\gamma(\tau)| \le (t_2-t_1)^{-1}\sum_{\tau=-\infty}^{\infty}|\gamma(\tau)|\\
\le&(t_2-t_1)^{-1}A^2.
\end{aligned} \tag{9}$$

To show (8), let m be an arbitrary integer. For each pair of integers, t and s, between t_1 and t_2, we have that $Y_tY_s = (Z_t+\mu)(Z_s+\mu) = Z_tZ_s + \mu Z_t + \mu Z_s + \mu^2$, and $\mathrm{E}[Y_tY_s] = \gamma(s-t)+\mu^2$. It follows that $Y_tY_s - \mathrm{E}[Y_tY_s] = (Z_tZ_s - \gamma(s-t)) + \mu Z_t + \mu Z_s$. Using this fact, we obtain that

$$\mathrm{var}\left[(t_2-t_1)^{-1}\sum_{t=t_1+1}^{t_2} Y_tY_{t+m}\right] = \mathrm{E}\left[\left((t_2-t_1)^{-1}\sum_{t=t_1+1}^{t_2}((Z_tZ_{t+m}-\gamma(m)) + \mu Z_t + \mu Z_{t+m})\right)^2\right].$$

It follows that

$$\begin{aligned}\operatorname{var}\left[(t_2-t_1)^{-1}\sum_{t=t_1+1}^{t_2} Y_tY_{t+m}\right]^{1/2} &= \left\|(t_2-t_1)^{-1}\sum_{t=t_1+1}^{t_2}((Z_tZ_{t+m}-\gamma(m))+\mu Z_t+\mu Z_{t+m})\right\| \\ \leq \left\|(t_2-t_1)^{-1}\sum_{t=t_1+1}^{t_2}(Z_tZ_{t+m}-\gamma(m))\right\| &+ |\mu|\left\|(t_2-t_1)^{-1}\sum_{t=t_1+1}^{t_2} Z_t\right\| + |\mu|\left\|(t_2-t_1)^{-1}\sum_{t=t_1+1}^{t_2} Z_{t+m}\right\|.\end{aligned} \tag{10}$$

For the first term on the right-hand side of (10), we have that

$$\begin{aligned}\left\|(t_2-t_1)^{-1}\sum_{t=t_1+1}^{t_2}(Z_tZ_{t+m}-\gamma(m))\right\|^2 &= (t_2-t_1)^{-2}\sum_{t=t_1+1}^{t_2}\sum_{s=t_1+1}^{t_2}\mathrm{E}[(Z_tZ_{t+m}-\gamma(m))(Z_sZ_{s+m}-\gamma(m))] \\ &= (t_2-t_1)^{-2}\sum_{t=t_1+1}^{t_2}\sum_{s=t_1+1}^{t_2}(\mathrm{E}[Z_tZ_{t+m}Z_sZ_{s+m}]-\gamma(m)^2) \\ &= (t_2-t_1)^{-2}\sum_{t=t_1+1}^{t_2}\sum_{s=t_1+1}^{t_2}(\kappa_4(t,t+m,s,s+m)+\gamma(s-t)^2+\gamma(s-t+m)\gamma(s-t-m)) \\ &= (t_2-t_1)^{-2}\sum_{t=t_1+1}^{t_2}\sum_{s=t_1+1}^{t_2}(\kappa_4(0,m,s-t,s-t+m)+\gamma(s-t)^2+\gamma(s-t+m)\gamma(s-t-m)) \\ &= (t_2-t_1)^{-2}\sum_{\tau=-(t_2-t_1-1)}^{t_2-t_1-1}(t_2-t_1-|\tau|)(\kappa_4(0,m,\tau,\tau+m)+\gamma(\tau)^2+\gamma(\tau+m)\gamma(\tau-m)) \\ &\leq (t_2-t_1)^{-1}\sum_{\tau=-(t_2-t_1-1)}^{t_2-t_1-1}\left(1-\frac{|\tau|}{t_2-t_1}\right)(|\kappa_4(0,m,\tau,\tau+m)|+\gamma(\tau)^2+|\gamma(\tau+m)\gamma(\tau-m)|) \\ &\leq (t_2-t_1)^{-1}\sum_{\tau=-(t_2-t_1-1)}^{t_2-t_1-1}(|\kappa_4(0,m,\tau,\tau+m)|+\gamma(\tau)^2+|\gamma(\tau+m)\gamma(\tau-m)|),\end{aligned}$$

where the third equality follows by the fact that

$$\begin{aligned}\kappa_4(t_1,t_2,t_3,t_4) &= \mathrm{E}[Z_{t_1}Z_{t_2}Z_{t_3}Z_{t_4}]-\mathrm{Cov}[Z_{t_1},Z_{t_2}]\mathrm{Cov}[Z_{t_3},Z_{t_4}] \\ &\quad -\mathrm{Cov}[Z_{t_1},Z_{t_3}]\mathrm{Cov}[Z_{t_2},Z_{t_4}]-\mathrm{Cov}[Z_{t_1},Z_{t_4}]\mathrm{Cov}[Z_{t_2},Z_{t_3}], \quad (t_1,t_2,t_3,t_4)\in\mathbb{Z}^4.\end{aligned}$$

Because

$$\sum_{\tau=-(t_2-t_1-1)}^{t_2-t_1-1}|\kappa_4(0,m,\tau,\tau+m)| \leq \sum_{\tau_1=-\infty}^{\infty}\sum_{\tau_2=-\infty}^{\infty}\sum_{\tau_3=-\infty}^{\infty}|\kappa_4(0,\tau_1,\tau_2,\tau_3)|,$$

$$\sum_{\tau=-(t_2-t_1-1)}^{t_2-t_1-1}\gamma(\tau)^2 \leq \sum_{\tau=-\infty}^{\infty}\gamma(\tau)^2,$$

and

$$\sum_{\tau=-(t_2-t_1-1)}^{t_2-t_1-1} |\gamma(\tau+m)\gamma(\tau-m)| \leq \left(\sum_{\tau=-(t_2-t_1-1)}^{t_2-t_1-1} \gamma(\tau+m)^2\right)^{1/2} \left(\sum_{\tau=-(t_2-t_1-1)}^{t_2-t_1-1} \gamma(\tau-m)^2\right)^{1/2}$$
$$\leq \sum_{\tau=-\infty}^{\infty} \gamma(\tau)^2$$

by the Cauchy-Schwartz inequality, it follows that

$$\left\|(t_2-t_1)^{-1} \sum_{t=t_1+1}^{t_2} (Z_t Z_{t+m} - \gamma(m))\right\|$$
$$\leq (t_2-t_1)^{-1/2} \left(\sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} \sum_{\tau_3=-\infty}^{\infty} |\kappa_4(0,\tau_1,\tau_2,\tau_3)| + 2 \sum_{\tau=-\infty}^{\infty} \gamma(\tau)^2\right)^{1/2}.$$

For the second and third term on the right-hand side of (10), we have that

$$\left\|(t_2-t_1)^{-1} \sum_{t=t_1+1}^{t_2} Z_{t+m}\right\| = \left\|(t_2-t_1)^{-1} \sum_{t=t_1+1}^{t_2} Z_t\right\|$$
$$= \operatorname{var}\left[(t_2-t_1)^{-1} \sum_{t=t_1+1}^{t_2} Y_t\right]^{1/2} \leq (t_2-t_1)^{-1/2} \left(\sum_{\tau=-\infty}^{\infty} |\gamma(\tau)|\right)^{1/2},$$

by (9). Thus, we have that

$$\operatorname{var}\left[(t_2-t_1)^{-1} \sum_{t=t_1+1}^{t_2} Y_t Y_{t+m}\right]^{1/2}$$
$$\leq (t_2-t_1)^{-1/2} \left(\sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} \sum_{\tau_3=-\infty}^{\infty} |\kappa_4(0,\tau_1,\tau_2,\tau_3)| + 2 \sum_{\tau=-\infty}^{\infty} \gamma(\tau)^2\right)^{1/2}$$
$$+ 2|\mu|(t_2-t_1)^{-1/2} \left(\sum_{\tau=-\infty}^{\infty} |\gamma(\tau)|\right)^{1/2}.$$

Inequality (8) therefore follows. ∎

Lemma A.2 *Suppose that Assumptions 1, 2(a)(b), and 3 hold. Then for each* $k = 0, 1, 2, \ldots$,

$$\left|(n-p_n-k)^{-1} \sum_{t=p_n+1}^{n-k} X_{nt} X'_{nt} - R_n\right| = \mathrm{O}_P(p_n/n^{1/2}) \text{ as } n \to \infty \text{ prob-}P, \tag{11}$$

$$\left|(n-p_n-k)^{-1} \sum_{t=p_n+1}^{n-k} X_{nt} U_{nt}\right| = \mathrm{O}_P(p_n/n^{1/2}) \text{ as } n \to \infty \text{ prob-}P, \tag{12}$$

$$\left|(n-p_n-k)^{-1}\sum_{t=p_n+1}^{n-k}U_{n,t}^2-\mathrm{E}[U_{n0}^2]\right|=\mathrm{O}_P(p_n/n^{1/2})\ \textit{as}\ n\to\infty\ \textit{prob-P}, \tag{13}$$

and

$$\left|(n-p_n-k)^{-1}\sum_{t=p_n+1}^{n-k}U_{n,t}Y_{t+k}-\mathrm{E}[U_{n,-k}Y_0]\right|=\mathrm{O}_P(p_n^{1/2}/n^{1/2})\ \textit{as}\ n\to\infty\ \textit{prob-P}, \tag{14}$$

Proof of Lemma A.2. Let k be an arbitrary nonnegative integer. To prove (11), note that for each $n=\bar{n}_k,\bar{n}_k+1,\ldots,$

$$\begin{aligned}0&\le\left|(n-p_n-k)^{-1}\sum_{t=p_n+1}^{n-k}X_{nt}X_{nt}'-R_n\right|^2\\&\le\mathrm{tr}\left(\left((n-p_n-k)^{-1}\sum_{t=p_n+1}^{n-k}X_{nt}X_{nt}'-R_n\right)\left((n-p_n-k)^{-1}\sum_{t=p_n+1}^{n-k}X_{nt}X_{nt}'-R_n\right)'\right).\end{aligned}$$

Because $\mathrm{E}[X_{nt}X_{nt}']=R_n$, $t\in\mathbb{Z}$, $n\in\mathbb{N}$, the mean of the right-hand side in the above inequality is no greater than $(p_n+1)^2(n-p_n-k)^{-1}A^2$ by Lemma A.1. Thus, we have that for each $n=\bar{n}_k,\bar{n}_k+1,\ldots$

$$\mathrm{E}\left[\left(\frac{p_n^2}{n}\right)^{-1}\left|(n-p_n-k)^{-1}\sum_{t=p_n+1}^{n-k}X_{nt}X_{nt}'-R_n\right|^2\right]\le\frac{(p_n+1)^2/p_n^2}{(n-p_n-k)/n}A^2.$$

Because the right-hand side is bounded as $n\to\infty$, it follows by the Markov inequality that

$$\left(\frac{p_n^2}{n}\right)^{-1}\left|(n-p_n-k)^{-1}\sum_{t=p_n+1}^{n-k}X_{nt}X_{nt}'-R_n\right|^2=\mathrm{O}_P(1)\text{ as }n\to\infty.$$

This leads to the desired result (11).

To show (12), let $\eta_{nt}\equiv(Y_t,1,Y_{t-1},\ldots,Y_{t-p_n})'$, $t\in\mathbb{Z}$, $n=2,3,\ldots$. Then

$$U_{nt}=\eta_{nt}'\begin{pmatrix}1\\-\beta_n\end{pmatrix},\quad t\in\mathbb{Z},\ n=2,3,\ldots.$$

Because $\mathrm{E}[X_{nt}U_{nt}] = 0$, $t \in \mathbb{Z}$, $n = 2, 3, \ldots$, it follows that for each $n = \bar{n}_k, \bar{n}_k + 1, \ldots$

$$
\begin{aligned}
0 \le \left| (n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} X_{nt}U_{nt} \right|^2 &= \left| (n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} (X_{nt}U_{nt} - \mathrm{E}[X_{nt}U_{nt}]) \right|^2 \\
&= \left| (n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} (X_{nt}\eta_{nt}' - \mathrm{E}[X_{nt}\eta_{nt}']) \begin{pmatrix} 1 \\ -\beta_n \end{pmatrix} \right|^2 \\
&\le \left| (n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} (X_{nt}\eta_{nt}' - \mathrm{E}[X_{nt}\eta_{nt}']) \right|^2 \left| \begin{pmatrix} 1 \\ -\beta_n \end{pmatrix} \right|^2 \\
&\le (1 + |\beta_n|^2) \operatorname{tr}\left(\left((n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} (X_{nt}\eta_{nt}' - \mathrm{E}[X_{nt}\eta_{nt}']) \right) \right. \\
&\qquad \left. \times \left((n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} (X_{nt}\eta_{nt}' - \mathrm{E}[X_{nt}\eta_{nt}']) \right)' \right)
\end{aligned}
$$

Using Lemma A.1, we have that for each $n = \bar{n}_k, \bar{n}_k + 1, \ldots$

$$
\begin{aligned}
\mathrm{E}\Bigg[\operatorname{tr}\Bigg(\left((n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} (X_{nt}\eta_{nt}' - \mathrm{E}[X_{nt}\eta_{nt}']) \right) \\
\times \left((n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} (X_{nt}\eta_{nt}' - \mathrm{E}[X_{nt}\eta_{nt}']) \right)' \Bigg) \Bigg] \\
\le (p_n + 1)(p_n + 2)(n - p_n - k)^{-1} A^2.
\end{aligned}
$$

Hence it follows that for each $n = \bar{n}_k, \bar{n}_k + 1, \ldots$

$$
\mathrm{E}\left[\left(\frac{p_n^2}{n} \right)^{-1} \left| (n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} X_{nt}U_{nt} \right|^2 \right] \le (1 + |\beta_n|^2) \frac{(p_n + 1)(p_n + 2)/p_n^2}{(n - p_n - k)/n} A^2.
$$

The right-hand side is bounded as $n \to \infty$, so (12) follows by the Markov inequality.

To show (13), we use the facts that

$$
U_{nt}^2 = (1, -\beta_n')\eta_{nt}\eta_{nt}' \begin{pmatrix} 1 \\ -\beta_n \end{pmatrix}, \quad t \in \mathbb{Z},\ n = 2, 3, \ldots
$$

and

$$
\mathrm{E}[U_{n0}^2] = (1, -\beta_n')\mathrm{E}[\eta_{n0}\eta_{n0}'] \begin{pmatrix} 1 \\ -\beta_n \end{pmatrix}, \quad n = 2, 3, \ldots.
$$

Given these equalities, we have that for each $n = 2, 3, \ldots$

$$\begin{aligned}
&\left|(n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} U_{n,t}^2 - \mathrm{E}[U_{n0}^2]\right|^2 \\
&\leq \left|(1, -\beta_n')(n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} (\eta_{nt}\eta_{nt}' - \mathrm{E}[\eta_{n0}\eta_{n0}']) \begin{pmatrix} 1 \\ -\beta_n \end{pmatrix}\right|^2 \\
&\leq \mathrm{tr}\left(\left((n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} (\eta_{nt}\eta_{nt}' - \mathrm{E}[\eta_{n0}\eta_{n0}'])\right)\left((n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} (\eta_{nt}\eta_{nt}' - \mathrm{E}[\eta_{n0}\eta_{n0}'])\right)'\right) \\
&\qquad \times \left|\begin{pmatrix} 1 \\ -\beta_n \end{pmatrix}\right|^4.
\end{aligned}$$

Divide both sides of this inequality by (p_n^2/n) and take expectations on both sides. We can apply Lemma A.1 in a similar fashion and obtain that

$$\mathrm{E}\left[\left(\frac{p_n^2}{n}\right)^{-1} \left|(n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} U_{n,t}^2 - \mathrm{E}[U_{n0}^2]\right|^2\right] \leq (1 + |\beta_n|^2)^2 \frac{(p_n + 2)^2/p_n^2}{(n - p_n - k)/n} A^2, \quad n = 2, 3, \ldots.$$

Because the right-hand side is bounded as $n \to \infty$, (13) therefore follows by the Markov inequality.

To prove (14), note that $U_{nt}Y_{t+k} = (1, -\beta_n')\eta_{nt}Y_{t+k}$ and $\mathrm{E}[U_{nt}Y_{t+k}] = (1, -\beta_n')\mathrm{E}[\eta_{n,-k}Y_0]$, $t \in \mathbb{Z}$, $n = 2, 3, \ldots$. Given these facts, we have that

$$\begin{aligned}
&\left|(n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} U_{n,t}Y_{t+k} - \mathrm{E}[U_{n,-k}Y_0]\right| \\
&\qquad = \left|(1, -\beta_n')(n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} (\eta_{nt}Y_{t+k} - \mathrm{E}[\eta_{n,-k}Y_0])\right| \\
&\qquad\qquad \leq \left|\begin{pmatrix} 1 \\ -\beta_n \end{pmatrix}\right| \left|\left((n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} (\eta_{nt}Y_{t+k} - \mathrm{E}[\eta_{n,-k}Y_0])\right)\right|.
\end{aligned}$$

It follows by Lemma A.1 that

$$\begin{aligned}
&\mathrm{E}\left[(p_n/n)^{-1} \left|(n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} U_{n,t}Y_{t+k} - \mathrm{E}[U_{n,-k}Y_0]\right|^2\right] \\
&\qquad \leq (1 + |\beta_n|^2)\mathrm{E}\left[\left|(n - p_n - k)^{-1} \sum_{t=p_n+1}^{n-k} (\eta_{nt}Y_{t+k} - \mathrm{E}[\eta_{n,-k}Y_0])\right|^2\right] / (p_n/n) \\
&\qquad\qquad \leq (1 + |\beta_n|^2)(p_n + 2)\,(n - p_n - k)^{-1} A^2/(p_n/n) = (1 + |\beta_n|^2) \frac{(p_n + 2)/p_n}{(n - p_n - k)/n} A^2.
\end{aligned}$$

Because the right-hand side is bounded as $n \to \infty$, the desired result (14) follows by the Markov inequality.

■

Lemma A.3 *Suppose that Assumptions 1, 2(a)(b), and 3 hold. Let* λ_{n1} *and* λ_{n2} *be the minimum and maximum eigenvalues of* R_n*, respectively,* $n = 2, 3, \ldots$*. Also, let* $\hat{\lambda}_{n1}$ *and* $\hat{\lambda}_{n2}$ *be the minimum and maximum eigenvalues of* $\hat{R}_n$*, respectively,* $n = 2, 3, \ldots$*. Then* $|\hat{\lambda}_{n1} - \lambda_{n1}| = \mathrm{O}_P(p_n/n^{1/2})$ *and* $|\hat{\lambda}_{n2} - \lambda_{n2}| = \mathrm{O}_P(p_n/n^{1/2})$ *as* $n \to \infty$ *prob-P.*

Proof of Lemma A.3. Let S^{m-1} denote the unit sphere in the m-dimensional Euclidean space for each $m \in \mathbb{N}$. Then for each $n \in \mathbb{N}$,

$$\begin{aligned}\hat{\lambda}_{n1} &= \min_{a \in S^{p_n}} a'\hat{R}_n a = \min_{a \in S^{p_n}} a'(R_n + (\hat{R}_n - R_n))a \\ &\geq \min_{a \in S^{p_n}} a'R_n a - \max_{a \in S^{p_n}} |a'(\hat{R}_n - R_n)a| = \lambda_{n1} - |\hat{R}_n - R_n|,\end{aligned}$$

and

$$\begin{aligned}\lambda_{n1} &= \min_{a \in S^{p_n}} a'R_n a = \min_{a \in S^{p_n}} a'(\hat{R}_n + (R_n - \hat{R}_n))a \\ &\geq \min_{a \in S^{p_n}} a'\hat{R}_n a - \max_{a \in S^{p_n}} |a'(R_n - \hat{R}_n)a| = \hat{\lambda}_{n1} - |\hat{R}_n - R_n|.\end{aligned}$$

Rewriting these inequalities, we obtain that $|\hat{\lambda}_{n1} - \lambda_{n1}| \leq |\hat{R}_n - R_n|$, $n = 2, 3, \ldots$. Because $|\hat{R}_n - R_n| = \mathrm{O}_P(p_n/n^{1/2})$ as $n \to \infty$ by Lemma A.2, the first result follows. The second result can be analogously proved. ■

Lemma A.4 *Suppose that Assumptions 1–3 hold. Then* $|R_n| = O(p_n)$ *as* $n \to \infty$*.*

Proof of Lemma A.4. Let e_n be the $p_n \times 1$ vector, all of whose elements are ones, and Γ_n the covariance matrix of $(Y_{-1}, Y_{-2}, \ldots, Y_{-p_n})$, $n = 2, 3, \ldots$. Also, let S^{m-1} denote the unit sphere in the m-dimensional Euclidean space for each $m \in \mathbb{N}$. Then

$$R_n = \begin{pmatrix} 0 & 0 \\ 0 & \Gamma_n \end{pmatrix} + \begin{pmatrix} 1 \\ \mu e_n \end{pmatrix} (1, \mu e_n').$$

It follows that

$$\begin{aligned}|R_n| &= \max_{a \in S^{p_n}} a'R_n a \leq \max_{b \in S^{p_n - 1}} b'\Gamma_n b + \max_{a \in S^{p_n}} ((1, \mu e_n')a)^2 \\ &\leq 2\pi \bar{f}_2 + |(1, \mu e_n')|^2 = 2\pi \bar{f}_2 + (1 + \mu^2 p_n) = O(p_n) \text{ as } n \to \infty.\end{aligned}$$

■

Lemma A.5 *Suppose that Assumptions 1–3 hold. Then* $\{\lambda_{1n}\}_{n=2}^{\infty}$ *is uniformly positive (hence* R_n *is nonsingular for each* $n = 2, 3, \ldots$*), and* $\{|R_n^{-1}|\}_{n=2}^{\infty}$ *is bounded.*

Proof of Lemma A.5. Let $\{\Gamma_n\}_{n=2}^{\infty}$ and $\{e_n\}_{n=2}^{\infty}$ be as in the proof of Lemma A.4. Then we have that for each $n = 2, 3, \ldots$, $\lambda_{1n} = \min_{a \in S^{p_n}} a' R_n a \geq \min_{a \in S^{p_n - 1}} a' \Gamma_n a \geq 2\pi \bar{f}_1 > 0$. Thus, $\{\lambda_{1n}\}_{n \in \mathbb{N}}$ is uniformly positive, and $\{|R_n^{-1}| = 1/\lambda_{1n}\}_{n=2}^{\infty}$ is bounded. ■

Lemma A.6 *Suppose that Assumptions 1–3 hold. Then*

$$|\hat{R}_n^{+} - R_n^{-1}| = \mathrm{O}_P(p_n/n^{1/2}) \text{ as } n \to \infty \tag{15}$$

and

$$R_n \hat{R}_n^{+} = \mathrm{O}_P(1) \text{ as } n \to \infty. \tag{16}$$

Proof of Lemma A.6. Define a sequence of random matrices $\{\Xi_n\}_{n=2}^{\infty}$ by $\Xi_n \equiv (\hat{R}_n^{+} - R_n^{-1}) - \hat{R}_n^{+}(R_n - \hat{R}_n)R_n^{-1}$. Then we have that $\hat{R}_n^{+} - R_n^{-1} = \hat{R}_n^{+}(R_n - \hat{R}_n)R_n^{-1} + \Xi_n$, so that

$$|\hat{R}_n^{+} - R_n^{-1}| \leq |\hat{R}_n^{+}|\,|R_n - \hat{R}_n|\,|R_n^{-1}| + |\Xi_n|, \quad n = 2, 3, \ldots.$$

It is straightforward to verify that Ξ_n is zero if $\hat{R}_n$ is nonsingular. It follows that for each positive real number δ,

$$P[|(p_n/n^{1/2})^{-1}\Xi_n| > \delta] \leq P[|\Xi_n| > 0] \leq P[\hat{\lambda}_{n1} = 0] \to 0,$$

where the convergence of the last term to zero follows by Lemmas A.3 and A.5. Thus, $|\Xi_n| = \mathrm{o}_P(p_n/n^{1/2})$ as $n \to \infty$. We also have that $|\hat{R}_n^{+}| = \mathrm{O}_P(1)$ and $R_n^{-1} = O(1)$ as $n \to \infty$ by Lemmas A.3 and A.5. Because $|R_n - \hat{R}_n| = \mathrm{O}_P(p_n/n^{1/2})$ as $n \to \infty$ by Lemma A.2, result (15) therefore follows.

For (16), rewrite $R_n \hat{R}_n^{+}$ as

$$\begin{aligned} R_n \hat{R}_n^{+} &= (\hat{R}_n - (\hat{R}_n - R_n))\hat{R}_n^{+} = \hat{R}_n \hat{R}_n^{+} - (\hat{R}_n - R_n)\hat{R}_n^{+} \\ &= (\hat{R}_n \hat{R}_n^{+} - I) - (\hat{R}_n - R_n)\hat{R}_n^{+} + I, \quad n = 2, 3, \ldots. \end{aligned}$$

Thus, we have that

$$|R_n \hat{R}_n^{+}| \leq |\hat{R}_n \hat{R}_n^{+} - I| + |\hat{R}_n - R_n|\,|\hat{R}_n^{+}| + |I|, \quad n = 2, 3, \ldots.$$

It is straightforward to show that the first term on the right-hand side of this inequality converges to zero prob-P (it is actually $\mathrm{o}_P(c_n)$ for any sequence $\{c_n\}_{n\in\mathbb{N}}$ of nonzero real numbers). The second term is $\mathrm{O}_P(p_n/n^{1/2})$, because $|\hat{R}_n - R_n| = \mathrm{O}_P(p_n/n^{1/2})$ (Lemma A.2) and $|\hat{R}_n^+| = \mathrm{O}_P(1)$ (Lemmas A.3 and A.5) as $n \to \infty$. The third term is one for each n. Result (16) therefore follows. ∎

Lemma A.7 *Suppose that Assumptions 1–3 hold. Then*

$$\left|\hat{\beta}_n - \beta_n - \hat{R}_n^+ (n-p_n)^{-1} \sum_{t=p_n+1}^{n} X_{nt}U_{nt}\right| = \mathrm{o}_P(c_n) \text{ as } n \to \infty, \tag{17}$$

where $\{c_n\}_{n\in\mathbb{N}}$ is an arbitrary sequence of nonzero real numbers. Also, it holds that

$$\left|\hat{\beta}_n - \beta_n - R_n^{-1} (n-p_n)^{-1} \sum_{t=p_n+1}^{n} X_{nt}U_{nt}\right| = \mathrm{o}_P(p_n/n^{1/2}) \text{ as } n \to \infty, \tag{18}$$

and

$$|\hat{\beta}_n - \beta_n| = \mathrm{O}_P(p_n/n^{1/2}) \text{ as } n \to \infty. \tag{19}$$

Proof of Lemma A.7. By definition of $\{\hat{\beta}_n\}$ and $\{U_{nt}\}$, we have that for each $n = 2, 3, \ldots,$

$$\begin{aligned}\hat{\beta}_n =& \hat{R}_n^+ (n-p_n)^{-1} \sum_{t=p_n+1}^{n} X_{nt}Y_{nt} \\ =& \hat{R}_n^+ (n-p_n)^{-1} \sum_{t=p_n+1}^{n} X_{nt}(X_{nt}'\beta_n + U_{nt}) \\ =& \hat{R}_n^+ \hat{R}_n \beta_n + \hat{R}_n^+ (n-p_n)^{-1} \sum_{t=p_n+1}^{n} X_{nt}U_{nt}.\end{aligned}$$

Therefore,

$$\hat{\beta}_n - \beta_n - \hat{R}_n^+ (n-p_n)^{-1} \sum_{t=p_n+1}^{n} X_{nt}U_{nt} = (\hat{R}_n^+\hat{R}_n - I)\beta_n, \quad n = 2, 3, \ldots.$$

It follows that

$$\left|\hat{\beta}_n - \beta_n - \hat{R}_n^+ (n-p_n)^{-1} \sum_{t=p_n+1}^{n} X_{nt}U_{nt}\right| \le |\hat{R}_n^+\hat{R}_n - I|\,|\beta_n|, \quad n = 2, 3, \ldots.$$

For any sequence $\{c_n\}_{n\in\mathbb{N}}$ of nonzero real numbers, we have that for each positive real number δ,

$$P[|c_n|^{-1}|\hat{R}_n^+\hat{R}_n - I| > \delta] \le P[|\hat{R}_n^+\hat{R}_n - I| > 0] \le P[\hat{\lambda}_{n1} = 0] \to 0$$

by Lemmas A.3 and A.5 in the Appendix. Therefore, it follows that $|\hat{R}_n^+\hat{R}_n - I| = \mathrm{o}_P(c_n)$ as $n \to \infty$.

Moreover, because $\{|\beta_n|\}_{n\in\mathbb{N}}$ is bounded by (5) of Lemma 1, the desired result (17) therefore follows.

Setting $c_n = p_n/n^{1/2}$ in (17), we have that $\left|\hat{\beta}_n - \beta_n - \hat{R}_n^+ (n-p_n)^{-1} \sum_{t=p_n+1}^n X_{nt}U_{nt}\right| = \mathrm{o}_P(p_n/n^{1/2})$ as $n\to\infty$. Given the fact that $\hat{R}_n^+ = R_n^{-1} + (\hat{R}_n^+ - R_n^{-1})$, we can obtain that

$$\left|\hat{\beta}_n - \beta_n - R_n^{-1} (n-p_n)^{-1} \sum_{t=p_n+1}^n X_{nt}U_{nt}\right| - \left|(\hat{R}_n^+ - R_n^{-1}) (n-p_n)^{-1} \sum_{t=p_n+1}^n X_{nt}U_{nt}\right| \\ \leq \left|\hat{\beta}_n - \beta_n - \hat{R}_n^+ (n-p_n)^{-1} \sum_{t=p_n+1}^n X_{nt}U_{nt}\right|.$$

The second term on the left-hand side is dominated by $|\hat{R}_n^+ - R_n^{-1}| \left|(n-p_n)^{-1} \sum_{t=p_n+1}^n X_{nt}U_{nt}\right| = \mathrm{O}_P(p_n^2/n)$ as $n\to\infty$, where the convergence rate is given by (12) of Lemma A.2 and (15) of Lemma A.6 in the Appendix. Result (18) therefore follows.

To prove (19), note that $\left|R_n^{-1} (n-p_n)^{-1} \sum_{t=p_n+1}^n X_{nt}U_{nt}\right| \leq |R_n^{-1}| \left|(n-p_n)^{-1} \sum_{t=p_n+1}^n X_{nt}U_{nt}\right| = \mathrm{O}_P(p_n/n^{1/2})$ as $n\to\infty$ by (12) of Lemma A.2 and Lemma A.5 in the Appendix. Because

$$\left||\hat{\beta}_n - \beta_n| - \left|R_n^{-1} (n-p_n)^{-1} \sum_{t=p_n+1}^n X_{nt}U_{nt}\right|\right| \\ \leq \left|\hat{\beta}_n - \beta_n - R_n^{-1} (n-p_n)^{-1} \sum_{t=p_n+1}^n X_{nt}U_{nt}\right|, \quad n = 2,3,\ldots,$$

where the second term on the left-hand side is $\mathrm{O}_P(p_n/n^{1/2})$ as shown above and the right-hand side is $\mathrm{o}_P(p_n/n^{1/2})$, result (19) follows immediately. ■

Corollary A.8 *Suppose that Assumptions 1–3 hold. Then for each* $k = 0,1,2,\ldots,$

$$\left|(\hat{\beta}_n - \beta_n)' \left((n-p_n-k)^{-1} \sum_{t=p_n+k+1}^n X_{n,t-k}X_{n,t-k}'\right) (\hat{\beta}_n - \beta_n)\right| = \mathrm{O}_P(p_n^2/n) \text{ as } n\to\infty.$$

Proof of Corollary A.8. Because

$$\begin{aligned}
&(\hat{\beta}_n - \beta_n)' \left((n-p_n-k)^{-1} \sum_{t=p_n+k+1}^n X_{n,t-k}X_{n,t-k}'\right) (\hat{\beta}_n - \beta_n) \\
&= (\hat{\beta}_n - \beta_n)' \left(\left((n-p_n-k)^{-1} \sum_{t=p_n+k+1}^n X_{n,t-k}X_{n,t-k}' - R_n\right) + R_n\right) (\hat{\beta}_n - \beta_n) \\
&= (\hat{\beta}_n - \beta_n)' \left((n-p_n-k)^{-1} \sum_{t=p_n+k+1}^n X_{n,t-k}X_{n,t-k}' - R_n\right) (\hat{\beta}_n - \beta_n) \\
&\quad + (\hat{\beta}_n - \beta_n)' R_n (\hat{\beta}_n - \beta_n), \quad n = \bar{n}_k, \bar{n}_k+1,\ldots,
\end{aligned}$$

we have that

$$
\begin{aligned}
&\left|(\hat{\beta}_n-\beta_n)'\left((n-p_n-k)^{-1}\sum_{t=p_n+k+1}^{n}X_{n,t-k}X'_{n,t-k}\right)(\hat{\beta}_n-\beta_n)\right| \\
&\qquad\le|\hat{\beta}_n-\beta_n|^2\left|(n-p_n-k)^{-1}\sum_{t=p_n+k+1}^{n}X_{n,t-k}X'_{n,t-k}-R_n\right| \\
&\qquad\qquad+\left|(\hat{\beta}_n-\beta_n)'R_n(\hat{\beta}_n-\beta_n)\right|,\quad n=\bar{n}_k,\bar{n}_k+1,\ldots.
\end{aligned}
$$

By (11) of Lemma A.2 and (19) of Lemma A.7, the first term on the right-hand side of this inequality is $\mathrm{O}_P(p_n^3/n^{3/2})$. It thus suffices to show that the second term is $\mathrm{O}_P(p_n^2/n)$. Define

$$
\xi_n\equiv\hat{\beta}_n-\beta_n-\hat{R}_n^+\,(n-p_n)^{-1}\sum_{t=p_n+1}^{n}X_{nt}U_{nt},\quad n=2,3,\ldots.
$$

Then $|\xi_n|=\mathrm{o}_P(1/n^{1/2})$ as $n\to\infty$ by (17) of Lemma A.7. Using this $\{\xi_n\}$, we have that $\hat{\beta}_n-\beta_n=\hat{R}_n^+\,(n-p_n)^{-1}\sum_{t=p_n+1}^{n}X_{nt}U_{nt}+\xi_n$, $n=2,3,\ldots$. It follows that

$$
\begin{aligned}
&(\hat{\beta}_n-\beta_n)'R_n(\hat{\beta}_n-\beta_n) \\
&\qquad=\left((n-p_n)^{-1}\sum_{t=p_n+1}^{n}X_{nt}U_{nt}\right)'\hat{R}_n^+R_n\hat{R}_n^+\left((n-p_n)^{-1}\sum_{t=p_n+1}^{n}X_{nt}U_{nt}\right) \\
&\qquad\qquad+2\xi_n'R_n\hat{R}_n^+\left((n-p_n)^{-1}\sum_{t=p_n+1}^{n}X_{nt}U_{nt}\right)+\xi_n'R_n\xi_n,\quad n=2,3,\ldots.
\end{aligned}
$$

We thus have that

$$
\begin{aligned}
&\left|(\hat{\beta}_n-\beta_n)'R_n(\hat{\beta}_n-\beta_n)\right|\le|\hat{R}_n^+|\,|R_n\hat{R}_n^+|\left|(n-p_n)^{-1}\sum_{t=p_n+1}^{n}X_{nt}U_{nt}\right|^2 \\
&\qquad\qquad+2|\xi_n|\,|R_n\hat{R}_n^+|\left|(n-p_n)^{-1}\sum_{t=p_n+1}^{n}X_{nt}U_{nt}\right|+|\xi_n|^2\,|R_n|,\quad n=2,3,\ldots.
\end{aligned}
$$

By applying Lemmas A.2, A.3, A.4, A.5, and A.6, we can verify that the first term on the right-hand side of the above inequality is $\mathrm{O}_P(p_n^2/n)$, the second term is $\mathrm{o}_P(p_n/n)$, and the third term is $\mathrm{o}_P(p_n/n)$ as $n\to\infty$. The desired result therefore follows. ■

Lemma A.9 *Suppose that Assumptions 1–3 hold. Then*

$$
(n-p_n)^{-1}\sum_{t=p_n+1}^{n}(\hat{\epsilon}_{nt}-U_{nt})^2=\mathrm{O}_P(p_n^2/n)\ \textit{as}\ n\to\infty.
$$

Proof of Lemma A.9. Because $\hat{\epsilon}_{nt} - U_{nt} = -X'_{nt}(\hat{\beta}_n - \beta_n)$, $t \in \mathbb{Z}$, $n = 2, 3, \ldots$, we have that for each $n = 2, 3, \ldots$, $(\hat{\epsilon}_{nt} - U_{nt})^2 = (\hat{\beta}_n - \beta_n)' X_{nt} X'_{nt} (\hat{\beta}_n - \beta_n)$, $t = p_n + 1, \ldots, n$, so that

$$\left| (n - p_n)^{-1} \sum_{t=p_n+1}^{n} (\hat{\epsilon}_{nt} - U_{nt})^2 \right| = \left| (\hat{\beta}_n - \beta_n)' \left((n - p_n)^{-1} \sum_{t=p_n+1} X_{nt} X'_{nt} \right) (\hat{\beta}_n - \beta_n) \right|,$$

$n = 2, 3, \ldots$. Since the right-hand side of the above equality is $\mathrm{O}_P(p_n^2/n)$ as $n \to \infty$ by Corollary A.8, the desired result therefore follows. ■

Lemma A.10 *Suppose that Assumptions 1–3 hold. Also, let* $\{\{\ddot{\epsilon}_{nt}\}_{t=p_n+1}^{n}\}_{n=2}^{\infty}$ *be a double array of random variables such that*

$$\ddot{\eta}_n \equiv (n - p_n) \sum_{t=p_n+1}^{n} (\ddot{\epsilon}_{nt} - U_{nt})^2 \to 0 \text{ as } n \to \infty \text{ prob-}P.$$

For each $k \in \mathbb{N}$, *define a sequence of random variables* $\{\ddot{\psi}_{nk}\}_{n=2}^{\infty}$ *by*

$$\ddot{\psi}_{nk} \equiv \begin{cases} \left((n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} \ddot{\epsilon}_{n,t-k}^2 \right)^+ & \\ \qquad \times (n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} \ddot{\epsilon}_{n,t-k} Y_t & \text{if } k \le n - p_n - 1, \\ 0 & \text{otherwise,} \end{cases}$$

$n = 2, 3, \ldots$. *Then for each* $k \in \mathbb{N}$,

$$(n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} \ddot{\epsilon}_{n,t-k} Y_t - \mathrm{E}[U_{n,-k} Y_0] = \mathrm{O}_P(p_n^{1/2}/n^{1/2}) + \mathrm{O}_P(\ddot{\eta}_n^{1/2}) \text{ as } n \to \infty, \tag{20}$$

$$(n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} \ddot{\epsilon}_{n,t-k}^2 - \mathrm{E}[U_{n,0}^2] = \mathrm{O}_P(p_n/n^{1/2}) + \mathrm{O}_P(\ddot{\eta}_n^{1/2}) \text{ as } n \to \infty, \tag{21}$$

and

$$\ddot{\psi}_{nk} - \bar{\psi}_{nk} = \mathrm{O}_P(p_n/n^{1/2}) + \mathrm{O}_P(\ddot{\eta}_n^{1/2}) \text{ as } n \to \infty. \tag{22}$$

Proof of Lemma A.10. Let k be an arbitrary natural number. For convenience, write

$$\ddot{C}_n \equiv (n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} \ddot{\epsilon}_{n,t-k} Y_t, \quad n = \bar{n}_k, \bar{n}_k + 1, \ldots,$$

$$C_n \equiv \mathrm{E}[U_{n,-k} Y_0], \quad n = 2, 3, \ldots,$$

$$\ddot{D}_n \equiv (n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} \ddot{\epsilon}_{n,t-k}^2, \quad n = \bar{n}_k, \bar{n}_k + 1, \ldots,$$

and

$$D_n \equiv \mathrm{E}[U_{n0}^2], \quad n = 2, 3, \ldots.$$

Then for each $n = \bar{n}_k, \bar{n}_k + 1, \ldots$, $\ddot{\psi}_{nk} = \ddot{C}_n / \ddot{D}_n$, and $\bar{\psi}_{nk} = C_n / D_n$. Because for each $n = \bar{n}_k, \bar{n}_k + 1, \ldots$,

$$\begin{aligned}\ddot{C}_n - C_n =& (n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} (\ddot{\epsilon}_{n,t-k} - U_{n,t-k}) Y_t \\ &+ (n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} U_{n,t-k} Y_t - \mathrm{E}[U_{n,-k} Y_0],\end{aligned}$$

we have that

$$\begin{aligned}|\ddot{C}_n - C_n| \leq& \left| (n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} (\ddot{\epsilon}_{n,t-k} - U_{n,t-k}) Y_t \right| \\ &+ \left| (n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} U_{n,t-k} Y_t - \mathrm{E}[U_{n,-k} Y_0] \right| \\ \leq& \left((n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} (\ddot{\epsilon}_{n,t-k} - U_{n,t-k})^2 \right)^{1/2} \left((n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} Y_t^2 \right)^{1/2} \\ &+ \left| (n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} U_{n,t-k} Y_t - \mathrm{E}[U_{n,-k} Y_0] \right|, \quad n = \bar{n}_k, \bar{n}_k + 1, \ldots. \end{aligned} \tag{23}$$

By (8) of Lemma A.1, we have that $\mathrm{var}\left[(n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} Y_t^2\right] \leq (n - p_n - k)^{-1} A^2 \to 0$ as $n \to \infty$. Moreover, it holds that $\mathrm{E}\left[(n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} Y_t^2\right] = \mathrm{E}[Y_0^2]$. It follows by the Chebyshev inequality that

$$(n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} Y_t^2 \to \mathrm{E}[Y_0^2] \text{ as } n \to \infty \text{ prob-}P.$$

Therefore, the first term on the right-hand side of (23) is $\mathrm{O}_P(\ddot{\eta}_n^{1/2})$ as $n \to \infty$. Furthermore, the second term is $\mathrm{O}_P(p_n^{1/2}/n^{1/2})$ as shown in (14) of Lemma A.2. Thus, we have that $\ddot{C}_n - C_n = \mathrm{O}_P(p_n^{1/2}/n^{1/2}) + \mathrm{O}_P(\ddot{\eta}_n^{1/2})$ as $n \to \infty$, and (20) holds.

Next, we examine $\ddot{D}_n - D_n$. Because for each $n = \bar{n}_k, \bar{n}_k + 1, \ldots$,

$$\begin{aligned}\ddot{D}_n - D_n =& (n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} (\ddot{\epsilon}_{n,t-k}^2 - U_{n,t-k}^2) + (n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} U_{n,t-k}^2 - \mathrm{E}[U_{n,-k}^2] \\ =& (n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} (\ddot{\epsilon}_{n,t-k} - U_{n,t-k})(\ddot{\epsilon}_{n,t-k} + U_{n,t-k}) \\ &+ (n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} U_{n,t-k}^2 - \mathrm{E}[U_{n,-k}^2],\end{aligned}$$

we have that for each $n = \bar{n}_k, \bar{n}_k + 1, \ldots$,

$$|\ddot{D}_n - D_n| \le \left|(n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} (\ddot{\epsilon}_{n,t-k} - U_{n,t-k})(\ddot{\epsilon}_{n,t-k} + U_{n,t-k})\right| + \left|(n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} U_{n,t-k}^2 - \mathrm{E}[U_{n,-k}^2]\right|.$$

The second term on the right-hand side of this inequality is $\mathrm{O}_P(p_n/n^{1/2})$ by (13) of Lemma A.2. For the first term, applying the Cauchy-Schwartz inequality and the Minkowsky inequality, we have that for each $n = \bar{n}_k, \bar{n}_k + 1, \ldots$,

$$\begin{aligned}
&\left|(n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} (\ddot{\epsilon}_{n,t-k} - U_{n,t-k})(\ddot{\epsilon}_{n,t-k} + U_{n,t-k})\right| \\
&\le \left((n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} (\ddot{\epsilon}_{n,t-k} - U_{n,t-k})^2\right)^{1/2} \left((n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} (\ddot{\epsilon}_{n,t-k} + U_{n,t-k})^2\right)^{1/2} \\
&= \left((n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} (\ddot{\epsilon}_{n,t-k} - U_{n,t-k})^2\right)^{1/2} \\
&\quad \times \left((n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} ((\ddot{\epsilon}_{n,t-k} - U_{n,t-k}) + 2U_{n,t-k})^2\right)^{1/2} \\
&\le \left((n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} (\ddot{\epsilon}_{n,t-k} - U_{n,t-k})^2\right)^{1/2} \\
&\quad \times \left(\left((n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} (\ddot{\epsilon}_{n,t-k} - U_{n,t-k})^2\right)^{1/2} + 2\left((n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} U_{n,t-k}^2\right)^{1/2}\right).
\end{aligned}$$

The second term in the second factor on the right-hand side of this inequality is $\mathrm{O}_P(1)$ by Lemma 1 and (13) of Lemma A.2. For the first factor on the right-hand side and the first term in the second factor, we have that

$$\begin{aligned}
\left((n - p_n - k)^{-1} \sum_{t=p_n+k+1}^{n} (\ddot{\epsilon}_{n,t-k} - U_{n,t-k})^2\right)^{1/2} &\le \left((n - p_n - k)^{-1} \sum_{t=p_n+1}^{n} (\ddot{\epsilon}_{n,t} - U_{n,t})^2\right)^{1/2} \\
&= \left(\frac{n - p_n}{n - p_n - k}\right)^{1/2} \left((n - p_n)^{-1} \sum_{t=p_n+1}^{n} (\ddot{\epsilon}_{n,t} - U_{n,t})^2\right)^{1/2}, \quad n = \bar{n}_k, \bar{n}_k + 1, \ldots,
\end{aligned}$$

where the right-hand side is $\mathrm{O}_P(\ddot{\eta}_n^{1/2})$ by assumption. It follows that

$$\left|(n-p_n-k)^{-1}\sum_{t=p_n+k+1}^{n}(\ddot{\epsilon}_{n,t-k}-U_{n,t-k})(\ddot{\epsilon}_{n,t-k}+U_{n,t-k})\right|=\mathrm{O}_P(\ddot{\eta}_n^{1/2}) \text{ as } n\to\infty.$$

Therefore, $\ddot{D}_n - D_n = \mathrm{O}_P(p_n/n^{1/2}) + \mathrm{O}_P(\ddot{\eta}_n^{1/2})$ as $n \to \infty$, and (21) holds.

We now apply the standard linearization method to (21) and (13) of Lemma A.2 to obtain that

$$\ddot{D}_n^{+} - D_n^{-1} = -D_n^{-2}\left(\ddot{D}_n - D_n\right) + \mathrm{o}_P(p_n/n^{1/2}) + \mathrm{o}_P(\ddot{\eta}_n^{1/2}) = \mathrm{O}_P(p_n/n^{1/2}) + \mathrm{O}_P(\ddot{\eta}_n^{1/2}) \text{ as } n\to\infty.$$

Using $\ddot{C}_n$, C_n, $\ddot{D}_n$, and D_n, we can write $\ddot{\psi}_{nk} - \bar{\psi}_{nk}$ as

$$\begin{aligned}\ddot{\psi}_{nk} - \bar{\psi}_{nk} =& \ddot{D}_n^{+}\ddot{C}_n - D_n^{-1}C_n\\ =& C_n(\ddot{D}_n^{+} - D_n^{-1}) + D_n^{-1}(\ddot{C}_n - C_n) + (\ddot{D}_n^{+} - D_n^{-1})(\ddot{C}_n - C_n).\end{aligned}$$

By applying the rate of convergence stated above to each term on the right-hand side of the above equality, we can derive the desired result (22). ■

REFERENCES

Andrews, D. W. K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

Bhansali, R. J. (1976), "Estimation of the Moving Average Representation of a Stationary Nondeterministic Process," *Biometrika*, 63, 408–410.

Brockwell, P. J. and Davis, R. A. (1991), *Time Series: Theory and Methods*, Springer Series in Statistics, New York, Berlin, London and Tokyo: Springer-Verlag, 2nd ed.

Magnus, J. R. and Neudecker, H. (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, New York, Chichester, Toronto and Brisbane: Wiley.

Murray, C. J. and Papell, D. H. (2002), "The Purchasing Power Parity Persistence Paradigm," *Journal of International Economics*, 56, 1–19.

Newey, W. K. and West, K. D. (1994), "Automatic Lag Selection in Covariance Matrix Estimation," *The Review of Economic Studies*, 61, 631–653.

Sims, C. A. (1971), "Distributed Lag Estimation When the Parameter Space is Explicitly Infinite-Dimensional," *Annals of Mathematical Statistics*, 42.

— (1972), "The Role of Approximate Prior Restrictions in Distributed Lag Estimation," *Journal of the American Statistical Association*, 67, 169–175.

White, H. (1994), *Estimation, Inference and Specification Analysis*, Econometric Society Monographs, no. 22, Cambridge, New York and Melbourne: Cambridge University Press.

Wright, J. H. (1999), "Frequency Domain Inference for Univariate Impulse Responses," *Economics Letters*, 63, 269–277.

Table 1: The Estimated AR(1) Model for the French Real Exchange Rate

Parameter	Estimate	(Std. Err.)
α	0.0375	(0.0274)
ϕ_1	0.9791	(0.0156)

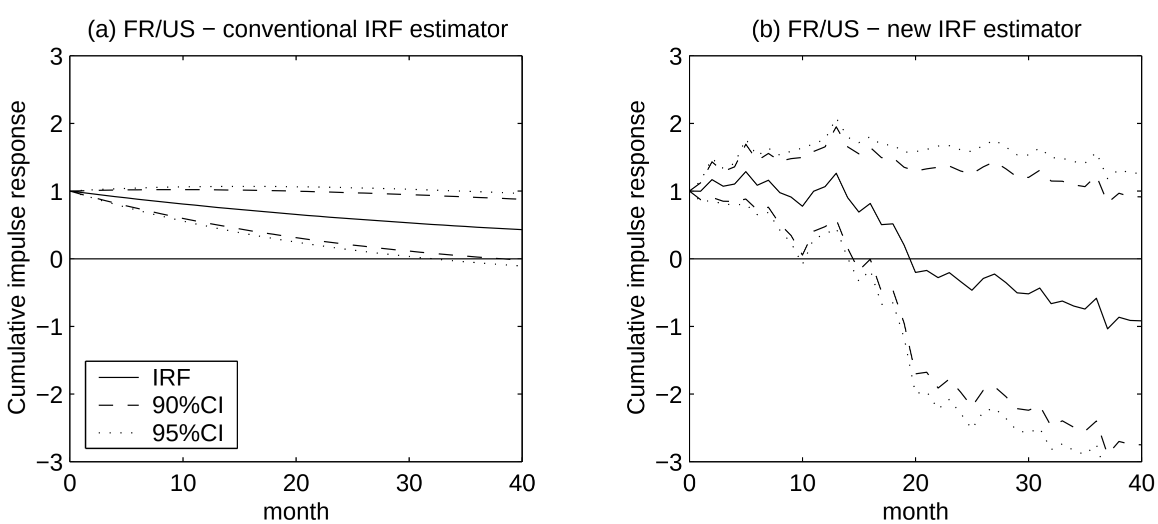

Figure 1: The IRF Estimates of the French Real Exchange Rate

Table 2: The IRF Estimates of the French Real Exchange Rate

k	ψ_{nk}^{conv}	(std)	$\tilde{\psi}_{nk}$	(std)
0	1.0000	(0.0000)	1.0000	(0.0000)
1	0.9791	(0.0156)	0.9972	(0.0758)
2	0.9585	(0.0306)	1.1697	(0.1601)
3	0.9385	(0.0449)	1.0709	(0.1338)
4	0.9188	(0.0586)	1.1025	(0.1556)
5	0.8996	(0.0717)	1.2873	(0.2460)
6	0.8807	(0.0842)	1.0863	(0.2211)
7	0.8623	(0.0962)	1.1595	(0.2418)
8	0.8442	(0.1077)	0.9775	(0.2820)
9	0.8265	(0.1186)	0.9110	(0.3453)
10	0.8092	(0.1290)	0.7760	(0.4391)
11	0.7923	(0.1389)	0.9963	(0.3578)
12	0.7757	(0.1484)	1.0642	(0.3579)
13	0.7594	(0.1574)	1.2649	(0.4158)
14	0.7435	(0.1659)	0.9036	(0.4554)
15	0.7279	(0.1741)	0.6891	(0.5232)
16	0.7127	(0.1818)	0.8186	(0.5024)
17	0.6978	(0.1891)	0.5043	(0.6019)
18	0.6831	(0.1960)	0.5158	(0.5955)
19	0.6688	(0.2026)	0.2010	(0.6997)
20	0.6548	(0.2088)	-0.2035	(0.9112)
21	0.6411	(0.2146)	-0.1736	(0.9133)
22	0.6277	(0.2201)	-0.2794	(0.9920)
23	0.6145	(0.2253)	-0.2045	(0.9572)
24	0.6017	(0.2302)	-0.3378	(0.9940)
25	0.5891	(0.2348)	-0.4660	(1.0458)
26	0.5767	(0.2390)	-0.2924	(1.0036)
27	0.5646	(0.2430)	-0.2272	(1.0103)
28	0.5528	(0.2468)	-0.3545	(1.0229)
29	0.5412	(0.2502)	-0.5043	(1.0409)
30	0.5299	(0.2534)	-0.5182	(1.0461)
31	0.5188	(0.2564)	-0.4328	(1.0574)
32	0.5079	(0.2591)	-0.6639	(1.1007)
33	0.4973	(0.2616)	-0.6255	(1.0769)
34	0.4869	(0.2639)	-0.6985	(1.0883)
35	0.4767	(0.2660)	-0.7422	(1.0987)
36	0.4667	(0.2678)	-0.5848	(1.1018)
37	0.4569	(0.2695)	-1.0347	(1.1314)
38	0.4473	(0.2710)	-0.8657	(1.1150)
39	0.4380	(0.2723)	-0.9127	(1.1119)
40	0.4288	(0.2734)	-0.9186	(1.1140)

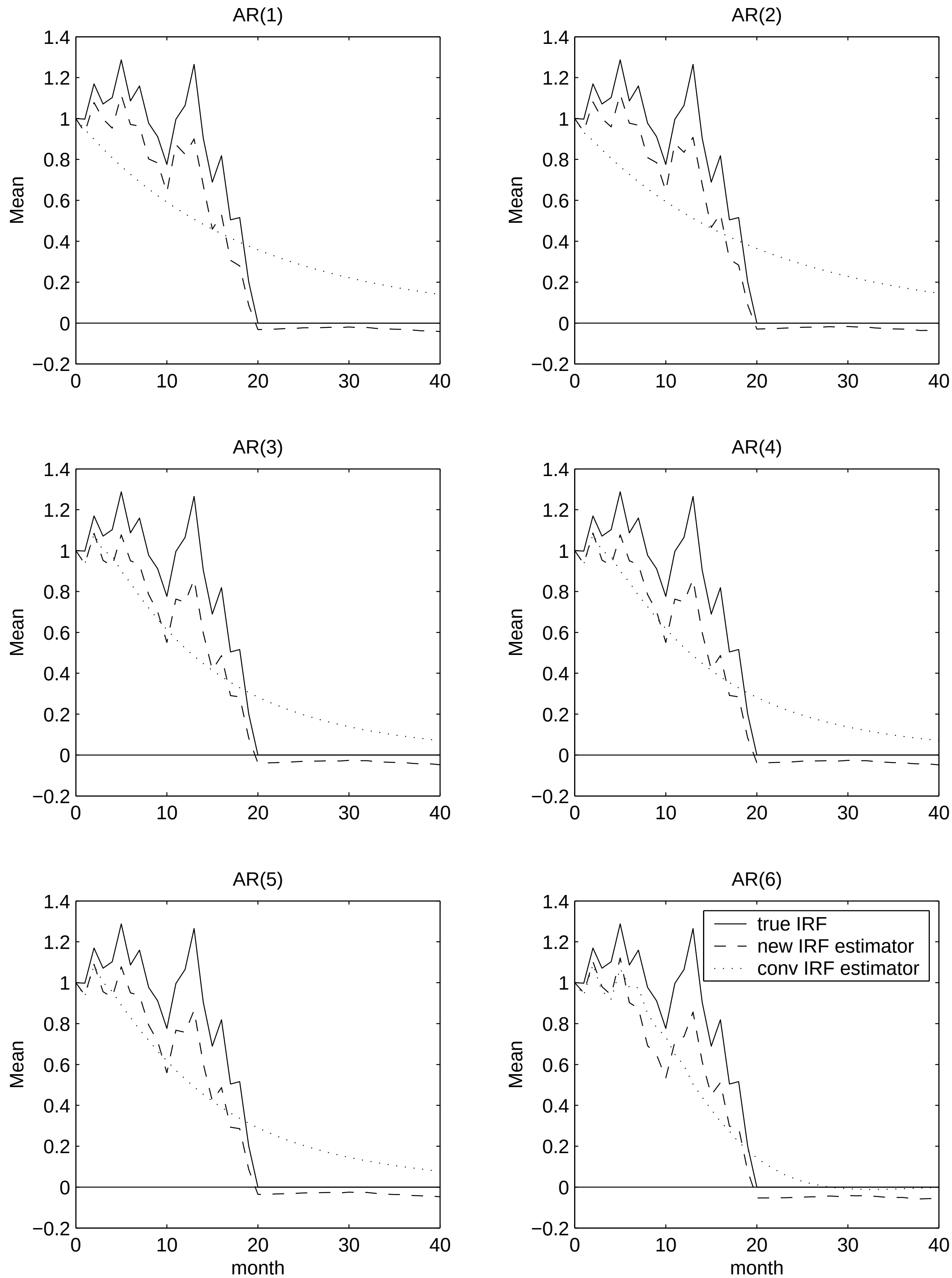

Figure 2: Monte Carlo Simulation I - Mean of the IRF Estimators

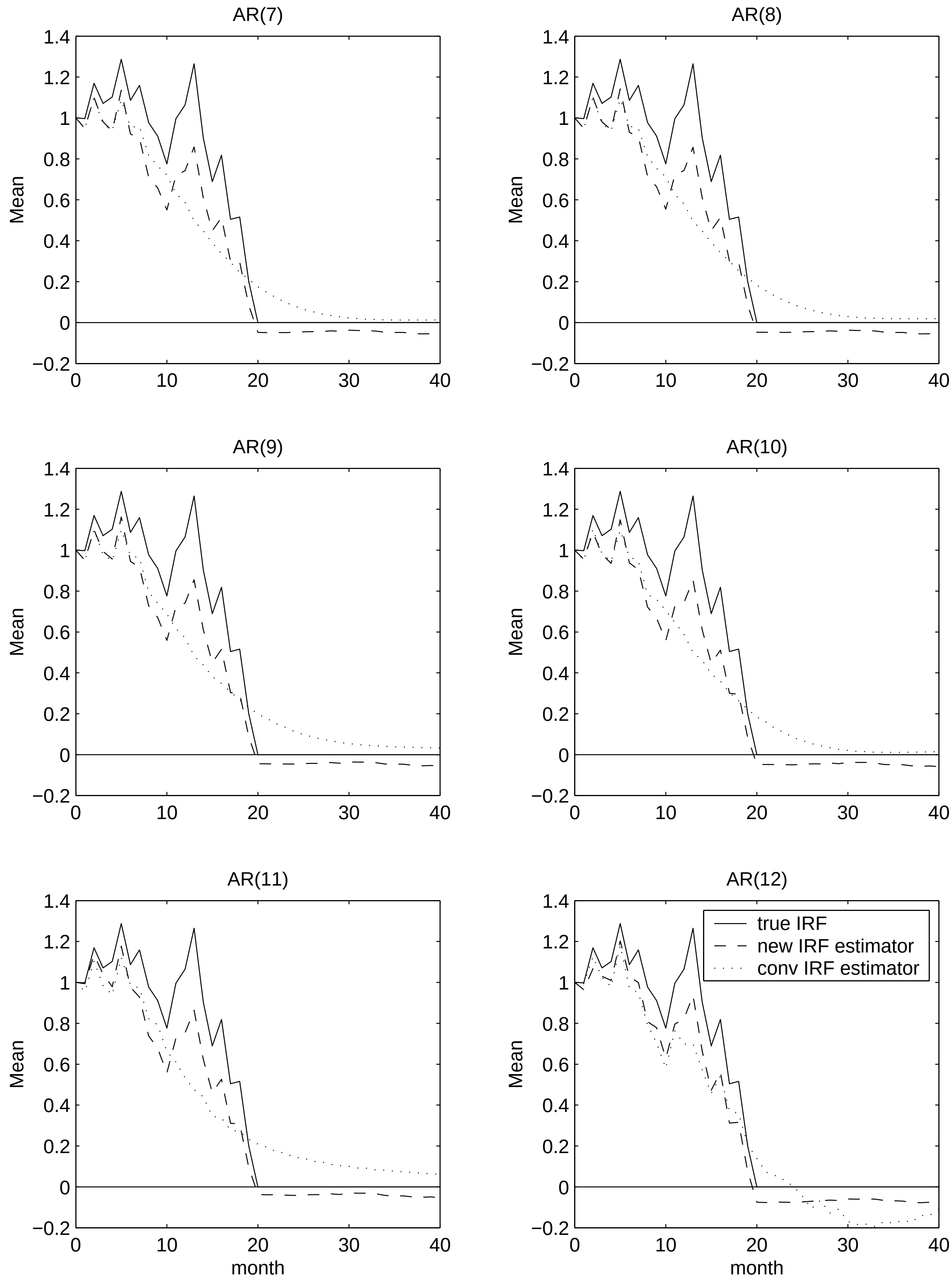

Figure 2: Monte Carlo Simulation I - Mean of the IRF Estimators (continued)

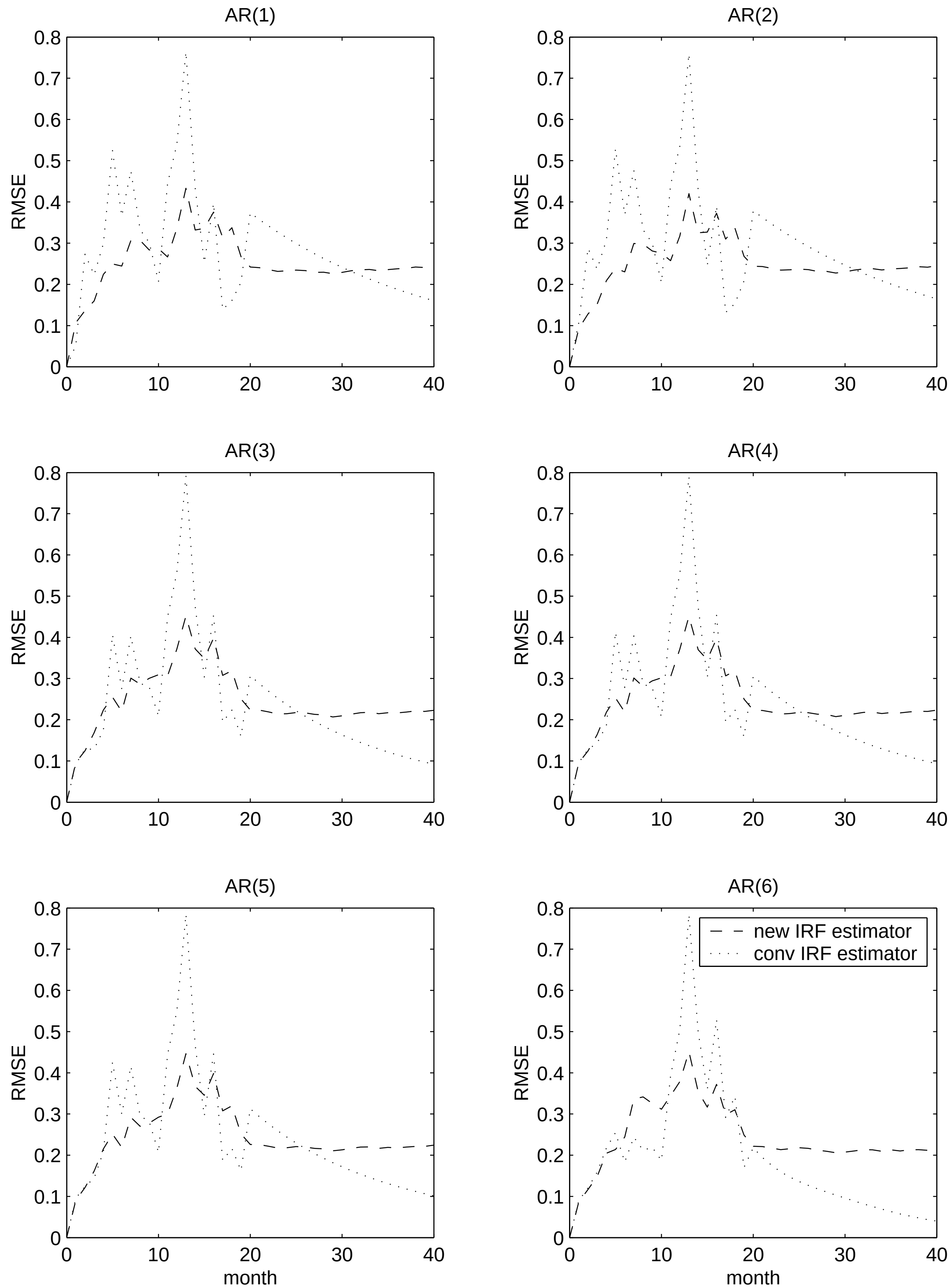

Figure 3: Monte Carlo Simulation I - RMSE of the IRF Estimators

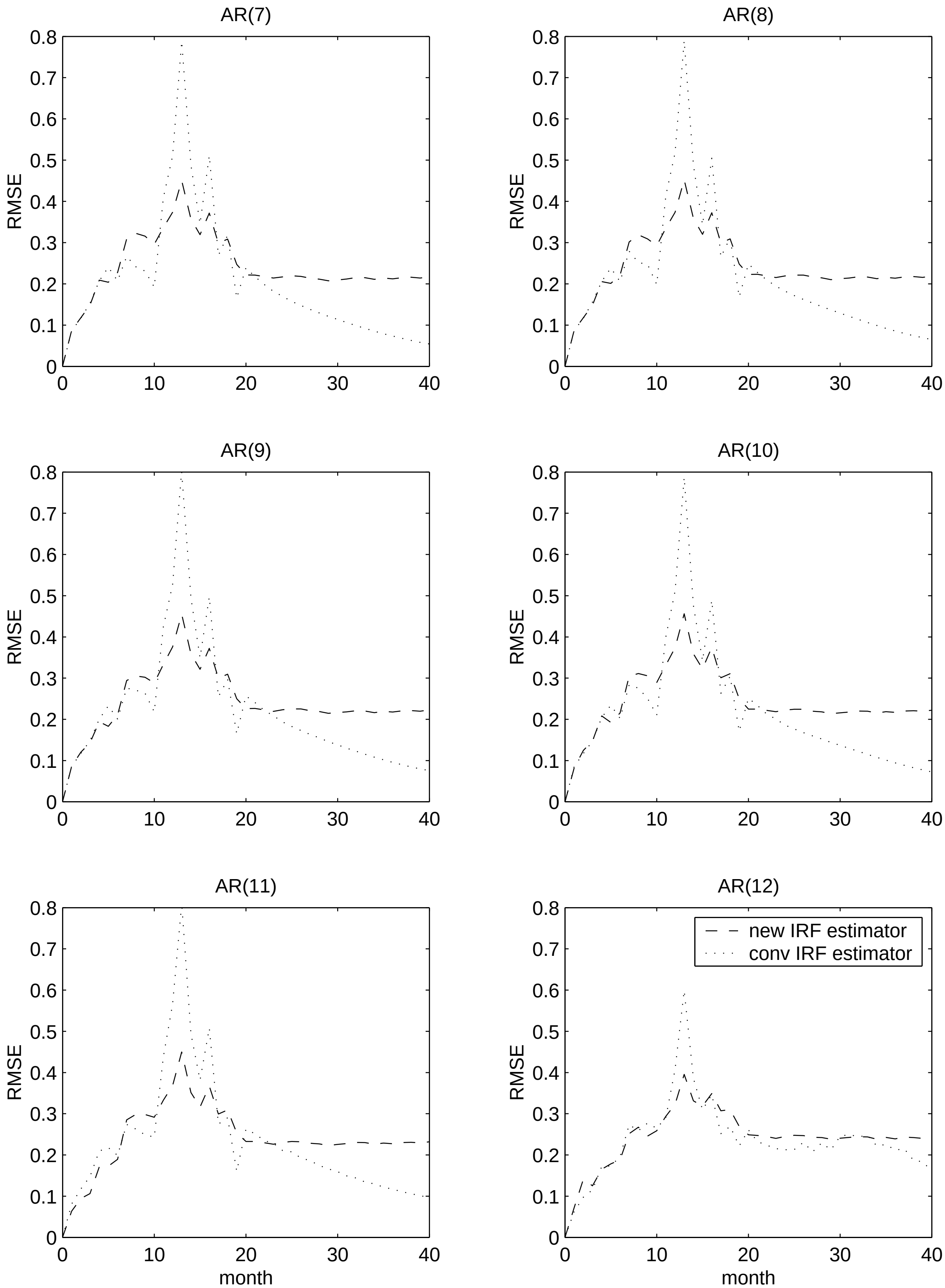

Figure 3: Monte Carlo Simulation I - RMSE of the IRF Estimators (continued)

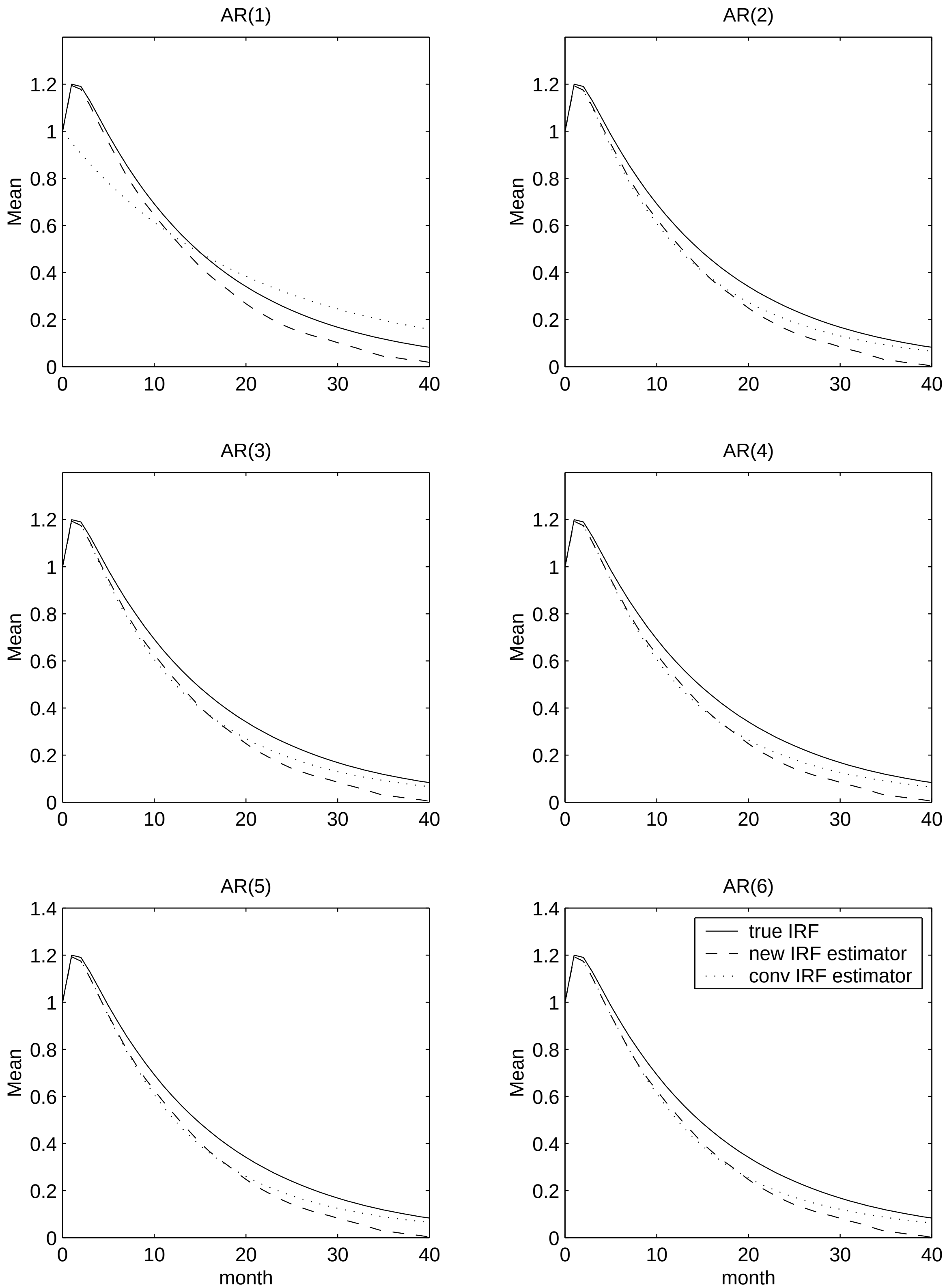

Figure 4: Monte Carlo Simulation II - Mean of the IRF Estimators

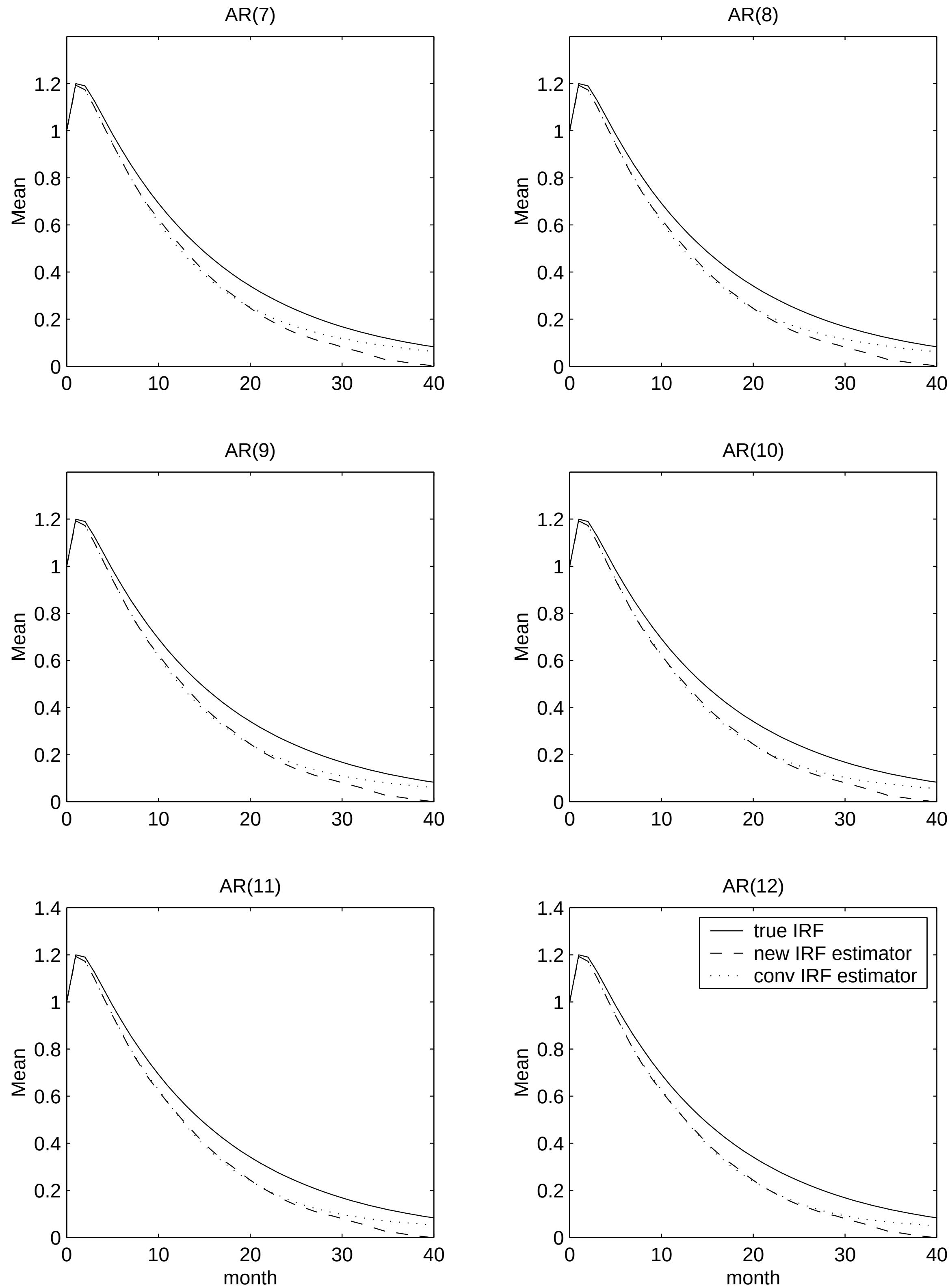

Figure 4: Monte Carlo Simulation II - Mean of the IRF Estimators (continued)

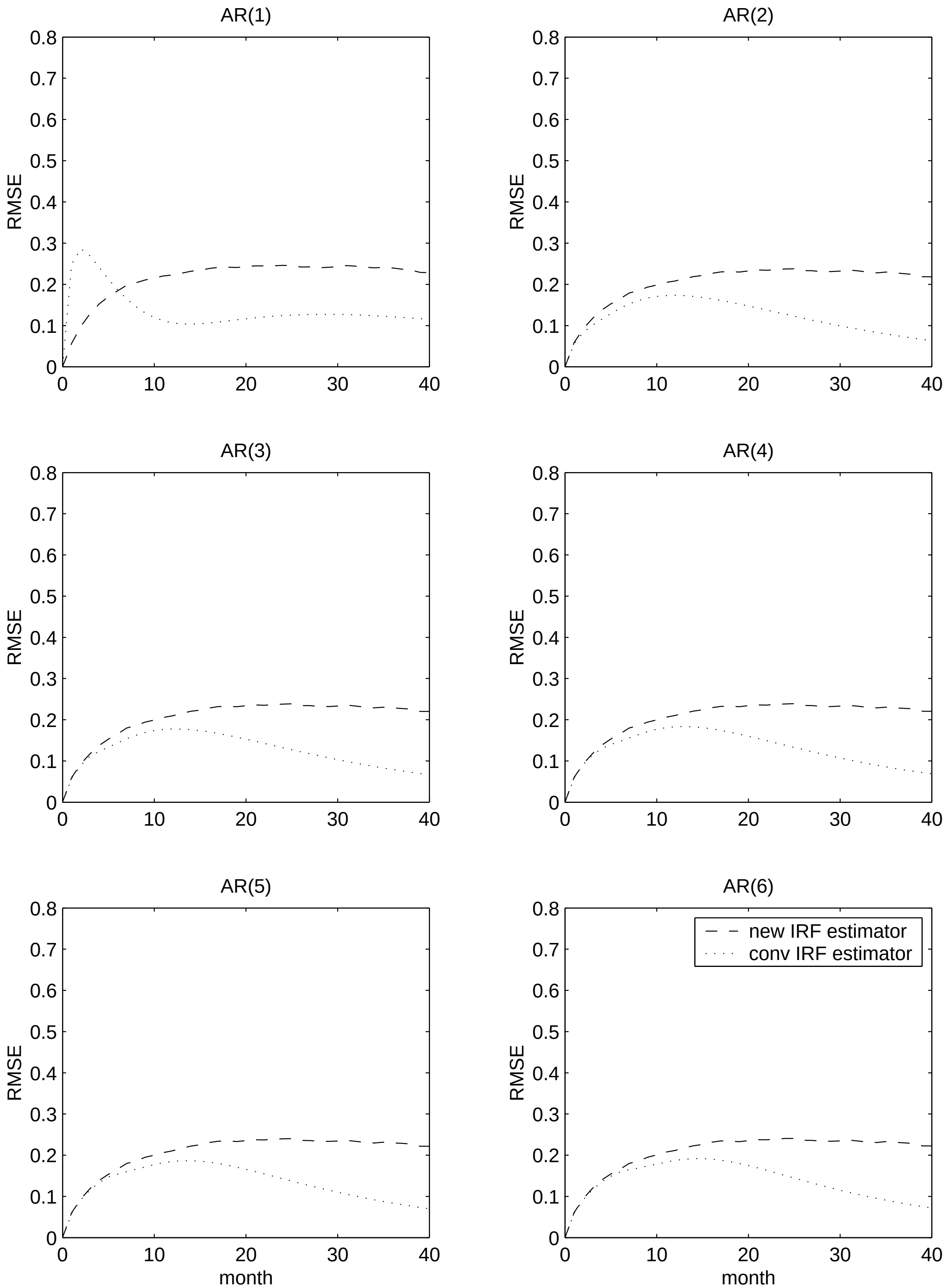

Figure 5: Monte Carlo Simulation II - RMSE of the IRF Estimators

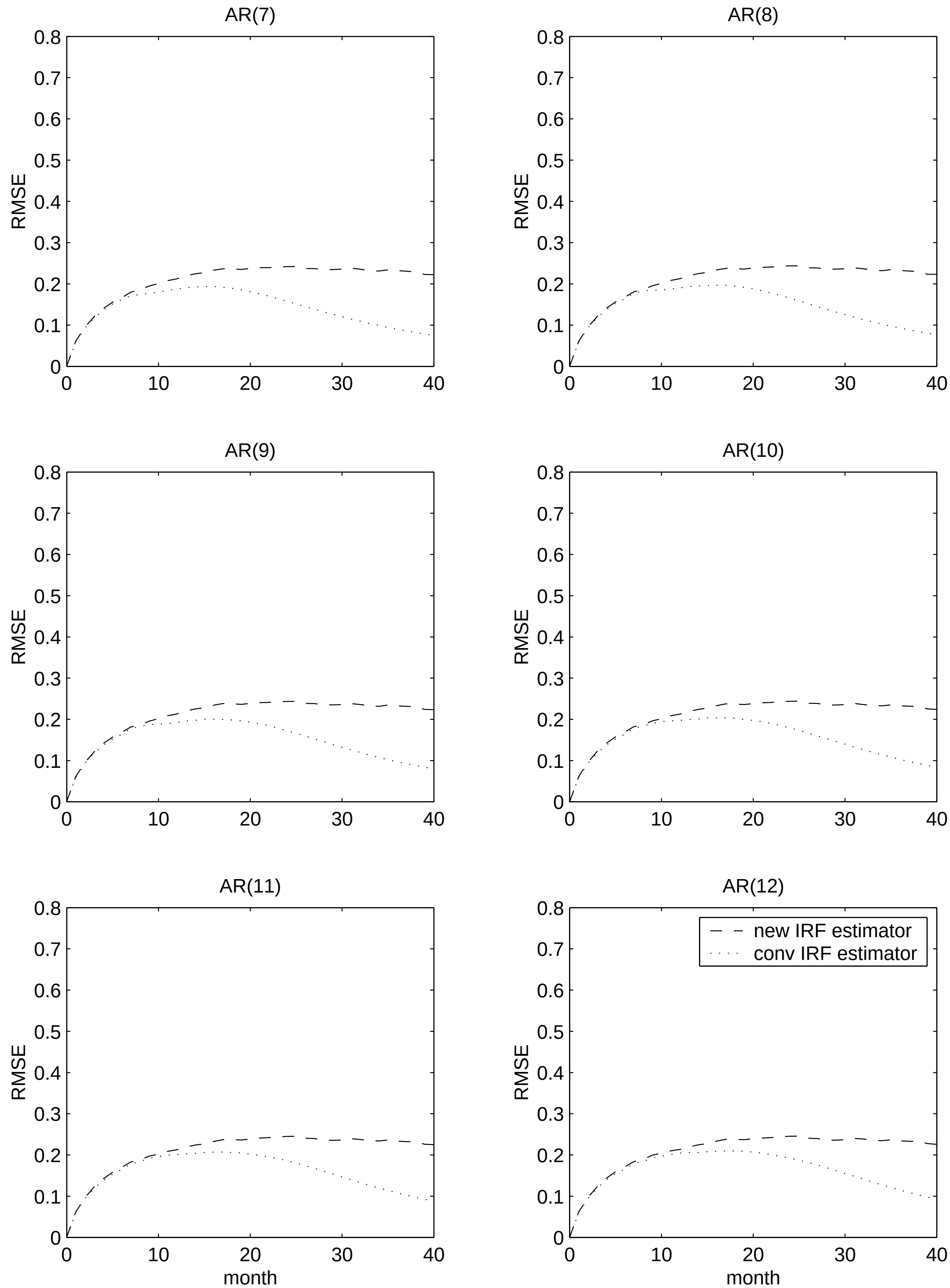

Figure 5: Monte Carlo Simulation II - RMSE of the IRF Estimators (continued)