Identifiability and Estimation of Autoregressive ARCH Models with Measurement Error

Mustafa Salamh and Liqun Wang

Abstract The autoregressive conditional heteroscedasticity (ARCH) model and its various generalizations have been widely used to analyze economic and financial data. Although many variables like GDP, inflation, and commodity prices are imprecisely measured, research focusing on the mismeasured response processes in GARCH models is sparse. We study a dynamic model with ARCH error where the underlying process is latent and subject to additive measurement error. We show that, in contrast to the case of covariate measurement error, this model is identifiable by using the observations of the proxy process only and no extra information is needed. We construct GMM estimators for the unknown parameters which are consistent and asymptotically normally distributed under general conditions. We also propose a procedure to test the presence of measurement error, which avoids the usual boundary problem of testing variance parameters. We carry out Monte Carlo simulations to study the impact of measurement error on the naive maximum likelihood estimators and have found interesting patterns of their biases. Moreover, the proposed estimators have fairly good finite sample properties.

Keywords Dynamic ARCH model · Errors in variables · Generalized method of moments · Measurement error · Semiparametric estimation

The original version of the chapter has been revised. A correction to this chapter can be found at https://doi.org/10.1007/978-3-031-08329-7_16

M. Salamh
Department of Statistics, Cairo University, Giza, Egypt
e-mail: Mustafa.Salamh@feps.edu.eg

L. Wang (✉)
Department of Statistics, University of Manitoba, Winnipeg, MB, Canada
e-mail: Liqun.Wang@umanitoba.ca

W. He et al. (eds.), *Advances and Innovations in Statistics and Data Science*, ICSA Book Series in Statistics, https://doi.org/10.1007/978-3-031-08329-7_11

1 Introduction

Since the seminal works of Engle (1982) and Bollerslev (1986), the autoregressive conditional heteroscedasticity (ARCH) model and its various generalizations have been widely used to analyze economic and financial data, such as GDP, inflation, stock prices, and interest rates, see, e.g., Grier and Perry (2000), Engle et al. (2008), Fang and Miller (2009), Teräsvirta (2009), Francq and Zakoian (2011), and Caporale et al. (2012). Moreover, there is also a large number of empirical studies of agricultural and industrial commodity prices using ARCH/GARCH models, e.g., Ramirez and Fadiga (2003), Roche and McQuinn (2003), and Reitz and Westerhoff (2007). However, it is well documented in the literature that many economic variables including GDP, inflation and commodity prices are imprecisely measured. For example, Wansbeek and Meijer (2000) and Buonaccorsi (2013) provide broad surveys on the issues of measurement errors and their impacts in econometric models. In particular, Alberini and Filippini (2011) emphasize that the US energy prices are mismeasured, while Fan and Wang (2007) point out that high-frequency financial data are particularly noisy. Furthermore, Handbury et al. (2013) investigate the informativeness and bias of the consumer price index (CPI) as a proxy for the "true" inflation and use a classical measurement error model to test for bias in Japanese CPI. This raises an interesting question whether the "ARCH behavior" is only a manifest phenomenon in empirical (observed) processes, or it is an intrinsic property of the underlying (unobserved) processes. Therefore it is of theoretical and practical interests to investigate the problem and impact of measurement error in ARCH-type models.

The errors-in-variables problem has been extensively studied in statistics and econometrics, see, e.g., Carroll et al. (2006); Chen et al. (2011); Wang and Hsiao (2011); Yi et al. (2021), and the references therein. However, most of the research focuses mainly on the problem of measurement error in covariates in regression models. For dynamic models, Staudenmayer and Buonaccorsi (2005) studied autoregressive (AR) model with white noise errors and mismeasured response process, while Buonaccorsi (2010) gives an overview of estimation in dynamic models. Some researchers, e.g., Harvey et al. (1992), Gourieroux et al. (1993) and Francq and Zakoïan (2000), have considered GARCH models where the innovation term contains an unobserved white noise component. However, research focusing on the mismeasured response processes in GARCH models is sparse and even answers to very basic questions are not known. For example, what is the impact of measurement error on parameter estimation and inference? Under what conditions is the model identifiable? How to quantify and correct the estimation bias caused by measurement error?

In this paper we attempt to address some of these questions. To simplify notation and analysis, we start with an autoregressive model with ARCH innovation where the true latent process is measured with additive white noise error process. In contrast to the models with covariate measurement error, we show that all model parameters are identifiable by the observed proxy process only and no extra information is needed. Moreover, we propose a set of moment conditions that are sufficient for the identifiability and therefore can used to construct GMM estimators for the unknown parameters. We investigate the impact of measurement error on the parameter estimation in dynamic ARCH models through Monte Carlo simulations. In particular, we show that the measurement error induces biases in the naive maximum likelihood estimators and the relative biases have certain functional forms. We also develop a statistical test for the presence of measurement error, which is useful because more efficient GMM or maximum quasi-likelihood estimators can be used if the measurement error is found to be absent or ignorable. Finally, we carry out Monte Carlo simulations to examine the finite sample behavior of our proposed estimators and compare them with the naive maximum likelihood estimators.

The paper is organized as follows. In Sect. 2 we introduce the model and show it is identifiable without extra information. In Sect. 3 we construct the GMM estimators and provide their asymptotic properties. In Sect. 4 we propose a test for the measurement error. Further, we carry out Monte Carlo simulations to study the impact of measurement error on the naive estimators in Sect. 5 and to examine the finite sample properties of the proposed estimators and compare them with the naive MLE in Sect. 6. Finally, conclusions and discussions are given in Sect. 7, while regularity assumptions and mathematical proofs are in the Appendix.

2 The Model and Identifiability

Let $\{X_t\}$ be the unique nonanticipative strictly stationary solution of the following AR(p)-ARCH(q) model (Francq and Zakoian 2011, Ch. 7)

$$X_t = \alpha_0(B) X_t + \epsilon_t, \quad t \in \mathbb{Z}, \tag{1}$$

$$\epsilon_t = \sqrt{h_t}\eta_t, \quad h_t = \omega_0 + \beta_0(B)\epsilon_t^2, \tag{2}$$

where $\alpha_0(B) = \sum_{i=1}^{p} \alpha_{0i} B^i$, $\beta_0(B) = \sum_{j=1}^{q} \beta_{0j} B^j$, B is the backshift operator, and $\{\eta_t\}$ is a sequence of iid random variables with $E(\eta_t) = 0$ and $E({\eta_t}^2) = 1$. Under this model $\{X_t\}$ is second-order stationary if the unknown parameters satisfy $\omega_0 > 0$, $\beta_{0j} \geq 0$, $j = 1, 2, \ldots, q$, $\sum_{j=1}^{q} \beta_{0j} < 1$ and $\alpha_0(z) \neq 1$ for all $|z| \leq 1$. Moreover, under these conditions $\{X_t\}$ is strictly stationary and ergodic (Francq and Zakoian 2011, Th. 2.5).

Assume that $\{X_t\}$ is not directly observable and instead we observe the proxy process

$$Z_t = X_t + \delta_t, \tag{3}$$

where the measurement error process $\{\delta_t\}$ is iid with $E(\delta_t) = 0$, $E(\delta_t^2) = \sigma_0^2$ and is independent of $\{\eta_t\}$. Note that such a classical measurement error model is commonly used in the literature and is also used by Handbury et al. (2013). Our main interest is consistent estimation of unknown parameters $\boldsymbol{\theta}_0 = (\boldsymbol{\alpha}_0', \boldsymbol{\beta}_0', \omega_0, \sigma_0^2)'$, where $\boldsymbol{\alpha}_0' = (\alpha_{01}, \ldots, \alpha_{0p})$ and $\boldsymbol{\beta}_0' = (\beta_{01}, \ldots, \beta_{0q})$. If $\{X_t\}$ were observable, then this can be done by using standard methods such as least squares or quasi-likelihood methods. However, when only observations on $\{Z_t\}$ are available, several issues arise and one of them is the model identifiability.

It is well-known that in a regression model with covariate measurement error usually extra information such as replicate or instrumental data are needed in order for all parameters to be identifiable. Here we demonstrate that, in contrary, all parameters in model (1)–(3) are identifiable based on the observations on $\{Z_t\}$ only. To simplify notation, we consider the case where $p = q = 1$ and let $Y_t = Z_t - \alpha_0 Z_{t-1}$.

Then under model assumptions we have

$$\begin{aligned}
E(Y_t|Z_s, s \le t-2) &= E(Z_t|Z_s, s \le t-2) - \alpha_0 E(Z_{t-1}|Z_s, s \le t-2) \\
&= E(\epsilon_t|Z_s, s \le t-2) + E(\delta_t - \alpha_0\delta_{t-1}|Z_s, s \le t-2) \\
&= 0.
\end{aligned}$$

Since both $E(Z_t|Z_s, s \le t-2)$ and $E(Z_{t-1}|Z_s, s \le t-2)$ are observable functions, α_0 is uniquely identified by the above equation. In order to see the identifiability of other parameters, we consider higher moments. In particular, since

$$E(Y_t Y_{t-1}) = E(-\alpha_0\delta_{t-1}^2) = -\alpha_0\sigma_0^2,$$

σ_0^2 is identified given that α_0 is identified. Further, from

$$E(Y_t^2|Z_s, s \le t-3) = \omega_0 + (1-\beta_0)(1+\alpha_0^2)\sigma_0^2 + \beta_0 E(Y_{t-1}^2|Z_s, s \le t-3),$$

it is easy to see that β_0 and $\omega_0 + (1-\beta_0)(1+\alpha_0^2)\sigma_0^2$ are uniquely determined and hence ω_0 is identified.

3 GMM Estimation

Motivated by the above discussion of identifiability, in this section we propose an estimation procedure based on the following conditional moments. Specifically, let

$$Y_t(\boldsymbol{\alpha}_0) = [1 - \boldsymbol{\alpha}_0(B)]Z_t. \tag{4}$$

Then under the model assumptions we have (w.p.1)

$$\begin{aligned}
E\{Y_t(\boldsymbol{\alpha}_0)|Z_s, s < t-p\} &= E\{\epsilon_t|Z_s, s < t-p\} + E\{[1-\boldsymbol{\alpha}_0(B)]\delta_t|Z_s, s < t-p\} \\
&= 0
\end{aligned} \tag{5}$$

and

$$E\left\{[1-\boldsymbol{\beta}_0(B)]Y_t^2(\boldsymbol{\alpha}_0)|Z_s, s < t-p-q\right\} = \omega_0 + [1x\boldsymbol{\beta}_0(1)][1+\boldsymbol{\alpha}_0^2(1)]\sigma_0^2, \tag{6}$$

where $\boldsymbol{\beta}_0(1) = \sum_{j=1}^{q} \beta_{0j}$ and $\boldsymbol{\alpha}_0^2(1) = \sum_{j=1}^{p} \alpha_{0j}^2$. In addition, we have the following unconditional moment condition

$$E\left\{Y_t(\boldsymbol{\alpha}_0)Y_{t-1}(\boldsymbol{\alpha}_0)\right\} = \left[\sum_{j=1}^{p-1} \alpha_{0j}\alpha_{0(j+1)} - \alpha_{01}\right]\sigma_0^2. \tag{7}$$

Therefore a GMM estimator for $\boldsymbol{\theta}_0$ can be constructed as follows. Denote

$$\boldsymbol{r}_t(\boldsymbol{\theta}) = \begin{pmatrix} Y_t(\boldsymbol{\alpha}) \\ [1-\boldsymbol{\beta}(B)]\,Y_t^2(\boldsymbol{\alpha}) - \omega - [1-\boldsymbol{\beta}(1)][1+\boldsymbol{\alpha}^2(1)]\sigma^2 \\ Y_t(\boldsymbol{\alpha})Y_{t-1}(\boldsymbol{\alpha}) - [\sum_{j=1}^{p-1}\alpha_j\alpha_{j+1} - \alpha_1]\sigma^2 \end{pmatrix} \tag{8}$$

and the matrix of instrumental functions

$$\boldsymbol{G}_t = \begin{pmatrix} f_1(\tilde{Z}_{t-p-1}) & 0 & 0 \\ 0 & f_2(\tilde{Z}_{t-p-q-1}) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{9}$$

where $f_1(\tilde{Z}_{t-p-1})$ is a k_1-vector of measurable functions of $\tilde{Z}_{t-p-1} = (Z_{t-p-1}, Z_{t-p-2}, \ldots)$, $f_2(\tilde{Z}_{t-p-q-1})$ is a k_2-vector of functions of $\tilde{Z}_{t-p-q-1} = (Z_{t-p-q-1}, Z_{t-p-q-2}, \ldots)$, and $k_1 \geq p, k_2 \geq q+1$ are chosen to achieve identification and efficiency. Then from (5)–(7) we have

$$E\left\{\boldsymbol{G}_t\boldsymbol{r}_t(\boldsymbol{\theta}_0)\right\} = 0. \tag{10}$$

To simplify notation, in the following we assume that $\tilde{Z}_{t-p-1} = (Z_{t-p-1}, \ldots, Z_{t-p-k_1})$ and $\tilde{Z}_{t-p-q-1} = (Z_{t-p-q-1}, \ldots, Z_{t-p-q-k_2})$. Given the observations $Z_\tau, Z_{\tau+1}, \ldots, Z_n$, $\tau = \min\{1-p-k_1, 1-p-q-k_2\}$, the GMM estimator is given by

$$\hat{\boldsymbol{\theta}}_n = \underset{\boldsymbol{\Theta}}{\operatorname{argmin}}[\sum_{t=1}^{n}\boldsymbol{G}_t\boldsymbol{r}_t(\boldsymbol{\theta})]'\boldsymbol{\Lambda}_{\boldsymbol{n}}[\sum_{t=1}^{n}\boldsymbol{G}_t\boldsymbol{r}_t(\boldsymbol{\theta})], \tag{11}$$

where $\boldsymbol{\Lambda}_{\boldsymbol{n}}$ is a nonnegative definite matrix which may depend on the observed data and converges to a positive definite matrix $\boldsymbol{\Lambda}$ as $n \to \infty$. The parameter space $\boldsymbol{\Theta} \subset \mathbb{R}^p \times [0,\infty)^q \times (0,\infty) \times [0,\infty)$ is assumed to be compact and contain $\boldsymbol{\theta}_0$ as an interior point. The asymptotic properties of $\hat{\boldsymbol{\theta}}_n$ can be established in a usual GMM framework. Specifically, denote $\boldsymbol{\Sigma} = E[\boldsymbol{G}_0\boldsymbol{r}_0(\boldsymbol{\theta}_0)\boldsymbol{r}_0'(\boldsymbol{\theta}_0)\boldsymbol{G}_0']$ and $\boldsymbol{\Delta}' = E[\nabla_{\boldsymbol{\theta}}\boldsymbol{r}_0'(\boldsymbol{\theta}_0)\boldsymbol{G}_0']$ where $\nabla_{\boldsymbol{\theta}}\boldsymbol{r}_0'(\boldsymbol{\theta}) = \partial\boldsymbol{r}_0'(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$. Then we have the following asymptotic results for the GMM estimator, the proof of which and further regularity conditions are given in Section 8.

Theorem 1 *The GMM estimator $\hat{\boldsymbol{\theta}}_n$ has the following properties.*

(1) Under Assumption 1–2, $\hat{\boldsymbol{\theta}}_n \xrightarrow{a.s.} \boldsymbol{\theta}_0$ as $n \to \infty$.

(2) Under Assumption 1–6, $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1})$ as $n \to \infty$, where $\boldsymbol{A} = \boldsymbol{\Delta}'\boldsymbol{\Lambda}\boldsymbol{\Delta}$, $\boldsymbol{B} = \boldsymbol{\Delta}'\boldsymbol{\Lambda}\boldsymbol{\Sigma}\boldsymbol{\Lambda}\boldsymbol{\Delta}$.

Given the specified set of instruments, the optimal (efficient) GMM estimator is obtained by taking the weight $\boldsymbol{\Lambda}_n$ to be such that $\boldsymbol{\Lambda}_n^{-1} \xrightarrow{p} \boldsymbol{\Sigma}$ as $n \to \infty$. Then the optimal GMM has asymptotic variance-covariance matrix $\boldsymbol{A}_0^{-1} = (\boldsymbol{\Delta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\Delta})^{-1}$.

To compute the optimal weight, we propose to use a serial correlation robust estimator of $\boldsymbol{\Sigma}$. Specifically, let $\boldsymbol{e}_t = \boldsymbol{G}_t\boldsymbol{r}_t(\boldsymbol{\theta}_0)$. Then since $E(\boldsymbol{e}_t|\mathcal{F}_{t-p-q-1}) = 0$, we have

$$\begin{aligned}\boldsymbol{\Sigma}_n &= V\left(n^{-1/2}\sum_{i=1}^{n}\boldsymbol{e}_t\right)\\ &= E(\boldsymbol{e}_0\boldsymbol{e}_0') + \sum_{i=1}^{p+q}\frac{n-i}{n}[E(\boldsymbol{e}_0\boldsymbol{e}_{-i}') + E(\boldsymbol{e}_{-i}\boldsymbol{e}_0')].\end{aligned}$$

Similarly to White (2001, p.147) and Wooldridge (1994, Sec.4.5), we can find a positive definite matrix $\boldsymbol{\Lambda}_n$ such that $\boldsymbol{\Lambda}_n^{-1} - \boldsymbol{\Sigma}_n \xrightarrow{p} 0$ as $n \to \infty$, where

$$\boldsymbol{\Lambda}_n^{-1} = \frac{1}{n}\sum_{t=1}^{n}\hat{\boldsymbol{e}}_t\hat{\boldsymbol{e}}_t' + \sum_{i=1}^{p+q}\frac{m_i(n)}{n}\sum_{t=i+1}^{n}(\hat{\boldsymbol{e}}_t\hat{\boldsymbol{e}}_{t-i}' + \hat{\boldsymbol{e}}_{t-i}\hat{\boldsymbol{e}}_t'),$$

$\hat{\boldsymbol{e}}_t = \boldsymbol{G}_t\boldsymbol{r}_t(\hat{\boldsymbol{\theta}})$ and $m_i(n) \to 1, i = 1, 2, \ldots, p+q$ are suitably chosen to ensure that $\boldsymbol{\Lambda}_n > 0$. In practice, we can start with $m_i(n) = 1, i = 1, 2, \ldots, p+q$. If $\boldsymbol{\Lambda}_n$ is not positive definite, then we can modify $m_i(n)$ as $m_i(n) = (1 - n^{-1})^i$ or $m_i(n) = \exp(i/n)$ to achieve the desired result.

4 Testing for Measurement Error

Although our GMM framework does not rule out zero measurement error, from practical point of view it is of interest to verify its presence and severity. However, testing for measurement error is generally a challenging task because under the null hypothesis the value of the measurement error variance is on the boundary of the parameter space. The framework in the previous section provides a possibility to construct such a test by applying the similar idea of the incremental Sargan test (Arellano 2003, p.193). Specifically, we construct a test for the following hypotheses on the measurement error variance

$$H_0 : \sigma_0^2 = 0 \quad vs. \quad H_a : \sigma_0^2 > 0. \tag{12}$$

We first consider the problem of estimating a subset of unknown parameters $\boldsymbol{\gamma}_0 = (\boldsymbol{\alpha}_0', \boldsymbol{\beta}_0', \tau_0)'$, where $\boldsymbol{\alpha}_0, \boldsymbol{\beta}_0$ are defined as in the AR(p)-ARCH(q) model (1)–(2) and

$$\tau_0 = \omega_0 + [1 - \beta_0(1)][1 + \alpha_0^2(1)]\sigma_0^2. \tag{13}$$

Then it can be shown that $\boldsymbol{\gamma}_0$ can be identified by the following $k_1 + k_2$ moment conditions

$$E[\boldsymbol{G}_{1t}\tilde{\boldsymbol{r}}_t(\boldsymbol{\gamma}_0)] = 0, \tag{14}$$

where $\tilde{\boldsymbol{r}}_t(\boldsymbol{\gamma}) = (Y_t(\boldsymbol{\alpha}), [1 - \beta(B)]Y_t^2(\boldsymbol{\alpha}) - \tau)'$,

$$\boldsymbol{G}_{1t}' = \begin{pmatrix} Z_{t-p-1} & Z_{t-p-2} & \cdots & Z_{t-p-k_1} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & Z_{t-p-q-1} & Z_{t-p-q-2} & \cdots & Z_{t-p-q-k_2} \end{pmatrix}$$

and $k_1 > p, k_2 > q + 1$. Therefore the optimal GMM estimator for $\boldsymbol{\gamma}_0$ is given by

$$\hat{\boldsymbol{\gamma}}_1 = \underset{\Gamma}{\operatorname{argmin}}\, \boldsymbol{b}_{1n}'(\boldsymbol{\gamma})\boldsymbol{V}_{1n}^{-1}\boldsymbol{b}_{1n}(\boldsymbol{\gamma}), \tag{15}$$

where $\boldsymbol{b}_{1n}(\boldsymbol{\gamma}) = n^{-1}\sum_{t=1}^n \boldsymbol{G}_{1t}\tilde{\boldsymbol{r}}_t(\boldsymbol{\gamma})$, $\boldsymbol{V}_{1n}$ is positive definite and $\boldsymbol{V}_{1n} - V[n^{1/2}\boldsymbol{b}_{1n}(\boldsymbol{\gamma}_0)] \xrightarrow{p} 0$, and $\Gamma \subset \mathbb{R}^p \times [0, \infty)^q \times (0, \infty)$ is compact.

Next, under H_0 we consider additional $2p + q$ moment conditions

$$E[\boldsymbol{G}_{2t}\tilde{\boldsymbol{r}}_t(\boldsymbol{\gamma}_0)] = 0, \tag{16}$$

where

$$\boldsymbol{G}_{2t}' = \begin{pmatrix} Z_{t-1} & Z_{t-2} & \cdots & Z_{t-p} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & Z_{t-1} & Z_{t-2} & \cdots & Z_{t-p-q} \end{pmatrix}.$$

Similarly, the optimal GMM estimator is given by

$$\hat{\boldsymbol{\gamma}} = \underset{\Gamma}{\operatorname{argmin}}\, \boldsymbol{b}_n'(\boldsymbol{\gamma})\boldsymbol{V}_n^{-1}\boldsymbol{b}_n(\boldsymbol{\gamma}), \tag{17}$$

where $\boldsymbol{b}_n(\boldsymbol{\gamma}) = n^{-1}\sum_{t=1}^n (\boldsymbol{G}_{1t}' \vdots \boldsymbol{G}_{2t}')'\tilde{\boldsymbol{r}}_t(\boldsymbol{\gamma})$, $\boldsymbol{V}_n$ is positive definite and $\boldsymbol{V}_n - V[n^{1/2}\boldsymbol{b}_n(\boldsymbol{\gamma}_0)] \xrightarrow{p} 0$.

Then the test statistic is defined as

$$SW = n\boldsymbol{b}_n'(\hat{\boldsymbol{\gamma}})\boldsymbol{V}_n^{-1}\boldsymbol{b}_n(\hat{\boldsymbol{\gamma}}) - n\boldsymbol{b}_{1n}'(\hat{\boldsymbol{\gamma}}_1)\boldsymbol{V}_{1n}^{-1}\boldsymbol{b}_{1n}(\hat{\boldsymbol{\gamma}}_1). \tag{18}$$

For this test we have the following result, the proof of which is given in section 8.

Theorem 2 *Under Assumption 1–6 and H_0, $SW \xrightarrow{d} \chi^2_{2p+q}$ as $n \to \infty$.*

5 Impact of Measurement Error

It is well-known that in a linear errors-in-variables model with iid data the usual OLS or ML estimators are attenuated towards zero. The impact of the measurement error in dynamic ARCH models, however, has not been studied before. In this section we carry out Monte Carlo simulations to investigate the behavior of the naive MLE of a Gaussian AR(1)-ARCH(1) model with Gaussian classical additive measurement error. Specifically, we consider the model

$$X_t = \alpha_0 X_{t-1} + \epsilon_t, \tag{19}$$

$$\epsilon_t = \sqrt{h_t}\eta_t, \quad h_t = \omega_0 + \beta_0 \epsilon^2_{t-1}, \tag{20}$$

$$Z_t = X_t + \delta_t, \tag{21}$$

where $\eta_t \sim N(0, 1)$ and $\delta_t \sim N(0, \sigma_0^2)$ are independent and iid sequences. The parameter values are set to $\omega_0 = \sigma_0^2 = 1$, $\alpha_0 \in \{-0.9, -0.8, \ldots, 0.8, 0.9\}$ and $\beta_0 \in \{0.05, 0.1, \ldots, 0.9, 0.95\}$, respectively. In all simulations, 1000 samples of size $n = 10^5$ are generated to accurately estimate the asymptotic bias of ML$(\alpha_0, \beta_0, \omega_0)$.

We first calculate the relative bias of the ML(α_0) as

$$RB.ML(\alpha_0) = \frac{Bias.ML(\alpha_0)}{\alpha_0} \times 100.$$

Figure 1 shows clearly that the ML(α_0) is biased towards zero, similar to the OLS estimator of the slope parameter in a linear errors-in-variables model. More importantly, the bias has a pattern of a symmetric parabolic function in α_0 and a nearly linear function in β_0. The absolute RB is monotone decreasing in both α_0 and β_0. These observations indicate a similarity between the asymptotic bias of ML(α_0) and OLS(α_0) calculated by regressing Z_t on Z_{t-1}. By direct calculation we can obtain the OLS relative bias as

$$RB.OLS(\alpha_0) = -\frac{1}{1 + \omega_0/(1-\beta_0)(1-\alpha_0^2)\sigma_0^2}.$$

This raises an interesting question: To what extend can the OLS(α_0) bias formula be used to approximate and therefore to correct the bias of ML(α_0)? To further investigate this question, we examine the ratio $Bias.OLS(\alpha_0)/Bias.ML(\alpha_0)$ as a function of α_0 and β_0. Figure 2 shows that the formula of Bias.OLS provides good approximation to Bias.ML in a fairly large area of the parameter space. However,

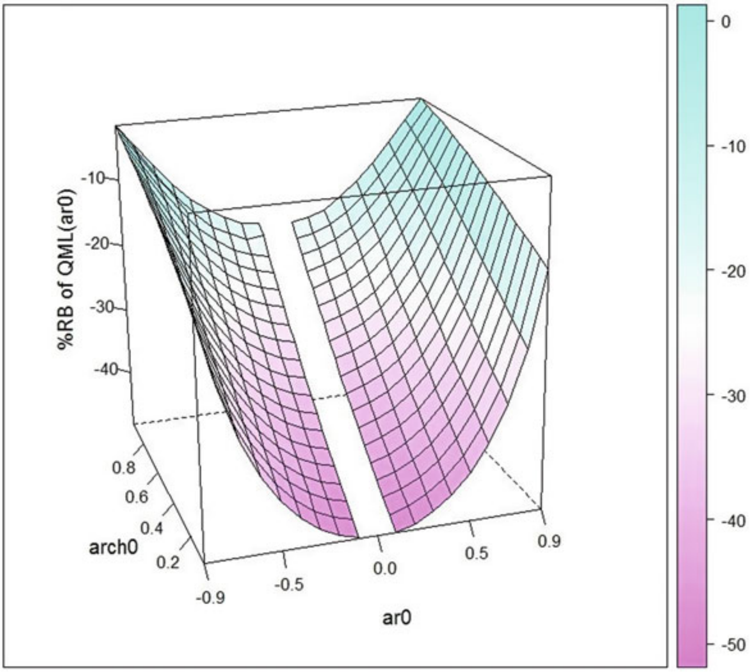

Fig. 1 Relative bias of ML(α_0)

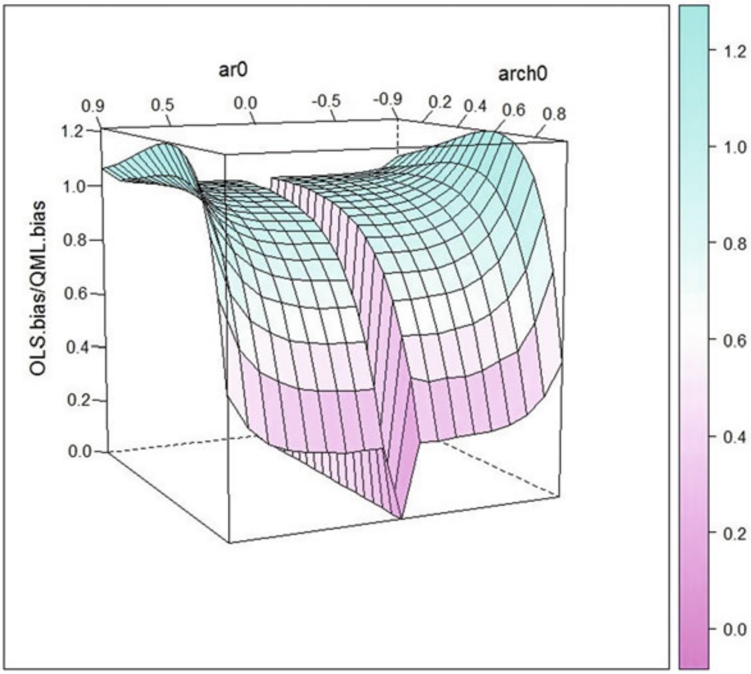

Fig. 2 Ratio of Bias.OLS(α_0)/Bias.ML(α_0)

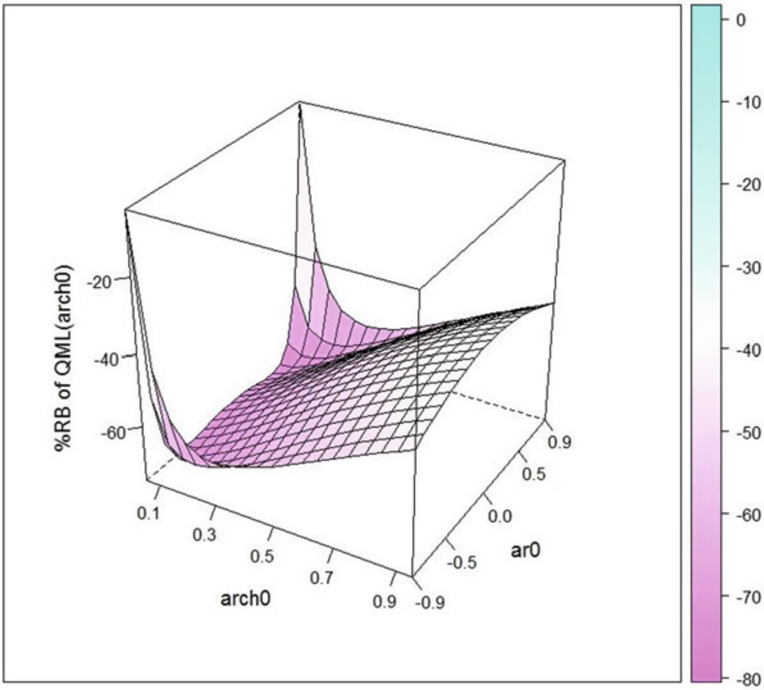

Fig. 3 Relative bias of ML(β_0)

the Bias.OLS formula underestimates the Bias.ML for large values of β_0, which is understandable because the two estimators are more different when β_0 gets larger. Again, it is interesting to see that there is a clear (unknown) functional relationship between Bias.ML and Bias.OLS.

Further, we have also calculated the relative bias of ML(β_0) and ML(ω_0), which are shown in Figs. 3 and 4, respectively. From these figures we can see that the ML(β_0) has downward bias and the absolute relative bias is generally decreasing for $\beta_0 \geq 0.3$ or $|\alpha_0| \leq 0.7$. In contrast, ML(ω_0) has an upward bias pattern, which is similar to the intercept estimator in a linear errors-in-variables model. Again, both Figs. 3 and 4 show clear (but unknown) functional patterns of the asymptotic bias of the MLE. Overall, Figs. 1, 2, 3, and 4 show that the measurement error has more severe effect on the estimate of ω_0 than on α_0 and β_0.

6 Finite Sample Properties

In this section we carry out Monte Carlo simulations to investigate the finite sample properties of the proposed GMM estimator and compare it with the corresponding naive ML estimator. Again we use the model (19)–(21) in the previous section,

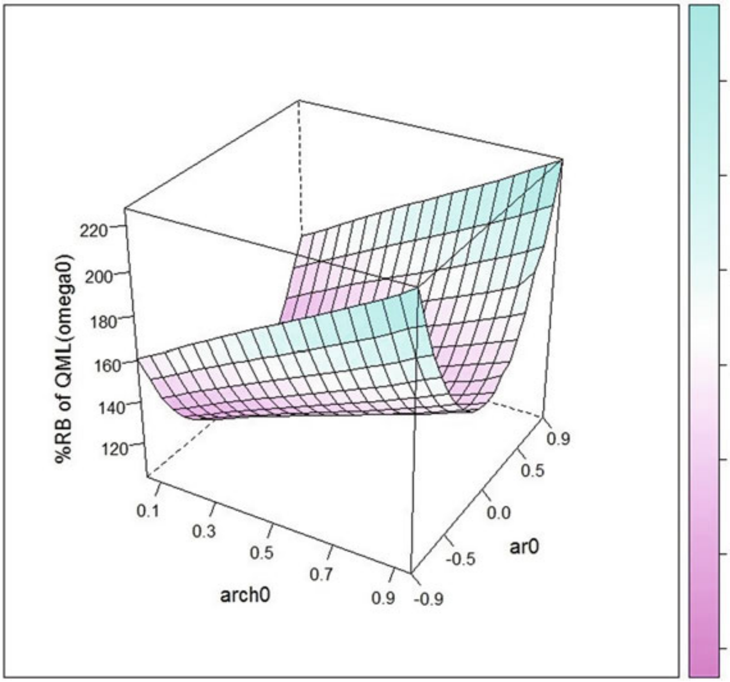

Fig. 4 Relative bias of ML(ω_0)

under which the optimal choice of the instrument matrix $\boldsymbol{G}_t$ depends on the quantities such as $E[Z_{t-h}|\tilde{Z}_{t-p-1}]$, $h = 1, 2, \ldots, p$, $E[Y^2_{t-h}(\alpha_0)|\tilde{Z}_{t-p-q-1}]$, $h = 1, 2, \ldots, q$, and $E[(\boldsymbol{r}_t(\boldsymbol{\theta}_0)\boldsymbol{r}'_t(\boldsymbol{\theta}_0))_{ij}|I_{ij}]$ for some suitably chosen information set I_{ij}. Unfortunately some of these instrumental functions cannot be computed easily without simplification which would require further distributional assumptions on the latent and error processes. Consequently we have attempted with several constructions and found the following procedure to be most practical. Since the number of moment equations used here is the same of the number of unknown parameters, the estimators can be calculated in the following sequential process.

First, compute

$$\hat{\alpha} = \underset{-1<\alpha<1}{\operatorname{argmin}} \left[\sum_{t=k_3}^{n} \hat{Z}_{t-1} Y_t(\alpha) \right]^2, \tag{22}$$

where $k_3 = 4 + k_2$ and $\hat{Z}_{t-1}$ is the linear projection of Z_{t-1} onto $\{Z_{t-2}, Z_{t-3}, \ldots, Z_{t-1-k_1}\}$. Second, let $Y_t = Y_t(\hat{\alpha})$ and compute

$$\hat{\sigma}^2 = \underset{\sigma^2 \geq 0}{\operatorname{argmin}} \left[\sum_{t=k_3}^{n} (Y_t Y_{t-1} + \hat{\alpha}\sigma^2) \right]^2 . \tag{23}$$

Third, compute

$$\hat{\beta} = \underset{0<\beta<1}{\operatorname{argmin}} \left[\sum_{t=k_3}^{n} \hat{Y}_{t-1}^2 (y_t^2 - \beta y_{t-1}^2) \right]^2 , \tag{24}$$

where $\hat{Y}_{t-1}^2$ is the linear projection of Y_{t-1}^2 onto

$$\left\{ \frac{Y_{t-3}^2}{1+Y_{t-3}^2}, \frac{Y_{t-4}^2}{1+Y_{t-4}^2}, \dots, \frac{Y_{t-2-k_2}^2}{1+Y_{t-2-k_2}^2} \right\},$$

and $y_t^2 = Y_t^2 - \overline{Y^2}$, $y_{t-1}^2 = Y_{t-1}^2 - \overline{Y_{-1}^2}$ with

$$\overline{Y^2} = \frac{1}{n-k_3+1} \sum_{t=k_3}^{n} Y_t^2, \quad \overline{Y_{-1}^2} = \frac{1}{n-k_3+1} \sum_{t=k_3}^{n} Y_{t-1}^2.$$

Finally, compute

$$\hat{\omega} = \underset{\omega>0}{\operatorname{argmin}} \left[\overline{Y^2} - \hat{\beta}\overline{Y_{-1}^2} - \omega - (1-\hat{\beta})(1+\hat{\alpha}^2)\hat{\sigma}^2 \right]^2 . \tag{25}$$

It is worthwhile to note that $\hat{Y}_{t-1}^2$ is defined in terms of bounded instruments to guarantee the consistency of the proposed estimator over a wide range of the parameter space.

We generate the data using parameter values $\{0.05, 0.2, 0.35, 0.5, 0.65, 0.8, 0.95\}$ for α_0 and β_0, respectively. In addition, we set $\omega_0 = 1$ and let σ_0^2 vary proportionally to $\sigma_X^2 = \omega_0/(1-\beta_0)(1-\alpha_0^2)$ such as $\sigma_0^2 = a\sigma_X^2$, where the noise-to-signal ratio $a \in \{0, 0.25, 0.5, \dots, 1.75, 2\}$, respectively, and $a = 0$ corresponds to the case of no measurement error. Again, we generate 1000 samples for each of the sizes $n = 100, 1000, 10000$ and $n = 100000$ to approximate the asymptotic scenario. In each simulation we compute the naive ML (nML) and two GMM estimators using $k_1 = k_2 = 1$ (GMM1) and $k_1 = k_2 = 5$ (GMM5) instruments, respectively.

The bias and root mean squared error (RMSE) of the estimators are calculated and numerical results for AR and ARCH parameters $\alpha_0, \beta_0, \omega_0$ for some selected cases are reported in Tables 1, 2, and 3. The numerical results for negative α_0 values are similar to those for positive values and therefore are not reported here.

From Tables 1, 2, and 3 we can see that, in the case of no measurement error ($a = 0$), the GMM estimators have both larger bias and RMSE than the naive

Table 1 Bias and **RMSE** of nML, GMM1 and GMM5 estimators for AR parameter α_0 (with $\omega_0 = 1$)

		$a = 0.00$			$a = 0.75$			$a = 1.5$		
		nML	GMM1	GMM5	nML	GMM1	GMM5	nML	GMM1	GMM5
n	α_0	$\beta_0 = 0.2$								
100	0.2	−0.007	−0.022	−0.029	−0.095	0.020	−0.093	−0.120	−0.055	−0.146
		0.109	**0.510**	**0.340**	**0.142**	**0.580**	**0.402**	**0.159**	**0.622**	**0.420**
	0.5	−0.011	−0.020	−0.023	−0.227	−0.003	−0.122	−0.304	−0.082	−0.249
		0.094	**0.217**	**0.180**	**0.254**	**0.367**	**0.331**	**0.323**	**0.498**	**0.448**
	0.8	−0.015	−0.019	−0.021	−0.359	−0.012	−0.096	−0.489	−0.073	−0.234
		0.067	**0.092**	**0.089**	**0.382**	**0.192**	**0.208**	**0.505**	**0.298**	**0.373**
10^5	0.2	−0.000	0.000	0.000	−0.086	−0.001	−0.001	−0.120	−0.002	−0.003
		0.003	**0.016**	**0.016**	**0.086**	**0.028**	**0.028**	**0.120**	**0.040**	**0.039**
	0.5	−0.000	0.000	0.000	−0.214	−0.000	−0.000	−0.300	−0.001	−0.001
		0.003	**0.005**	**0.005**	**0.214**	**0.010**	**0.010**	**0.300**	**0.016**	**0.014**
	0.8	−0.000	−0.000	−0.000	−0.339	−0.000	−0.000	−0.479	−0.001	−0.001
		0.002	**0.002**	**0.002**	**0.339**	**0.005**	**0.004**	**0.479**	**0.009**	**0.005**
		$\beta_0 = 0.5$								
100	0.2	−0.004	−0.063	−0.053	−0.099	−0.010	−0.106	−0.123	−0.086	−0.148
		0.107	**0.565**	**0.353**	**0.149**	**0.614**	**0.409**	**0.165**	**0.655**	**0.418**
	0.5	−0.008	−0.025	−0.038	−0.240	−0.034	−0.143	−0.313	−0.118	−0.263
		0.091	**0.244**	**0.207**	**0.268**	**0.432**	**0.351**	**0.334**	**0.548**	**0.464**
	0.8	−0.011	−0.021	−0.025	−0.375	−0.022	−0.114	−0.501	−0.094	−0.260
		0.060	**0.101**	**0.098**	**0.402**	**0.210**	**0.235**	**0.521**	**0.342**	**0.406**
10^5	0.2	−0.000	0.001	0.001	−0.090	−0.001	−0.002	−0.123	−0.001	−0.002
		0.003	**0.026**	**0.026**	**0.090**	**0.037**	**0.034**	**0.123**	**0.044**	**0.043**
	0.5	−0.000	0.000	0.000	−0.222	−0.000	−0.001	−0.307	−0.001	−0.001
		0.003	**0.009**	**0.009**	**0.222**	**0.012**	**0.011**	**0.307**	**0.017**	**0.015**
	0.8	−0.000	0.000	0.000	−0.343	−0.000	−0.000	−0.484	−0.000	−0.000
		0.002	**0.003**	**0.003**	**0.343**	**0.006**	**0.004**	**0.484**	**0.009**	**0.005**
		$\beta_0 = 0.8$								
100	0.2	−0.003	−0.109	−0.071	−0.120	−0.064	−0.131	−0.138	−0.109	−0.156
		0.107	**0.614**	**0.371**	**0.169**	**0.676**	**0.423**	**0.180**	**0.701**	**0.421**
	0.5	−0.006	−0.049	−0.061	−0.290	−0.082	−0.215	−0.350	−0.173	−0.315
		0.093	**0.326**	**0.254**	**0.320**	**0.520**	**0.418**	**0.373**	**0.628**	**0.502**
	0.8	−0.009	−0.029	−0.036	−0.452	−0.058	−0.185	−0.559	−0.172	−0.359
		0.053	**0.142**	**0.135**	**0.483**	**0.297**	**0.329**	**0.581**	**0.482**	**0.516**
10^5	0.2	−0.000	−0.000	−0.005	−0.105	−0.016	−0.017	−0.136	−0.006	−0.013
		0.003	**0.170**	**0.130**	**0.105**	**0.173**	**0.128**	**0.137**	**0.154**	**0.129**
	0.5	−0.000	0.001	−0.001	−0.264	−0.007	−0.008	−0.342	−0.001	−0.003
		0.004	**0.059**	**0.051**	**0.264**	**0.093**	**0.066**	**0.342**	**0.053**	**0.043**
	0.8	−0.000	−0.000	−0.001	−0.411	−0.002	−0.003	−0.545	−0.001	−0.001
		0.001	**0.020**	**0.019**	**0.411**	**0.032**	**0.027**	**0.545**	**0.018**	**0.014**

Table 2 Bias and **RMSE** of nML, GMM1 and GMM5 estimators for ARCH slope parameter β_0 (with $\omega_0 = 1$)

		$a = 0.00$			$a = 0.75$			$a = 1.5$		
		nML	GMM1	GMM5	nML	GMM1	GMM5	nML	GMM1	GMM5
n	β_0	$\alpha_0 = 0.2$								
100	0.2	−0.030	0.177	0.041	−0.133	0.216	0.039	−0.150	0.225	0.043
		0.146	**0.461**	**0.335**	**0.161**	**0.495**	**0.363**	**0.169**	**0.502**	**0.361**
	0.5	−0.064	−0.044	−0.201	−0.353	−0.079	−0.268	−0.407	−0.074	−0.283
		0.219	**0.424**	**0.388**	**0.385**	**0.447**	**0.429**	**0.425**	**0.451**	**0.431**
	0.8	−0.107	−0.326	−0.460	−0.601	−0.366	−0.580	−0.669	−0.381	−0.615
		0.253	**0.519**	**0.567**	**0.636**	**0.580**	**0.669**	**0.692**	**0.590**	**0.696**
10^5	0.2	−0.000	−0.000	−0.009	−0.134	−0.036	−0.044	−0.167	−0.040	−0.029
		0.005	**0.098**	**0.123**	**0.134**	**0.150**	**0.133**	**0.167**	**0.198**	**0.125**
	0.5	−0.000	−0.002	−0.003	−0.316	−0.005	−0.009	−0.391	−0.003	−0.022
		0.006	**0.032**	**0.035**	**0.316**	**0.064**	**0.064**	**0.391**	**0.125**	**0.108**
	0.8	−0.001	−0.021	−0.024	−0.528	−0.025	−0.036	−0.616	−0.032	−0.051
		0.008	**0.048**	**0.052**	**0.528**	**0.066**	**0.073**	**0.616**	**0.080**	**0.086**
		$\alpha_0 = 0.5$								
100	0.2	−0.028	0.152	0.047	−0.142	0.188	0.021	−0.155	0.219	0.020
		0.145	**0.441**	**0.342**	**0.166**	**0.484**	**0.349**	**0.171**	**0.496**	**0.335**
	0.5	−0.062	−0.041	−0.188	−0.380	−0.103	−0.289	−0.421	−0.071	−0.295
		0.219	**0.417**	**0.384**	**0.405**	**0.448**	**0.435**	**0.436**	**0.452**	**0.432**
	0.8	−0.104	−0.317	−0.449	−0.629	−0.394	−0.612	−0.685	−0.378	−0.615
		0.252	**0.502**	**0.561**	**0.659**	**0.593**	**0.687**	**0.704**	**0.593**	**0.696**
10^5	0.2	−0.000	−0.000	−0.009	−0.150	−0.040	−0.035	−0.175	0.017	0.030
		0.005	**0.098**	**0.123**	**0.150**	**0.175**	**0.121**	**0.175**	**0.290**	**0.218**
	0.5	−0.000	−0.002	−0.003	−0.351	−0.003	−0.014	−0.410	0.004	−0.040
		0.006	**0.032**	**0.035**	**0.351**	**0.091**	**0.087**	**0.411**	**0.204**	**0.185**
	0.8	−0.001	−0.020	−0.022	−0.562	−0.027	−0.041	−0.636	−0.035	−0.059
		0.022	**0.050**	**0.053**	**0.562**	**0.074**	**0.081**	**0.636**	**0.099**	**0.101**
		$\alpha_0 = 0.8$								
100	0.2	−0.024	0.141	0.045	−0.150	0.200	0.029	−0.161	0.230	0.010
		0.145	**0.431**	**0.342**	**0.173**	**0.486**	**0.340**	**0.176**	**0.506**	**0.328**
	0.5	−0.059	−0.041	−0.182	−0.424	−0.113	−0.286	−0.446	−0.089	−0.329
		0.217	**0.411**	**0.380**	**0.439**	**0.455**	**0.439**	**0.454**	**0.460**	**0.448**
	0.8	−0.102	−0.309	−0.429	−0.690	−0.379	−0.635	−0.723	−0.415	−0.670
		0.252	**0.487**	**0.545**	**0.708**	**0.594**	**0.710**	**0.735**	**0.618**	**0.728**
10^5	0.2	−0.000	−0.000	−0.009	−0.170	0.075	0.164	−0.185	0.219	0.393
		0.005	**0.098**	**0.123**	**0.170**	**0.362**	**0.382**	**0.185**	**0.494**	**0.575**
	0.5	−0.000	−0.002	−0.003	−0.410	0.010	−0.053	−0.442	−0.043	0.067
		0.006	**0.032**	**0.035**	**0.410**	**0.286**	**0.266**	**0.442**	**0.417**	**0.396**
	0.8	−0.000	−0.020	−0.022	−0.627	−0.031	−0.063	−0.670	−0.076	−0.108
		0.008	**0.050**	**0.054**	**0.627**	**0.127**	**0.118**	**0.670**	**0.273**	**0.219**

Table 3 Bias and **RMSE** of nML, GMM1 and GMM5 estimators for ARCH intercept parameter ω_0 (with $\omega_0 = 1$)

		$a = 0.00$			$a = 0.75$			$a = 1.5$		
		nML	GMM1	GMM5	nML	GMM1	GMM5	nML	GMM1	GMM5
n	β_0	$\alpha_0 = 0.2$								
100	0.2	0.003	−0.045	0.025	1.043	−0.708	−0.460	1.994	−0.607	−0.343
		0.190	**0.776**	**0.512**	**1.099**	**0.927**	**0.902**	**2.054**	**1.001**	**1.052**
	0.5	0.042	0.361	0.462	1.908	−0.531	−0.139	3.521	−0.420	0.128
		0.240	**1.401**	**0.969**	**1.985**	**1.116**	**1.267**	**3.612**	**1.301**	**1.726**
	0.8	0.110	1.001	1.599	4.802	−0.066	0.917	8.823	0.539	2.016
		0.355	**2.175**	**4.624**	**4.957**	**2.468**	**3.345**	**9.012**	**5.836**	**7.898**
10^5	0.2	−0.000	0.000	0.011	1.101	0.072	0.085	2.127	0.117	0.104
		0.007	**0.122**	**0.153**	**1.101**	**0.269**	**0.258**	**2.127**	**0.410**	**0.338**
	0.5	0.000	0.003	0.005	1.905	0.047	0.059	3.590	0.075	0.116
		0.007	**0.055**	**0.062**	**1.905**	**0.245**	**0.244**	**3.590**	**0.412**	**0.396**
	0.8	0.000	0.054	0.072	4.676	0.103	0.165	8.728	0.209	0.320
		0.008	**0.176**	**0.175**	**4.676**	**0.606**	**0.633**	**8.729**	**0.849**	**0.911**
		$\alpha_0 = 0.5$								
100	0.2	0.001	−0.180	−0.051	1.460	−0.294	0.054	2.714	−0.206	0.313
		0.191	**0.584**	**0.475**	**1.521**	**0.979**	**0.989**	**2.782**	**1.252**	**1.466**
	0.5	0.041	0.108	0.356	2.595	0.062	0.639	4.682	0.148	1.082
		0.241	**0.947**	**0.868**	**2.684**	**1.487**	**1.722**	**4.792**	**1.919**	**2.578**
	0.8	0.113	0.820	1.348	6.366	1.009	2.435	11.536	1.389	3.592
		0.403	**1.938**	**2.657**	**6.559**	**3.708**	**4.954**	**11.784**	**4.891**	**7.102**
10^5	0.2	−0.000	−0.000	0.011	1.545	0.050	0.044	2.900	−0.016	−0.033
		0.007	**0.122**	**0.154**	**1.545**	**0.222**	**0.156**	**2.900**	**0.369**	**0.278**
	0.5	0.000	0.002	0.004	2.617	0.006	0.028	4.809	−0.004	0.086
		0.007	**0.055**	**0.062**	**2.617**	**0.185**	**0.174**	**4.809**	**0.414**	**0.377**
	0.8	0.002	0.042	0.051	6.262	0.085	0.151	11.497	0.121	0.238
		0.053	**0.156**	**0.176**	**6.262**	**0.316**	**0.337**	**11.498**	**0.453**	**0.466**
		$\alpha_0 = 0.8$								
100	0.2	−0.002	−0.188	−0.067	3.453	−0.170	0.222	6.240	0.258	1.096
		0.191	**0.563**	**0.468**	**3.541**	**1.100**	**1.227**	**6.354**	**2.031**	**2.573**
	0.5	0.037	0.062	0.316	5.835	0.452	1.077	10.319	1.187	2.588
		0.241	**0.880**	**0.817**	**5.982**	**2.061**	**2.437**	**10.520**	**3.693**	**4.750**
	0.8	0.107	0.743	1.271	13.969	2.189	4.419	24.908	4.305	7.359
		0.371	**1.664**	**4.663**	**14.324**	**6.165**	**8.307**	**25.415**	**10.275**	**12.976**
10^5	0.2	−0.000	−0.000	0.011	3.663	−0.095	−0.205	6.673	−0.271	−0.489
		0.007	**0.122**	**0.153**	**3.663**	**0.453**	**0.478**	**6.673**	**0.620**	**0.720**
	0.5	0.000	0.002	0.004	5.979	−0.020	0.105	10.721	0.092	−0.129
		0.007	**0.055**	**0.063**	**5.980**	**0.571**	**0.530**	**10.722**	**0.842**	**0.799**
	0.8	0.000	0.039	0.046	13.859	0.087	0.235	24.931	0.306	0.461
		0.008	**0.160**	**0.184**	**13.866**	**0.563**	**0.490**	**24.932**	**1.249**	**1.006**

MLE at sample size $n = 100$, while the bias reduces markedly at sample size $n = 10^5$. However, when the measurement error is present ($a = 0.75$ or $a = 1.5$), the GMMs has significantly smaller bias than the naive MLE at all sample sizes, and significantly smaller RMSE at $n = 10^5$. In particular, while the bias in GMMs reduces significantly at large sample size $n = 10^5$, the bias in naive MLE remains persistently at high level. Overall, the GMM5 using $k_1 = k_2 = 5$ instruments have smaller RMSE but larger bias than the GMM1 using $k_1 = k_2 = 1$ instrument. In general, the AR parameter α_0 has the smallest bias and RMSE, while the ARCH intercept ω_0 has the largest values.

In the following we provide a more detailed summary of findings for each parameter based on over 300 various configurations of parameter values.

AR Parameter α

The nML estimator is clearly downward biased and its absolute relative bias (ARB) is fast increasing (from 20% to 80%) with the noise-to-signal ratio a. The GMM estimators are downward biased in small samples and their ARB are decreasing with α_0 but increasing with β_0. While the ARB of GMM1 has no relation with a that of GMM5 is a fast increasing function of a. The RMSE of the nML estimator has, respectively, a shape of square-root function in a, a clear increasing linear function in α_0, and a slightly increasing linear function of β_0. In contrast, the RMSE of the GMM estimators are decreasing with α_0, but increasing with β_0 and a, respectively. However, the RMSE of GMM5 vanishes in large samples faster than that of GMM1 estimator.

ARCH slope β

Again the nML estimator is clearly downward biased in large samples and most of small sample cases, and its ARB is fast increasing with a (from 40% to 80%). In small samples, the biases of the GMM estimators have a shape of concave function with respect to β_0, while their ARB have a shape of convex function. Furthermore, when the sample size increases the GMM biases vanish very slowly for large α_0 (0.95) and small β_0 (0.05). The RMSE of the nML estimator takes the shape of a square-root function in a, a fast increasing linear function in β_0, has no relation to α_0. The RMSE of the GMM estimators are a convex function of β_0 but have no relation with α_0 or a.

ARCH intercept ω

The nML estimator has an upward bias and the bias is increasing with a, α_0, and β_0, respectively. In small samples the biases of the GMM estimators have, respectively, the shape of a linear function in α_0, an increasing function in β_0, and a fast increasing function in a. However, the bias vanishes slowly when the sample size increases. The RMSE of all three estimators have similar patterns as their respective biases.

7 Conclusions and Discussion

We have studied a dynamic model with autoregressive heteroscedastic error where the underlying process is latent and subject to additive measurement error. We have shown that the model is identifiable by using the observations of a proxy process only. This is in contrast to the case of measurement error in the covariates, where extra information such as external instrumental variables or replicate observations is needed for model identifiability. Moreover, we proposed a set of identifying moment conditions and used them to construct GMM estimators for the unknown parameters. The proposed estimators are consistent and asymptotically normally distributed under usual regularity conditions. As a byproduct, this framework allows us to construct a test for the presence of measurement error. Our Monte Carlo simulation studies show that the measurement error causes downward bias in the naive MLE, and the relative biases have certain functional forms. This is interesting because it provides a possibility to find the formulas that can be used to correct the biases in the naive MLE. Furthermore, the proposed estimators possess fairly good finite sample properties and comparisons with the naive MLE are also presented.

This work attempts to address some basic measurement error problems in dynamic models with ARCH-type errors. There are many more questions and issues remaining to be investigated. For example, it would be interesting to explore other possible moment conditions that can be used to achieve identification and to obtain more efficient estimators. It would also be interesting to study more general measurement error processes. We used a simple ARCH model in order to be able to gain insights of the problem and to obtain some concrete results. From both theoretical and practical point of view, it is important to investigate the measurement error problem in more general GARCH models. Our theoretical framework should apply to GARCH processes as well, but the estimation will be based on a different set of moments than (5)–(7) used here. Another way is to convert the GARCH process to an infinite order ARCH and then truncate it to finite order, so that the estimators based on the moments (5)–(7) can be used directly.

Acknowledgments The research was partially supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Appendix

Regularity Assumptions and Mathematical Proofs

In this Appendix we provide the regularity assumptions that are sufficient for the theoretical results in Theorems 1 and 2. We also provide a sketch of the proofs of them, since they follow the general framework of GMM estimation.

Regularity Assumptions

For the asymptotic properties of the GMM estimator $\hat{\boldsymbol{\theta}}_n$ we make the following assumptions, where $\|\cdot\|$ denotes the Euclidean norm.

Assumption 1 *The instrumental functions satisfy* $E\left\|Z_{t-h}f_1(\tilde{Z}_{t-p-1})\right\| < \infty$ *for* $h = 0, 1, \ldots, p$ *and* $E\left\|Z_{t-h}^2 f_2(\tilde{Z}_{t-p-q-1})\right\| < \infty$ *for* $h = 0, 1, \ldots, p+q$.

Assumption 2 $\boldsymbol{\theta}_0$ *is the unique solution to* $E\{\boldsymbol{G}_0\boldsymbol{r}_0(\boldsymbol{\theta})\} = 0$ *in* $\boldsymbol{\Theta}$.

Assumption 3 *The covariance matrix* $V[n^{-1/2}\sum_{t=1}^{n}\boldsymbol{G}_t\boldsymbol{r}_t(\boldsymbol{\theta_0})]$ *is uniformly positive definite.*

Assumption 4 *The instrumental functions satisfy* $E\left\|Z_{t-h}^2 f_1^2(\tilde{Z}_{t-p-1})\right\| < \infty$ *for* $h = 0, 1, 2, \ldots, p$, $E\left\|Z_{t-h}^4 f_2^2(\tilde{Z}_{t-p-q-1})\right\| < \infty$ *for* $h = 0, 1, 2, \ldots, p+q$, *and* $E(Z_0^4) < \infty$.

Assumption 5 $E[\left\|E(\boldsymbol{G}_0\boldsymbol{r}_0(\boldsymbol{\theta}_0)|\mathcal{F}_{-j})\right\|^2] < \infty$, $j = 1, 2, \ldots, p+q$, *where* $\mathcal{F}_t = \sigma(Z_s, s \le t)$.

Assumption 6 $\boldsymbol{\Delta}' = E[\nabla_{\boldsymbol{\theta}}\boldsymbol{r}_0'(\boldsymbol{\theta}_0)\boldsymbol{G}_0']$ *has full rank* $p+q+2$, *where* $\nabla_{\boldsymbol{\theta}}\boldsymbol{r}_0'(\boldsymbol{\theta}) = \partial\boldsymbol{r}_0'(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$.

Note that the above assumptions are not more restrictive than the usual assumptions for the asymptotic properties of GMM estimators in the literature. They are formulated for the general forms of the instrumental functions f_1 and f_2 (which are also on the diagonal of matrix $\boldsymbol{G}_t$). For example, if f_1 and f_2 are taken to be the linear projections of the lagged Z_t, then Assumption 1 simply means the Z_t process has finite second and third moments. Similarly, Assumption 4 means Z_t has finite fourth and sixth moments. In particular, the identifiability Assumption 3 is based on the moment conditions (5)–(7) which is given in $\boldsymbol{r}_t(\boldsymbol{\theta})$. Again, if f_1 and f_2 are taken to be the linear projections then this assumption follows directly from (5)–(7).

Proof of Theorem 1

To simplify notation in the following we will omit the subscript n in $\hat{\boldsymbol{\theta}}_n$ and denote it as $\hat{\boldsymbol{\theta}}$. First, since $\{X_t\}$ is strictly stationary and ergodic, and $\{\delta_t\}$ is iid and independent of $\{\eta_t\}$, $\{Z_t\}$ is strictly stationary and ergodic. It follows from Assumption 1 and White (1996, Th. A.2.2) that $E[\boldsymbol{G}_t\boldsymbol{r}_t(\boldsymbol{\theta})]$ is continuous on Θ and, furthermore, by the strong uniform law of large numbers (ULLN), as $n \to \infty$,

$$\sup_{\Theta} \left\| \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{G}_t \boldsymbol{r}_t(\boldsymbol{\theta}) - E[\boldsymbol{G}_0 \boldsymbol{r}_0(\boldsymbol{\theta})] \right\| \xrightarrow{a.s.} 0. \tag{26}$$

The result (1) follows then from Assumption 2 and White (1996, Th. 3.4).

To prove the asymptotic normality, note that for sufficiently large n, the score

$$S_n(\hat{\boldsymbol{\theta}}, \Lambda_n) = \left[\sum_{t=1}^{n} \nabla_{\boldsymbol{\theta}} \boldsymbol{r}_t'(\hat{\boldsymbol{\theta}}) \boldsymbol{G}_t' \right] \Lambda_n \left[\sum_{t=1}^{n} \boldsymbol{G}_t \boldsymbol{r}_t(\hat{\boldsymbol{\theta}}) \right] = \mathbf{0}, \quad w.p.1, \tag{27}$$

where $\nabla_{\boldsymbol{\theta}} \boldsymbol{r}_t'(\boldsymbol{\theta}) = \partial \boldsymbol{r}_t'(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}$. Then using the mean-value theorem (Jennrich 1969), we have

$$\left[\sum_{t=1}^{n} \nabla_{\boldsymbol{\theta}} \boldsymbol{r}_t'(\hat{\boldsymbol{\theta}}) \boldsymbol{G}_t' \right] \Lambda_n \left[\sum_{t=1}^{n} \boldsymbol{G}_t \boldsymbol{r}_t(\boldsymbol{\theta}_0) \right] = - \left[\sum_{t=1}^{n} \nabla_{\boldsymbol{\theta}} \boldsymbol{r}_t'(\hat{\boldsymbol{\theta}}) \boldsymbol{G}_t' \right] \Lambda_n \left[\sum_{t=1}^{n} \boldsymbol{G}_t \nabla_{\boldsymbol{\theta}} \boldsymbol{r}_t(\tilde{\boldsymbol{\theta}}) \right] (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0), \tag{28}$$

where $\left\| \tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \right\| \le \left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \right\|$. Again by the ULLN (White 1996, Th. A.2.2 and Cor. 3.8), we have, as $n \to \infty$,

$$\frac{1}{n} \sum_{t=1}^{n} \nabla_{\boldsymbol{\theta}} \boldsymbol{r}_t'(\hat{\boldsymbol{\theta}}) \boldsymbol{G}_t' \xrightarrow{a.s.} E\left[\nabla_{\boldsymbol{\theta}} \boldsymbol{r}_0'(\boldsymbol{\theta}_0) \boldsymbol{G}_0' \right] \tag{29}$$

which has full rank by Assumption 6. Further, by Assumption 3–5 and (White 1996, Th. A.3.2), we can use the so-called Cramer–Wold device (Rao 1973) to show that

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \boldsymbol{G}_t \boldsymbol{r}_t(\boldsymbol{\theta}_0) \xrightarrow{d} N(0, \boldsymbol{\Sigma}), \tag{30}$$

where $\boldsymbol{\Sigma} = E[\boldsymbol{G}_t \boldsymbol{r}_t(\boldsymbol{\theta}_0) \boldsymbol{r}_t'(\boldsymbol{\theta}_0) \boldsymbol{G}_t']$. Finally the result follows from (28)–(30) and Assumption 6.

Proof of Theorem 2

First, using the nonsingular factorization we can write

$$\boldsymbol{V}_1 = \operatorname*{plim}_{n \to \infty} V[n^{1/2} \boldsymbol{b}_{1n}(\boldsymbol{\gamma}_0)] = \boldsymbol{C}_1 \boldsymbol{C}_1'$$

and $\boldsymbol{V}_{1n} = \boldsymbol{C}_{1n} \boldsymbol{C}_{1n}'$ such that $\boldsymbol{C}_1 = \operatorname{plim}_{n \to \infty} \boldsymbol{C}_{1n}$. Then by the mean-value theorem and Slutsky's theorem we have

$$\ell_1 = n^{1/2} \boldsymbol{C}_{1n}' \boldsymbol{b}_{1n}(\hat{\boldsymbol{\gamma}}_1) = n^{1/2} \boldsymbol{M}_1 \boldsymbol{C}_1' \boldsymbol{b}_{1n}(\boldsymbol{\gamma}_0) + o_p(1), \tag{31}$$

where $\boldsymbol{M}_1 = \boldsymbol{I}_{k_1+k_2} - \boldsymbol{A}_1(\boldsymbol{A}_1'\boldsymbol{A}_1)^{-1}\boldsymbol{A}_1'$, $\boldsymbol{A}_1 = \boldsymbol{C}_1'\boldsymbol{D}_1$, and $\boldsymbol{D}_1' = E[\nabla_{\boldsymbol{\gamma}}\tilde{\boldsymbol{r}}_t'(\boldsymbol{\gamma}_0)\boldsymbol{G}_{1t}']$. Similarly, let $\boldsymbol{V}_n = \boldsymbol{C}_n\boldsymbol{C}_n'$, where

$$\boldsymbol{C}_n' = \begin{pmatrix} \boldsymbol{C}_{1n}' & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_{2n}' \end{pmatrix} \begin{pmatrix} \boldsymbol{I}_{k_1+k_2} & \boldsymbol{0} \\ -\boldsymbol{H}_n & \boldsymbol{I}_{2p+q} \end{pmatrix},$$

$\boldsymbol{C}_{2n}\boldsymbol{C}_{2n}' = V^{-1}[n^{1/2}(\boldsymbol{b}_{2n}(\boldsymbol{\gamma}_0) - \boldsymbol{H}_n\boldsymbol{b}_{1n}(\boldsymbol{\gamma}_0))]$, and

$$\operatorname*{plim}_{n\to\infty} \boldsymbol{H}_n = E[\boldsymbol{b}_{2n}(\boldsymbol{\gamma}_0)\boldsymbol{b}_{1n}(\boldsymbol{\gamma}_0)']E^{-1}[\boldsymbol{b}_{1n}(\boldsymbol{\gamma}_0)\boldsymbol{b}_{1n}(\boldsymbol{\gamma}_0)'].$$

Then it is easy to show that $\operatorname{plim}_{n\to\infty} \boldsymbol{C}_n\boldsymbol{C}_n' = \lim_{n\to\infty} V^{-1}[n^{1/2}\boldsymbol{b}_n(\boldsymbol{\gamma}_0)]$, and similarly to (31), we have

$$\ell = n^{1/2}\boldsymbol{C}_n'\boldsymbol{b}_n(\hat{\boldsymbol{\gamma}}) = n^{1/2}\boldsymbol{M}\boldsymbol{C}'\boldsymbol{b}_n(\boldsymbol{\gamma}_0) + o_p(1), \tag{32}$$

where $\boldsymbol{M} = \boldsymbol{I}_{k_1+k_2+2p+q} - \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'$, $\boldsymbol{C} = \operatorname{plim}_{n\to\infty} \boldsymbol{C}_n$, $\boldsymbol{A} = \boldsymbol{C}'\boldsymbol{D}$, $\boldsymbol{D}' = (\boldsymbol{D}_1', \boldsymbol{D}_2')$, and $\boldsymbol{D}_2' = E[\nabla_{\boldsymbol{\gamma}}\tilde{\boldsymbol{r}}_t'(\boldsymbol{\gamma}_0)\boldsymbol{G}_{2t}']$. Further, denote

$$\boldsymbol{M}_2 = \begin{pmatrix} \boldsymbol{M}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}.$$

Then out test statistic is

$$SW = n\boldsymbol{b}_n'(\boldsymbol{\gamma}_0)\boldsymbol{C}(\boldsymbol{M} - \boldsymbol{M}_2)\boldsymbol{C}'\boldsymbol{b}_n(\boldsymbol{\gamma}_0) + o_p(1). \tag{33}$$

Finally, since clearly $(\boldsymbol{M} - \boldsymbol{M}_2)\boldsymbol{M}_2 = 0$ and $n^{1/2}\boldsymbol{C}'\boldsymbol{b}_n(\boldsymbol{\gamma}_0) \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{I}_{k_1+k_2+2p+q})$ under H_0, we have $SW \xrightarrow{d} \chi^2_{2p+q}$.

References

Alberini, A., & Filippini, M. (2011). Response of residential electricity demand to price: The effect of measurement error. *Energy Economics, 33*(5), 889–895.

Arellano, M. (2003). *Panel data econometrics*. Oxford: Oxford University Press.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*(3), 307–327.

Buonaccorsi, J. P. (2010). *Measurement Error: Models, Methods, and Applications*. New York: CRC Press.

Buonaccorsi, J. P. (2013). Measurement error in dynamic models. In *ISS-2012 Proceedings Volume On Longitudinal Data Analysis Subject to Measurement Errors, Missing Values, and/or Outliers* (pp. 53–76). Berlin: Springer.

Caporale, G. M., Onorante, L., & Paesani, P. (2012). Inflation and inflation uncertainty in the Euro area. *Empirical Economics, 43*(2), 597–615.

Carroll, R. J., Ruppert, D., Crainiceanu, C. M., & Stefanski, L. A. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*. West Palm Beach: Chapman and Hall/CRC.

Chen, X., Hong, H., & Nekipelov, D. (2011). Nonlinear models of measurement errors. *Journal of Economic Literature, 49*(4), 901–937.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica, 50*(4), 987–1007.

Engle, R. F., Focardi, S. M., & Fabozzi, F. J. (2008). ARCH/GARCH models in applied financial econometrics. In *Handbook of Finance, 3*.

Fan, J., & Wang, Y. (2007). Multi-scale jump and volatility analysis for high-frequency financial data. *Journal of the American Statistical Association, 102*(480), 1349–1362.

Fang, W., & Miller, S. M. (2009). Modeling the volatility of real GDP growth: The case of Japan revisited. *Japan and the World Economy, 21*(3), 312–324.

Francq, C., & Zakoïan, J.-M. (2000). Estimating weak GARCH representations. *Econometric Theory, 16*(5), 692–728.

Francq, C., & Zakoian, J.-M. (2011). *GARCH Models: Structure, Statistical Inference and Financial Applications*. New York: Wiley.

Gourieroux, C., Monfort, A., & Renault, E. (1993). Indirect inference. *Journal of applied econometrics, 8*(S1), S85–S118.

Grier, K. B., & Perry, M. J. (2000). The effects of real and nominal uncertainty on inflation and output growth: some GARCH-M evidence. *Journal of Applied Econometrics, 15*(1), 45–58.

Handbury, J., Watanabe, T., & Weinstein, D. E. (2013). How much do official price indexes tell us about inflation? Technical report, National Bureau of Economic Research.

Harvey, A., Ruiz, E., & Sentana, E. (1992). Unobserved component time series models with arch disturbances. *Journal of Econometrics, 52*(1–2), 129–157.

Jennrich, R. I. (1969). Asymptotic properties of non-linear least squares estimators. *The Annals of Mathematical Statistics, 40*(2), 633–643.

Ramirez, O. A., & Fadiga, M. (2003). Forecasting agricultural commodity prices with asymmetric-error GARCH models. *Journal of Agricultural and Resource Economics, 28*(1), 71–85.

Rao, C. R. (1973). Linear Statistical Inference and Its Applications. New York: Wiley.

Reitz, S., & Westerhoff, F. (2007). Commodity price cycles and heterogeneous speculators: a STAR–GARCH model. *Empirical Economics, 33*(2), 231–244.

Roche, M. J., & McQuinn, K. (2003). Grain price volatility in a small open economy. *European Review of Agricultural Economics, 30*(1), 77–98.

Staudenmayer, J., & Buonaccorsi, J. P. (2005). Measurement error in linear autoregressive models. *Journal of the American Statistical Association, 100*(471), 841–852.

Teräsvirta, T. (2009). An introduction to univariate GARCH models. In *Handbook of Financial Time Series* (pp. 17–42). Berlin: Springer.

Wang, L., & Hsiao, C. (2011). Method of moments estimation and identifiability of semiparametric nonlinear errors-in-variables models. *Journal of Econometrics, 165*(1), 30–44.

Wansbeek, T. J., & Meijer, E. (2000). *Measurement Error and Latent Variables in Econometrics*, vol. 37. Amsterdam: Elsevier.

White, H. (1996). *Estimation, Inference and Specification Analysis*. Cambridge: Cambridge University Press.

White, H. (2001). Asymptotic Theory for Econometricians. New York: Academic Press.

Wooldridge, J. M. (1994). Estimation and inference for dependent processes. In Handbook of Econometrics, vol. 4, pp 2639-2738.

Yi, G. Y., Delaigle, A., & Gustafson, P. (2021). *Handbook of Measurement Error Models*. London: Chapman and Hall/CRC.