**Discussion Paper Series**

No.82

**Small Sample Bias Properties of the System GMM
Estimator in Dynamic Panel Data Models**

Kazuhiko Hayakawa

April 2005

# Small Sample Bias Properties of the System GMM Estimator in Dynamic Panel Data Models

Kazuhiko Hayakawa [*][†]

Graduate School of Economics

Hitotsubashi University

April, 2005

## Abstract

This paper examines analytically and experimentally why the system GMM estimator in dynamic panel data models is less biased than the first differencing or the level estimators even though the former uses more instruments. We find that the bias of the system GMM estimator is a weighted sum of the biases in opposite directions of the first differencing and the level estimator. We also find that an important condition for the system GMM estimator to have small bias is that the variances of the individual effects and the disturbances are almost of the same magnitude. If the variance of individual effects is much larger than that of disturbances, then all GMM estimators are heavily biased. To reduce such biases, we propose bias-corrected GMM estimators. On the other hand, if the variance of individual effects is smaller than that of disturbances, the system estimator has a more severe downward bias than the level estimator.

# 1    Introduction

A major topic in recent theoretical analyses of panel data models is the estimation of dynamic panel data models.[1] Since the work of Arellano and Bond (1991), the generalized method of moments (GMM) technique has been widely used in the estimation of dynamic panel data models. However, subsequent examinations of the finite sample performance of the GMM estimator have shown that it is substantially biased. One source of the bias, first discovered by Nelson and Startz (1990a, b), is weak instruments. Staiger and Stock (1997) have shown that the instrumental variables estimator would be inconsistent under weak instrument asymptotics.[2] We call this the "weak instruments problem". The other source of bias is the relative number of instruments to sample size. Especially in linear simultaneous equation models, Kunitomo (1980), Morimune (1983) and Bekker (1994) have shown that the 2SLS estimator is inconsistent as the number of the instruments tends to infinity. And Hahn and Hausman (2002) have shown that the finite sample bias of 2SLS estimator is monotonically increasing in the number of instruments. One important finding of the papers listed above is that the magnitude of the bias is proportional to the relative size of the number of instruments to the sample size. We label this the "many instruments problem". What is important with regards to using many instruments is the trade-off between the efficiency and the bias of the estimator. Although using many instruments is desirable to improve the efficiency of the estimator in terms of conventional first order asymptotic theory, it is problematic in terms of bias. This trade-off in GMM estimation was first highlighted by Tauchen (1986) for the case of time series data, and Donald and Newey(2001) proposed a method to determine the number of instruments that minimizes the mean squared error of the estimator. [3]

These considerations are still binding in the estimation of dynamic panel data

---

[1]For an overview of recent developments in this field, see Baltagi (2001), Hsiao (2003), Arellano (2003) and Blundell, Bond and Windmeijer (2000).

[2]See Stock, Wright and Yogo (2002) for an excellent survey.

[3]Ziliak (1997) has shown by simulation that the trade-off between the efficiency and the bias of the estimator is still binding in static panel data models with predetermined variables, while Okui (2004) proposed a procedure for the optimal choice of instruments in dynamic panel data models.

models. [4] It is well known that in the first differencing models the bias is sizable when the parameter concerning the lagged dependent variable is close to unity (Alonso-Borrego and Arellano 1999). Blundell and Bond (1998) have shown that this is due to weak instruments, and to overcome the "weak instruments problem" they proposed the system GMM estimator. They first showed that the level GMM estimators by Arellano and Bover (1995) are free from weak instruments when the parameter concerning the lagged variable is close to unity and then combined the moment conditions which are used in the first differencing and the level GMM estimators to improve the efficiency of the estimator. The system GMM estimator is becoming widely used in empirical analyses. Empirical applications include the estimation of production functions (Blundell and Bond 2000) and empirical growth models (Bond, Hoeffler and Temple 2001), among others.

However, since the system GMM estimator uses more instruments than the first differencing and the level estimators, even in the fixed $N$ and $T$ case, the number of instruments relative to the sample size of the system GMM estimator is larger than that of the first differencing and the level estimators. For example, when T=4, the first differencing and the level estimators utilize 3 and 2 instruments respectively, while the system estimator uses 5; thus, the number of instruments relative to the sample size is largest, and the "many instruments problem" is most serious for the system GMM estimator. Therefore, one might suspect that the system GMM estimator is more biased than the first differencing and the level estimators, although it is more efficient.

However some simulation results, such as those presented in Table 5 in Blundell and Bond (1998) or in Figure 1 in this paper, [5] do not pose such a problem; rather the system GMM estimator is less biased than the first differencing and the level

---

[4]A number of papers have suggesting ways to reduce the bias of the estimator in dynamic panel data models. Kiviet (1995) and Hahn and Kuersteiner (2002) propose bias-corrected within-groups estimators, while Hahn, Hausman and Kuersteiner (2001) propose a long-differenced estimator to strengthen the instruments.

[5]In the figures through the paper, the horizontal axis represents the autoregressive parameter $\alpha$ (which is defined in Section 2) from 0.1 to 0.9, while the vertical axis represents the size of the bias relative to the true $\alpha$ multiplied by 100. "-th" denotes theoretical values and "-sml" denotes simulation values based on 10,000 replications. The simulation design is explained in Section 4.

GMM estimators even in moderate sample sizes of cross-section data. The purpose of this paper is to show theoretically why the system GMM estimator, despite using more instruments, is less biased than the first differencing and level GMM estimators.

We find that the bias of the system GMM estimator is composed of a weighted sum of the biases of the first differencing and the level estimators with opposite direction, and this is the primary reason why the system estimator is less biased. Also, we find that the size of the variances of the individual effects relative to the disturbances is an important factor determining the magnitude of the biases. If the variances of the individual effects and disturbances are equal, as considered in Blundell and Bond (1998), the bias of the system estimator is very small. However in other cases with unequal variances of the individual effects and disturbances, the bias of the system estimator is fairly large. Hence, we can say that the simulation results reported by Blundell and Bond (1998) are special cases in which the bias becomes small.

The remainder of the paper is organized as follows. In Section 2, we define the model and the estimators used in this paper. For simplicity, we consider inefficient one-step GMM estimators with $T = 4$. Section 3 provides the main results concerning the small sample bias of the GMM estimators. Based on the results of Section 3, we compare the estimators numerically in Section 4. We find that all the GMM estimators are heavily biased when the variance of the individual effects is larger than that of the disturbances. Section 5 suggests bias-corrected GMM estimators to reduce such bias. Finally, Section 6 concludes.

## 2   The model and estimators

We consider an AR(1) panel data model given by

$$y_{it} = \alpha y_{i,t-1} + \eta_i + v_{it} \quad i = 1, ..., N; \quad t = 2, 3, 4 \tag{1}$$

where $\alpha$ is the parameter of interest with $|\alpha| < 1$ and $v_{it}$ has mean zero given $\eta_i, y_{i1}, ..., y_{i,t-1}$. Let $u_{it} = \eta_i + v_{it}$. We impose the following assumptions:

**Assumption 1.** $\{v_{it}\}$ $(t = 2, 3, 4; i = 1, ..., N)$ *are i.i.d across time and individuals*

*and independent of $\eta_i$ and $y_{i1}$ with $E(v_{it}) = 0$, $var(v_{it}) = \sigma_v^2$ and finite moments up to third order.*

**Assumption 2.** *$\eta_i$ are i.i.d across individuals with $E(\eta_i) = 0$, $var(\eta_i) = \sigma_\eta^2$, and finite third order moments*

**Assumption 3.** *The initial observations satisfy*

$$y_{i1} = \frac{\eta_i}{1-\alpha} + w_{i1} \qquad for \quad i = 1, ..., N \tag{2}$$

*where $w_{i1}$ is $w_{i1} = \sum_{j=0}^{\infty} \alpha^j v_{i,1-j}$ and independent of $\eta_i$.*

These assumptions are the same as Alvarez and Arellano's (2003) except that we assume finite third order moments.

## Linear GMM estimators

Given these assumptions, we consider three types of GMM estimators. These include the first differencing GMM estimator, the level GMM estimator and the system GMM estimator, abbreviated as GMM(DIF), GMM(LEV) and GMM(SYS) estimators, respectively.

### First differencing GMM estimator

Arellano and Bond (1991) proposed a first differencing GMM estimator which is based on the 3 moment conditions:

$$E[Z'_{d,i} \Delta u_i] = 0 \tag{3}$$

where

$$Z_{d,i} = \begin{bmatrix} y_{i1} & 0 & 0 \\ 0 & y_{i1} & y_{i2} \end{bmatrix} \qquad \Delta u_i = \begin{bmatrix} \Delta u_{i3} \\ \Delta u_{i4} \end{bmatrix} \tag{4}$$

and $\Delta u_{it} = u_{it} - u_{i,t-1}$. We consider a one-step GMM estimator based on moment condition (3) with weighting matrix $Z'_d Z_d$, where $Z'_d$ is $(Z'_{d,1}, \cdots, Z'_{d,N})$. Let $\Delta y'_i = (\Delta y_{i3}, \Delta y_{i4})'$, $\Delta y'_{i,-1} = (\Delta y_{i2}, \Delta y_{i3})'$, where $\Delta y$ and $\Delta y_{-1}$ are stacked across individuals. Then the one-step GMM estimator for $\alpha$ is

$$\widehat{\alpha}^{dif} = \left( \Delta y'_{-1} Z_d (Z'_d Z_d)^{-1} Z'_d \Delta y_{-1} \right)^{-1} \Delta y'_{-1} Z_d (Z'_d Z_d)^{-1} Z'_d \Delta y \tag{5}$$

## Level GMM estimator

Arellano and Bover (1995) proposed a level GMM estimator that is based on the moment conditions: [6]

$$E(Z'_{l,i}u_i) = 0 \tag{6}$$

where

$$Z_{l,i} = \begin{bmatrix} \Delta y_{i2} & 0 \\ 0 & \Delta y_{i3} \end{bmatrix} \quad u_i = \begin{bmatrix} u_{i3} \\ u_{i4} \end{bmatrix} \tag{7}$$

We consider a one-step GMM estimators based on moment conditions (6) with weighting matrix $Z'_l Z_l$, where $Z'_l = (Z'_{l,1}, \cdots, Z'_{l,N})$. Let $y'_i = (y_{i3}, y_{i4})'$, $y'_{i,-1} = (y_{i2}, y_{i3})'$, where $y$ and $y_{-1}$ are stacked across individuals. Then the one-step GMM estimator for $\alpha$ is

$$\widehat{\alpha}^{lev} = \left( y'_{-1} Z_l (Z'_l Z_l)^{-1} Z'_l y_{-1} \right)^{-1} y'_{-1} Z_l (Z'_l Z_l)^{-1} Z'_l y \tag{8}$$

We label (8) the GMM(LEV) estimator.

## System GMM estimator

In order to avoid weak instruments and improve the efficiency of the estimator, Blundell and Bond (1998) proposed a system GMM estimator in which the moment conditions of the GMM(DIF) and GMM(LEV) are used jointly. The moment conditions used in constructing the system GMM estimator are given by

$$E(Z'_{s,i}p_i) = 0 \tag{9}$$

where

$$Z_{s,i} = \begin{bmatrix} Z_{d,i} & 0 \\ 0 & Z_{l,i} \end{bmatrix} \quad p_i = \begin{bmatrix} \Delta u_i \\ u_i \end{bmatrix} \tag{10}$$

Let $q'_i = (\Delta y'_i \ y'_i)'$, $q'_{i,-1} = (\Delta y'_{i,-1} \ y'_{i,-1})'$ where $q$, $q_{-1}$ are stacked across individuals. Then the one-step GMM estimators for $\alpha$ based on moment condition (9) are

$$\widehat{\alpha}^{sys} = \left( q'_{-1} Z_s (Z'_s Z_s)^{-1} Z'_s q_{-1} \right)^{-1} q'_{-1} Z_s (Z'_s Z_s)^{-1} Z'_s q \tag{11}$$

---

[6]Although $\Delta y_{i2}$ is also available as an instrument for $t = 4$, we do not use it because it becomes redundant in the system estimation. See Blundell and Bond (1998).

where $Z'_s = (Z'_{s,1}, \cdots, Z'_{s,N})$. We label (11) the GMM(SYS) estimator. Note that $\hat{\alpha}^{sys}$ can be expressed as

$$\hat{\alpha}^{sys} = \hat{\gamma}\hat{\alpha}^{dif} + (1 - \hat{\gamma})\hat{\alpha}^{lev} \tag{12}$$

where

$$\hat{\gamma} = \frac{\Delta y'_{-1} Z_d (Z'_d Z_d)^{-1} Z'_d \Delta y_{-1}}{\Delta y'_{-1} Z_d (Z'_d Z_d)^{-1} Z'_d \Delta y_{-1} + y'_{-1} Z_l (Z'_l Z_l)^{-1} Z'_l y_{-1}} \tag{13}$$

# 3 Small sample biases of GMM estimators

In this section, we provide analytical forms of the biases of the GMM(DIF), GMM(LEV) and GMM(SYS) estimators. All proofs are given in the Appendix.

First we define $\pi_d$ and $\pi_l$ as

$$
\begin{aligned}
\pi_d &= E(Z'_{d,i} Z_{d,i})^{-1} E(Z'_{d,i} \Delta y_{i,-1}) \\
&= \left[ \begin{array}{ccc} \frac{-\sigma_v^2}{(1+\alpha)(C+D)} & \frac{\sigma_v^2(1-\alpha)C}{(1+\alpha)F} & \frac{\sigma_v^2(\alpha-1)}{(\alpha+1)F}[C + (\alpha+1)D] \end{array} \right]' \\
&= \left[ \begin{array}{ccc} \pi_{d,1} & \pi_{d,2} & \pi_{d,3} \end{array} \right]'
\end{aligned}
$$

$$
\begin{aligned}
\pi_l &= E(Z'_{l,i} Z_{l,i})^{-1} E(Z'_{l,i} y_{i,-1}) \\
&= \left[ \begin{array}{cc} \frac{1}{2} & \frac{1}{2} \end{array} \right]' \\
&= \left[ \begin{array}{cc} \pi_{l,1} & \pi_{l,2} \end{array} \right]'
\end{aligned}
$$

where

$$
\begin{aligned}
C &= \left(\frac{1}{1-\alpha}\right)^2 \sigma_\eta^2 \qquad D = \frac{1}{1-\alpha^2}\sigma_v^2 \\
F &= \left[\frac{1}{1+\alpha}\sigma_v^2\right]\left[2\left(\frac{1}{1-\alpha}\right)^2 \sigma_\eta^2 + \left(\frac{1}{1-\alpha}\right)\sigma_v^2\right]
\end{aligned}
$$

Also, we define $\phi_d$, $\phi_l$ and $\gamma$ as

$$
\begin{aligned}
\phi_d &= E(\Delta y'_{i,-1} Z_{d,i}) E(Z'_{d,i} Z_{d,i})^{-1} E(Z'_{d,i} \Delta y_{i,-1}) = \pi'_d E(Z'_{d,i} Z_{d,i})\pi_d \\
&= (\pi_{d,1}^2 + \pi_{d,2}^2 + \pi_{d,3}^2)(C + D) + 2\pi_{d,2}\pi_{d,3}(C + \alpha D)
\end{aligned}
$$

$$
\begin{aligned}
\phi_l &= E(y'_{i,-1} Z_{l,i}) E(Z'_{l,i} Z_{l,i})^{-1} E(Z'_{l,i} y_{i,-1}) = \pi'_l E(Z'_{l,i} Z_{l,i})\pi_l \\
&= \frac{\sigma_v^2}{1+\alpha}
\end{aligned}
$$

6

$$\gamma = \frac{\phi_d}{\phi_d + \phi_l}$$

Next, we provide the formulas for the small sample biases of $\hat{\alpha}^{dif}$, $\hat{\alpha}^{lev}$ and $\hat{\alpha}^{sys}$. The notations are explained in the appendix.

**Theorem 1.** *The second order bias of $\hat{\alpha}^{dif}$ is given by*

$$
\begin{aligned}
N \cdot Bias(\hat{\alpha}^{dif}) &= B_1^{dif} + B_2^{dif} + B_3^{dif} + B_4^{dif} \\
&= -\frac{\sigma_v^2}{\phi_d}\left[1 + \frac{2(C+D)^2}{F} - \frac{2(C+\alpha D)^2}{F}\right] \\
&\quad + \frac{2\sigma_v^2}{\phi_d^2}\left[\left\{\pi_{d,1}^2 + \pi_{d,2}^2 + \pi_{d,3}^2 + (\alpha-2)\pi_{d,1}\pi_{d,2}\right\}(C+D)\right. \\
&\qquad\qquad\left. -\pi_{d,1}\pi_{d,3}\left\{(2-\alpha)C - \alpha(2\alpha-3)D\right\} + 2\pi_{d,2}\pi_{d,3}(C+\alpha D)\right] \\
&\quad - \frac{2\sigma_v^2}{\phi_d^2}\left[\pi_{d,1}\pi_{d,3}^2(C+\alpha D) + \pi_{d,1}\pi_{d,2}\pi_{d,3}(C+D)\right]
\end{aligned}
$$

**Theorem 2.** *The second order bias of $\hat{\alpha}^{lev}$ is given by*

$$
\begin{aligned}
N \cdot Bias(\hat{\alpha}^{lev}) &= B_1^{lev} + B_2^{lev} + B_3^{lev} + B_4^{lev} \\
&= \frac{2\sigma_\eta^2}{\phi_l(1-\alpha)} - \frac{\sigma_v^2}{\phi_l^2}\left(\frac{2\alpha-1}{2(1+\alpha)}\sigma_v^2 - \frac{1}{1-\alpha}\sigma_\eta^2\right) + \frac{\alpha-1}{4\phi_l^2(\alpha+1)}\sigma_v^4
\end{aligned}
$$

**Theorem 3.** *The second order bias of $\hat{\alpha}^{sys}$ is given by*

$$
\begin{aligned}
N \cdot Bias(\hat{\alpha}^{sys}) &= \gamma(B_1^{dif} + B_2^{dif}) + (1-\gamma)(B_1^{lev} + B_2^{lev}) \\
&\quad + \gamma^2(B_3^{dif} + B_4^{dif}) + (1-\gamma)^2(B_3^{lev} + B_4^{lev}) \\
&\quad - \frac{2}{(\phi_d + \phi_l)^2}\Psi_3 + \frac{1}{(\phi_d + \phi_l)^2}\Psi_4 \qquad\qquad (14)
\end{aligned}
$$

*where*

$$
\Psi_3 = \frac{1}{2}\pi_{d,1}\left[Q_{11} + Q_{13}\right] + \frac{1}{2}\pi_{d,2}\left[Q_{21} + Q_{23}\right] + \frac{1}{2}\pi_{d,3}\left[Q_{31} + Q_{33}\right]
$$

$$
\Psi_4 = \frac{1}{2}\pi_{d,1}\left[P_{11} + P_{12}\right] + \frac{1}{2}\pi_{d,2}\left[P_{21} + P_{22}\right] + \frac{1}{2}\pi_{d,3}\left[P_{31} + P_{32}\right]
$$

**Remark 1** Although we do not give explicit formulas, the small sample biases of all the GMM estimators are characterized by $N$, $\alpha$ and $\sigma_\eta^2/\sigma_v^2$. Therefore, we can calculate the theoretical values of the biases, which will be done in the next section.

**Remark 2**   We find that the bias of the system GMM estimator is composed of two elements. The first is the weighted sum of the biases of the first differencing and the level estimators (the first two rows of eq(14)). The second element of the bias results from using the first differencing and the level estimators jointly (the last row in eq(14)). Therefore, if the biases of the first differencing and the level estimators are in opposite directions, they will cancel each other out and the system estimator will have small bias. In the next section we will show that this is indeed the case.

**Remark 3**   Even if the biases of the first differencing and the level estimators are in opposite directions, the bias of the system estimator will not be small in absolute value if $\gamma$ takes near zero or one, or the magnitude of one bias is much larger than the other. Hence, it is also necessary to consider $\gamma$ and the magnitude of the biases.

**Remark 4**   Although we consider the case of $T = 4$ in this paper, from the proof of Theorem 3 we find that what is stated in Remark 2 still holds when $T > 4$ and this case is left for future research.

# 4   Numerical analysis

Since the biases of all estimators are characterized by $N$, $\alpha$ and $\sigma_\eta^2/\sigma_v^2$, as noted in the previous section, we calculate the theoretical values of the biases for the cases $\sigma_\eta^2/\sigma_v^2 = 0.25, 1, 4$ with $\alpha = 0.1, ..., 0.9$. Here, we consider the case $N = 50$. Before we begin to analyze the direction and magnitude of the biases, we confirm how well the second order biases explain the actual biases by comparing theoretical values with simulation values. Figures 2, 3, and 4 describe the theoretical and simulation values of the biases of GMM(DIF), GMM(LEV) and GMM(SYS) estimators.[7] Looking at these figures, we find that although the theoretical and simulation values are close when $\alpha \leq 0.5$, the difference increases in the region of $\alpha > 0.5$. These results are similar to those obtained by Hahn, Hausman and Kuersteiner (2001). Although they state that the bad performance of the second order bias approach when $\alpha$ is close to one is due to weak instruments, the results

---

[7]$v_{it}$, $\eta_i$ and $w_{i1}$ are independently generated by $v_{it} \sim iidN(0,1)$, $\eta_i \sim iidN(0,\sigma_\eta^2)$ and $w_{i1} \sim iidN(0, 1/(1-\alpha^2))$ where $\sigma_\eta^2 = 0.25, 1, 4$.

shown above indicate that the second order bias does not explain the actual bias well even if the instruments are not weak, like GMM(LEV) estimator. Therefore, we can say that the second order bias approach used in this paper works well when $\alpha \leq 0.5$ and in the following we focus on the case with $\alpha \leq 0.5$.

The first concern is the direction of the biases of the GMM(DIF) and GMM(LEV) estimators. Looking at the figures, we find that the GMM(DIF) estimator has a downward bias, while the GMM(LEV) estimator has an upward bias, and both biases cancel each other out in the GMM(SYS) estimator. The second concern is the value of $\gamma$ and the magnitude of the biases. Even though they work in opposite directions and thus at least partly cancel each other out, the bias of the GMM(SYS) estimator will not be small in absolute value if $\gamma$ takes near zero or one, or if the bias of either estimator is much larger than the other in absolute value. Therefore, we need to think about $\gamma$ and the magnitude of the biases. Table 4 shows the values of $\gamma$ for $\alpha = 0.1, ..., 0.9$. We find that $\gamma$ moves between about 0.25 and 0.5 in the region of $\alpha \leq 0.5$. We also find that although the bias of the GMM(DIF) estimator is about -20 − -30% in the region of $\alpha \leq 0.5$, the bias of the GMM(LEV) estimator decreases from about 26% to 5% as $\alpha$ increases. Hence, the difference of the magnitude of the biases of the two estimators in absolute value gets larger as $\alpha$ increases. However, Table 5 shows that $1 - \gamma$, the weight on the GMM(LEV) estimator, increases as $\alpha$ increases, that is, the weight $\gamma$ is adjusting the difference of the magnitude of the biases to be almost the same in absolute value. The column labeled "weighted sum" in Table 4 reports the theoretical values of the first element of the bias of the system estimator. In the region of $\alpha \leq 0.5$, it takes negative stable values around -0.4. The theoretical values of the second element in the bias of the GMM(SYS) estimator are given in the column labeled "correlation" in Table 4. We find that the second element of the bias of the system estimator takes positive values and gets smaller as $\alpha$ grows. Hence, we find that in this case, too, the two elements of the bias of the GMM(SYS) estimator partly cancel each other out, and this is the reason why the small sample bias is almost zero around $\alpha = 0.3$ or 0.4. Thus, we find that there are two reasons why the GMM(SYS) estimator is less biased. The first is that the biases of the GMM(DIF) and the GMM(LEV) estimators are in opposite directions, and the second is that the weight $\gamma$ is adjusting the difference

of the magnitude of both biases.

The above results pertain to the case when $\sigma_\eta^2/\sigma_v^2 = 1$. Now we turn to the case when $\sigma_\eta^2/\sigma_v^2 = 4$. The theoretical and simulation values of the biases are given in Table 2 and plotted in Figures 5 to 8. The approximation of the small sample bias is a little worse than that of $\sigma_\eta^2/\sigma_v^2 = 1$. However, we find that there are sizable biases for any value of $\alpha$ for all estimators. Although Hahn, Hausman and Kuersteiner (2001) state that the GMM(SYS) estimator is heavily biased if the initial conditions fail, the result shown here suggests that even if the initial conditions are satisfied, the GMM(SYS) estimator becomes heavily biased. Therefore, using the GMM(SYS) estimator in this case is problematic and a reduction of the bias is required, which is considered in the next section.

Next, we consider the case $\sigma_\eta^2/\sigma_v^2 = 0.25$ (Table 3 and Figures 9 to 12). Looking at the figures, we see that the GMM(DIF) and the GMM(LEV) estimator have smaller biases than in the case where $\sigma_\eta^2/\sigma_v^2 = 1$. However, the bias of the GMM(SYS) estimator is downward for any $\alpha$, and for many regions of $\alpha$, the GMM(LEV) estimators display the smallest bias among the estimators. This result indicates that the fact that the GMM(SYS) estimator is a weighted sum of the GMM(DIF) and the GMM(LEV) estimator becomes a disadvantage, in contrast with the case of $\sigma_\eta^2/\sigma_v^2 = 1$, where the biases cancelled each other out.

# 5 Bias-corrected GMM estimators

In the previous section, we found that the small sample biases of the GMM estimators crucially depend on the ratio $\sigma_\eta^2/\sigma_v^2$. When $\sigma_\eta^2$ and $\sigma_v^2$ are almost equal, then the GMM(SYS) estimator is the least biased among the GMM estimators. However, if $\sigma_\eta^2$ is large relative to $\sigma_v^2$, the biases of all three GMM estimators are sizable and bias-reduced estimators are required. In this section, we provide bias-corrected GMM estimators and confirm their usefulness via simulation. The method used here to correct for biases is to simply subtract the second order bias from the estimators.[8] Since the second order bias is effective to predict small sample biases in the region of $\alpha \leq 0.5$ as shown in the previous section, we limit our investigation as

---

[8]Also see Newey and Smith (2004).

in the previous section to the case where $\alpha \leq 0.5$.

**Theorem 4.** *The bias corrected GMM estimators for $\hat{\alpha}^{dif}$, $\hat{\alpha}^{lev}$ and $\hat{\alpha}^{sys}$ are given by*

$$\widehat{\widehat{\alpha}}^{dif} = \hat{\alpha}^{dif} - Bias(\widehat{\alpha}^{dif}) \tag{15}$$

$$\widehat{\widehat{\alpha}}^{lev} = \hat{\alpha}^{lev} - Bias(\widehat{\alpha}^{lev}) \tag{16}$$

$$\widehat{\widehat{\alpha}}^{sys} = \hat{\alpha}^{sys} - Bias(\widehat{\alpha}^{sys}) \tag{17}$$

*where*

$$Bias(\widehat{\alpha}^{dif}) = \frac{B_1^{dif} + B_2^{dif} + B_3^{dif} + B_4^{dif}}{N}$$

$$Bias(\widehat{\alpha}^{lev}) = \frac{B_1^{lev} + B_2^{lev} + B_3^{lev} + B_4^{lev}}{N}$$

$$Bias(\widehat{\alpha}^{sys}) = \frac{B_1^{sys} + B_2^{sys} + B_3^{sys} + B_4^{sys}}{N}$$

Since the bias formulas contain expectations and unobservable disturbances, we replace expectations by their sample mean and disturbances by residuals.

## Monte Carlo Study

The data generating process is the same as that of Section 4. Table 6 shows the results of the simulation. Overall, the bias correction works well. However, as $\alpha$ gets larger, the bias-corrected estimators become more dispersed. Especially in the case of $\hat{\alpha}^{dif}$ and $\hat{\alpha}^{lev}$, although the biases are somewhat reduced, the standard deviations are extremely large. Therefore, bias correction for $\hat{\alpha}^{dif}$ and $\hat{\alpha}^{lev}$ may be problematic. Although the standard deviation of the bias corrected GMM(SYS) estimator gets larger as $\alpha$ grows, the degree of dispersion is much smaller than in the case of the bias-corrected GMM(DIF) and GMM(LEV) estimators. Hence, the bias-corrected GMM(SYS) estimator performs best among the bias-corrected GMM estimators considered here.

## 6 Conclusion

In this paper, we considered the small sample bias properties of GMM estimators in dynamic panel data models. We provided theoretical evidence why the system GMM

estimator has smaller bias. We found that the bias of the system GMM estimator is a weighted sum of the biases in opposite directions of the first differencing and the level GMM estimators. In addition, we found that the role of the weight is also important since it adjusts the difference of the magnitudes of the biases. The numerical analysis showed that the fact that $\sigma_\eta^2$ and $\sigma_v^2$ are of almost the same value is an important reason why the system estimator has small bias. In the case when $\sigma_\eta^2/\sigma_v^2 = 4$, the biases of all the GMM estimator are sizable. To reduce these biases, we proposed bias-corrected GMM estimators. Among the bias-corrected GMM estimators, the bias-corrected GMM(SYS) estimator performs well in terms of the magnitude of the bias, although its standard deviation increases. In the case when $\sigma_\eta^2/\sigma_v^2 = 0.25$, we found that the bias of the GMM(LEV) estimator is smaller than that of the GMM(SYS) estimator for a wide region of $\alpha$ and the GMM(SYS) estimator is not best in terms of the magnitude of bias.

# References

[1] Alonso-Borrego, C. and M. Arellano (1999): "Symmetrically normalized instrumental-variable estimation using panel data," *Journal of Business and Economic Statistics*, 17, 36-49.

[2] Alvarez, J. and M. Arellano (2003): "The time series and cross-section asymptotics of dynamic panel data estimators," *Econometrica*, 71, 1121-1159.

[3] Arellano, M. (2003): *Panel Data Econometrics*, Oxford University Press, Oxford.

[4] Arellano, M. and S.R. Bond (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies*, 58, 277-297.

[5] Arellano, M. and O. Bover (1995): "Another look am the instrumental variable estimation of error-component models," *Journal of Econometrics*, 68, 29-45.

[6] Baltagi, B.H (2001): *Econometric Analysis of Panel Data*, 2nd ed., John Wiley and Sons, Chichester.

[7] Bekker, P.A. (1994): "Alternative approximations to the distributions of instrumental variable estimators," *Econometrica*, 62, 657-681.

[8] Blundell, R. and S. Bond (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87, 115-43.

[9] Blundell, R. and S. Bond (2000): "GMM estimation with persistent panel data: An application to production functions," *Econometric Reviews*, 19, 321-340.

[10] Blundell, R., S. Bond and F. Windmeijer (2000): "Estimation in dynamic panel data models: Improving on the performance of the standard GMM estimator", in B.H. Baltagi (ed.), *Nonstationary Panels, Panel Cointegration and Dynamic Panels* Vol.5, pp.53-91, Elsevier Science, Amsterdam.

[11] Bond, S., A. Hoeffler and J. Temple (2001): "GMM estimation of Empirical Growth Models," University of Oxford, Nuffield College, Economics Group, Working Paper No. 2001-W21.

[12] Donald, S. G and W.K. Newey (2001): "Choosing the number of instruments", *Econometrica*, 69, 1161-1191.

[13] Hahn, J., J. Hausman and G. Kuersteiner (2001): "Bias corrected instrumental variables estimation for dynamic panel models with fixed effects," MIT Department of Economics Working Paper No. 01-24.

[14] Hahn, J. and J. Hausman (2002): "Notes on bias in estimators for simultaneous equation models," *Economics Letters*, 75, 237-241.

[15] Hahn, J. and G. Kuersteiner (2002): "Asymptotically unbiased inference for a dynamic panel model with fixed effects when both $n$ and $T$ are large," *Econometrica*, 70, 1639-1657.

[16] Hsiao, C.(2003): *Analysis of Panel Data*, Cambridge University Press, Cambridge and New York.

[17] Kiviet, J.F. (1995): "On bias, inconsistency, and efficiency of various estimators in dynamic panel data models," *Journal of Econometrics*, 68, 53-78.

[18] Koenker, R. and J.A.F. Machado (1999): "GMM inference when the number of moment conditions is large," *Journal of Econometrics*, 93, 327-344.

[19] Kunitomo, N. (1980): "Asymptotic expansions of the distributions of estimators in a linear functional relationship and simultaneous equations," *Journal of the American Statistical Association*, 75, 693-700.

[20] Morimune, K. (1983): "Approximate distributions of $k$-class estimators when the degree of overidentification is large compared with the sample size", *Econometrica*, 51, 821-841.

[21] Nelson, C.R. and R. Startz (1990a): "Some further results on the exact small sample properties of the instrumental variables estimator," *Econometrica*, 58, 967-976.

[22] Nelson, C.R and R. Startz (1990b): "The distribution of the instrumental variable estimator and its $t$ ratio when the instrument is a poor one," *Journal of Business*, 63, S125-140.

[23] Newey,W.K and R.J. Smith (2004): "Higher order properties of GMM and Generalized Empirical Likelihood estimators", *Econometrica*, 72, 219-255.

[24] Okui, R. (2004): "The optimal choice of moments in dynamic panel data models", unpublished manuscript.

[25] Staiger, D. and J.H. Stock (1997): "Instrumental variables regression with weak instruments," *Econometrica*, 65, 557-586.

[26] Stock, J.H, J.H. Wright and M. Yogo (2002): "A survey of weak instruments and weak identification in Generalized Method of Moments," *Journal of Business and Economics Statistics*, 20, 518-529.

[27] Tauchen, G. (1986): "Statistical properties of generalized method of moments estimators of structural parameters obtained from financial market data," *Journal of Business and Economic Statistics*, 4, 397-416.

[28] Ziliak, J.P. (1997): "Efficient estimation with panel data when instruments are predetermined: An empirical comparison of moment-condition estimators," *Journal of Business and Economic Statistics*, 15, 419-431.

# A    Theoretical proofs

To prove Theorems 1, 2 and 3, we consider the general one-step GMM estimator based on the moment condition $E[g(z_i, \theta_0) = E[g_i(\theta_0)] = 0$. An inefficient one-step GMM estimator is defined as

$$\widehat{\theta} = \arg\min_{\theta} \widehat{g}(\theta)'\widehat{W}^{-1}\widehat{g}(\theta) \tag{18}$$

where $\widehat{g}(\theta) = N^{-1}\sum_{i=1}^{N} g_i(\theta)$, $\widehat{W} = N^{-1}\sum_{i=1}^{N} W_i \rightarrow^p W$ and $W_i = W(z_i)$ are symmetric and positive definite matrices which do not depend on parameter $\theta$.

Generally, an estimator of $\theta$, $\widehat{\theta}$, based on a sample of size $N$ allows for an expansion of the form:

$$\sqrt{N}(\widehat{\theta} - \theta) = \theta^{(1)} + \frac{1}{\sqrt{N}}\theta^{(2)} + O_p\left(\frac{1}{N}\right) \tag{19}$$

where both $\theta^{(1)}$ and $\theta^{(2)}$ are $O_p(1)$. Typically, $\theta^{(1)}$ has a zero mean and converges in distribution to a normal distribution. By taking an expectation and ignoring the $O_p(1/N)$ term, we obtain the approximate mean of $\sqrt{N}(\widehat{\theta} - \theta)$, $\frac{1}{\sqrt{N}}E(\theta^{(2)})$. Therefore, we can regard $\frac{1}{N}E(\theta^{(2)})$ as the second order bias of $\widehat{\theta}$. In the following, we derive $N^{-1}E[\theta^{(2)}]$ for a one-step GMM estimator.

**Lemma 1.** *The second order bias for the one-step GMM estimator is given by*

$$
\begin{aligned}
N \cdot Bias(\widehat{\theta}) &= \frac{E(G'W^{-1}W_i W^{-1}g_i)}{G'W^{-1}G} - \frac{trace(W^{-1}E(g_i G_i'))}{G'W^{-1}G} \\
&\quad + 2\frac{G'W^{-1}E(g_i G_i')W^{-1}G}{(G'W^{-1}G)^2} - \frac{E(G'W^{-1}g_i G'W^{-1}W_i W^{-1}G)}{(G'W^{-1}G)^2} \\
&\quad + \frac{G'W^{-1}\Omega W^{-1}G_\theta}{(G'W^{-1}G)^2} - \frac{3}{2}\frac{(G'W^{-1}\Omega W^{-1}G)(G'W^{-1}G_\theta)}{(G'W^{-1}G)^3} \\
&= B_1 + B_2 + B_3 + B_4 + B_5 + B_6
\end{aligned}
$$

*where*

$$
G_i = \frac{\partial g_i(\theta_0)}{\partial\theta} \quad G = E(G_i) \quad G_\theta = E\left(\frac{\partial^2 g_i(\theta_0)}{\partial\theta^2}\right) \quad \Omega = E[g_i(\theta_0)g_i(\theta_0)'].
$$

**Proof.** The one-step GMM estimator is a special case of Theorem 1 in Hahn, Hausman and Kuersteiner (2001). All that is required is replacing the weighting matrix $\psi_i(c)\psi_i(c)'$ with $W_i$. $\square$

Furthermore, because the moment conditions considered in this paper are linear in the parameter, $G_\theta = 0$ holds, and hence the last two terms, $B_5$ and $B_6$, equal

zero. In the following, we derive $B_1, B_2, B_3, B_4$ each for GMM(DIF), GMM(LEV) and GMM(SYS), and to distinguish these estimators, we use the superscripts "dif", "lev" and "sys" for $B_1, ..., B_4$.

## Proof of Theorem 1

Since $G_i = -Z'_{d,i}\Delta y_{i,-1}$, $G = E[G_i]$, $W = E[Z'_{d,i}Z_{d,i}]$, $W_i = Z'_{d,i}Z_{d,i}$ and $g_i = Z'_{d,i}\Delta u_i$, $B_1^{dif}$ can be expressed as

$$
\begin{aligned}
B_1^{dif} &= -\frac{E\left[E(\Delta y'_{i,-1}Z_{d,i})E(Z'_{d,i}Z_{d,i})^{-1}Z'_{di}Z_{di}E(Z'_{di}Z_{di})^{-1}Z'_{di}\Delta u_i\right]}{E(\Delta y'_{i,-1}Z_{di})E(Z'_{di}Z_{di})^{-1}E(Z'_{di}\Delta y_{i,-1})} \\
&= -\frac{1}{\phi_d}\left[\pi_{d1}(C+D)^{-1}E(y_{i1}^3\Delta u_{i3})\right. \\
&\quad +\frac{\pi_{d2}}{F}\left\{(C+D)E(y_{i,1}^3\Delta u_{i4} + y_{i1}y_{i2}^2\Delta u_{i4})\right. \\
&\quad\quad \left.-(C+\alpha D)E(y_{i1}^2 y_{i2}\Delta u_{i4} + y_{i1}^2 y_{i2}\Delta u_{i4})\right\} \\
&\quad +\frac{\pi_{d3}}{F}\left\{(C+D)E(y_{i,1}^2 y_{i2}\Delta u_{i4} + y_{i2}^3\Delta u_{i4})\right. \\
&\quad\quad \left.\left.-(C+\alpha D)E(y_{i1}y_{i2}^2\Delta u_{i4} + y_{i1}y_{i2}^2\Delta u_{i4})\right\}\right].
\end{aligned}
$$

Because all the expectations are equal to zero,

$$B_1^{dif} = 0. \tag{20}$$

Next, $B_2^{dif}$ is expressed as

$$
\begin{aligned}
B_2^{dif} &= \frac{trace\left[E(Z'_{di}Z_{di})^{-1}E(Z'_{di}\Delta u_i\Delta y'_{i,-1}Z_{di})\right]}{E(\Delta y'_{i,-1}Z_{di})E(Z'_{di}Z_{di})^{-1}E(Z'_{di}\Delta y_{i,-1})} \\
&= \frac{1}{\phi_d}\left[(C+D)^{-1}E(y_{i1}^2\Delta y_{i2}\Delta u_{i3})\right. \\
&\quad +\frac{C+D}{F}E(y_{i1}^2\Delta y_{i3}\Delta u_{i4} + y_{i2}^2\Delta y_{i3}\Delta u_{i4}) \\
&\quad \left.-2\left(\frac{C+\alpha D}{F}\right)E(y_{i1}y_{i2}\Delta y_{i3}\Delta u_{i4})\right].
\end{aligned}
$$

Since

$$E(y_{i1}^2\Delta y_{i2}\Delta u_{i3}) = E(y_{i1}^2\Delta y_{i3}\Delta u_{i4}) = E(y_{i2}^2\Delta y_{i3}\Delta u_{i4}) = -\sigma_v^2(C+D)$$

$$E(y_{i1}y_{i2}\Delta y_{i3}\Delta u_{i4}) = -\sigma_v^2(C+\alpha D),$$

$B_2^{dif}$ reduces to

$$B_2^{dif} = \frac{-\sigma_v^2}{\phi_d}\left[1 + \frac{2(C+D)^2}{F} - \frac{2(C+\alpha D)^2}{F}\right]. \tag{21}$$

16

Note that $B_3^{dif}$ has the following form:

$$B_3^{dif} = -2\frac{E(\Delta y_{i,-1}'Z_{di})E(Z_{di}'Z_{di})^{-1}E(Z_{di}'\Delta u_i\Delta y_{i,-1}'Z_{di})E(Z_{di}'Z_{di})^{-1}E(Z_{di}'\Delta y_{i,-1})}{\left[E(\Delta y_{i,-1}'Z_{di})E(Z_{di}'Z_{di})^{-1}E(Z_{di}'\Delta y_{i,-1})\right]^2}$$

$$= \frac{-2}{\phi_d^2}\left[\pi_{d1}^2E(y_{i1}^2\Delta y_{i2}\Delta u_{i3}) + \pi_{d1}\pi_{d2}E(y_{i1}^2\Delta y_{i3}\Delta u_{i3}) + \pi_{d1}\pi_{d3}E(y_{i1}y_{i2}\Delta y_{i3}\Delta u_{i3})\right.$$

$$+\pi_{d1}\pi_{d2}E(y_{i1}^2\Delta y_{i2}\Delta u_{i4}) + \pi_{d2}^2E(y_{i1}^2\Delta y_{i3}\Delta u_{i4}) + 2\pi_{d2}\pi_{d3}E(y_{i1}y_{i2}\Delta y_{i3}\Delta u_{i4})$$

$$\left.\pi_{d1}\pi_{d3}E(y_{i1}y_{i2}\Delta y_{i2}\Delta u_{i4}) + \pi_{d3}^2E(y_{i2}^2\Delta y_{i3}\Delta u_{i4})\right].$$

Since

$$E(y_{i1}^2\Delta y_{i2}\Delta u_{i3}) = E(y_{i1}^2\Delta y_{i3}\Delta u_{i4}) = E(y_{i2}^2\Delta y_{i3}\Delta u_{i4}) = -\sigma_v^2(C + D)$$

$$E(y_{i1}^2\Delta y_{i3}\Delta u_{i3}) = \sigma_v^2(2 - \alpha)(C + D)$$

$$E(y_{i1}y_{i2}\Delta y_{i3}\Delta u_{i3}) = \sigma_v^2\left[(2 - \alpha)C - \alpha(2\alpha - 3)D\right]$$

$$E(y_{i1}y_{i2}\Delta y_{i3}\Delta u_{i4}) = -\sigma_v^2(C + \alpha D)$$

and $E(y_{i1}^2\Delta y_{i2}\Delta u_{i4}) = E(y_{i1}y_{i2}\Delta y_{i2}\Delta u_{i4}) = 0$, we get

$$B_3^{dif} = \frac{2\sigma_v^2}{\phi_d^2}\left[\left\{\pi_{d1}^2 + \pi_{d2}^2 + \pi_{d3}^2 + (\alpha - 2)\pi_{d1}\pi_{d2}\right\}(C + D)\right.$$

$$\left.-\pi_{d1}\pi_{d3}\left\{(2 - \alpha)C - \alpha(2\alpha - 3)D\right\} + 2\pi_{d2}\pi_{d3}(C + \alpha D)\right]. \quad (22)$$

Finally $B_4^{dif}$ can be expressed as

$$B_4^{dif} = \frac{E\left[E(\Delta y_{i,-1}'Z_{di})E(Z_{di}'Z_{di})^{-1}Z_{di}'\Delta u_i\right.}{\left[E(\Delta y_{i,-1}'Z_{di})E(Z_{di}'Z_{di})^{-1}E(Z_{di}'\Delta y_{i,-1})\right]^2}$$

$$\frac{\times E(\Delta y_{i,-1}'Z_{di})E(Z_{di}'Z_{di})^{-1}Z_i'Z_iE(Z_{di}'Z_{di})^{-1}E(Z_{di}'\Delta y_{i,-1})\Big]}{\left[E(\Delta y_{i,-1}'Z_{di})E(Z_{di}'Z_{di})^{-1}E(Z_{di}'\Delta y_{i,-1})\right]^2}$$

$$= \frac{1}{\phi_d^2}\left[\pi_{d1}(\pi_{d1}^2 + \pi_{d2}^2)E(y_{i1}^3\Delta u_{i3}) + \pi_{d2}(\pi_{d1}^2 + \pi_{d2}^2)E(y_{i1}^3\Delta u_{i4}) + \pi_{d3}^3E(y_{i2}^3\Delta u_{i4})\right.$$

$$\pi_{d1}\pi_{d3}^2E(y_{i1}y_{i2}^2\Delta u_{i3}) + 2\pi_{d1}\pi_{d2}\pi_{d3}E(y_{i1}^2y_{i2}\Delta u_{i3})$$

$$\left.3\pi_{d2}\pi_{d3}^2E(y_{i1}y_{i2}^2\Delta u_{i4}) + (\pi_{d1}^2 + \pi_{d3} + 3\pi_{d2}^2\pi_{d3})E(y_{i1}^2y_{i2}\Delta u_{i4})\right].$$

Since

$$E(y_{i1}^2y_{i2}\Delta u_{i3}) = -\sigma_v^2(C + D)$$

17

$$E(y_{i1}y_{i2}^2\Delta u_{i3}) = -2\sigma_v^2(C + \alpha D)$$

and the other expectations are zero, we have

$$B_4^{dif} = -\frac{-2\sigma_v^2}{\phi_d^2}\left[\pi_{d1}\pi_{d3}^2(C+\alpha D) + \pi_{d1}\pi_{d2}\pi_{d3}(C+D)\right]. \tag{23}$$

By adding the terms from eq.(20) to eq.(23), the proof is completed. $\square$

## Proof of Theorem 2

The flow of the proof is the same as Theorem 1: the four elements of the bias are calcurated in sequence. First $B_1^{lev}$ has an expression as follows:

$$
\begin{aligned}
B_1^{lev} &= -\frac{E\left[E(y'_{i,-1}Z_{l,i})E(Z'_{l,i}Z_{l,i})^{-1}Z'_{l,i}Z_{l,i}E(Z'_{l,i}Z_{l,i})^{-1}Z'_{l,i}u_i\right]}{E(y'_{i,-1}Z_{l,i})E(Z'_{l,i}Z_{l,i})^{-1}E(Z'_{l,i}y_{i,-1})} \\
&= -\frac{1}{\phi_l}\left[\frac{1}{2}\left(\frac{1+\alpha}{2\sigma_v^2}\right)E(\Delta y_{i2}^3 u_{i3}) + \frac{1}{2}\left(\frac{1+\alpha}{2\sigma_v^2}\right)E(\Delta y_{i3}^3 u_{i4})\right] \\
&= 0
\end{aligned} \tag{24}
$$

The last equality comes from the fact that all the expectations are equal to zero. Next we consider $B_2^{lev}$.

$$
\begin{aligned}
B_2^{lev} &= \frac{trace\left[E(Z'_{l,i}Z_{l,i})^{-1}E(Z'_{l,i}u_iy'_{i,-1}Z_{l,i})\right]}{E(y'_{i,-1}Z_{l,i})E(Z'_{l,i}Z_{l,i})^{-1}E(Z'_{l,i}y_{i,-1})} \\
&= \frac{1}{\phi_l}\left[\left(\frac{1+\alpha}{2\sigma_v^2}\right)E(y_{i2}\Delta y_{i2}^2 u_{i3}) + \left(\frac{1+\alpha}{2\sigma_v^2}\right)E(y_{i3}\Delta y_{i3}^2 u_{i4})\right]
\end{aligned}
$$

Since $E(y_{i2}\Delta y_{i2}^2 u_{i3}) = E(y_{i3}\Delta y_{i3}^2 u_{i4}) = \frac{2}{1-\alpha^2}\sigma_\eta^2\sigma_v^2$, we get

$$B_2^{lev} = \frac{2\sigma_\eta^2}{\phi_l(1-\alpha)}. \tag{25}$$

$B_3^{lev}$ can be expressed as

$$B_3^{lev} = -2\frac{E(y'_{i,-1}Z_{l,i})E(Z'_{l,i}Z_{l,i})^{-1}E(Z'_{l,i}u_iy'_{i,-1}Z_{l,i})E(Z'_{l,i}Z_{l,i})^{-1}E(Z'_{l,i}y_{i,-1})}{\left[E(y'_{i,-1}Z_{l,i})E(Z'_{l,i}Z_{l,i})^{-1}E(Z'_{l,i}y_{i,-1})\right]^2}.$$

Because

$$E(y_{i2}\Delta y_{i2}^2 u_{i3}) = E(y_{i3}\Delta y_{i3}^2 u_{i4}) = \frac{2}{1-\alpha^2}\sigma_\eta^2\sigma_v^2$$

$$E(y_{i3}\Delta y_{i2}\Delta y_{i3}u_{i3}) = \frac{-1}{1+\alpha}\sigma_\eta^2\sigma_v^2 + \frac{2\alpha-1}{1+\alpha}\sigma_v^4$$

18

$$E(y_{i3}\Delta y_{i2}\Delta y_{i3}u_{i4}) = E(y_{i2}\Delta y_{i2}\Delta y_{i3}u_{i4}) = \frac{-1}{1+\alpha}\sigma_\eta^2\sigma_v^2,$$

we have

$$
\begin{aligned}
B_3^{lev} &= \frac{-2}{\phi_l^2}
\begin{bmatrix} \pi_{l,1} & \pi_{l,2} \end{bmatrix}
\begin{bmatrix}
E(y_{i2}\Delta y_{i2}^2 u_{i3}) & E(y_{i3}\Delta y_{i2}\Delta y_{i3}u_{i3}) \\
E(y_{i2}\Delta y_{i2}\Delta y_{i3}u_{i4}) & E(y_{i3}\Delta y_{i3}^2 u_{i4})
\end{bmatrix}
\begin{bmatrix} \pi_{l,1} \\ \pi_{l,2} \end{bmatrix} \\
&= \frac{-2}{\phi_l^2}
\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}
\begin{bmatrix}
\frac{2}{1-\alpha^2}\sigma_v^2\sigma_\eta^2 & \frac{-1}{1+\alpha}\sigma_v^2\sigma_\eta^2 + \frac{2\alpha-1}{1+\alpha}\sigma_v^4 \\
\frac{-1}{1+\alpha}\sigma_v^2\sigma_\eta^2 & \frac{2}{1-\alpha^2}\sigma_v^2\sigma_\eta^2
\end{bmatrix}
\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \\
&= \frac{-\sigma_v^2}{\phi_l^2}\left[ \frac{1}{1-\alpha}\sigma_\eta^2 + \frac{2\alpha-1}{2(1+\alpha)}\sigma_v^2 \right]
\end{aligned}
\tag{26}
$$

Finally we consider $B_4^{lev}$.

$$
\begin{aligned}
B_4^{lev} &= \frac{E\left[ E(y_{i,-1}'Z_{l,i})E(Z_{l,i}'Z_{l,i})^{-1}Z_{l,i}'u_i \right.}{\left[ E(y_{i,-1}'Z_{l,i})E(Z_{l,i}'Z_{l,i})^{-1}E(Z_{l,i}'y_{i,-1}) \right]^2} \\
&\qquad \frac{\left. \times\, E(y_{i,-1}'Z_{l,i})E(Z_{l,i}'Z_{l,i})^{-1}Z_{l,i}'Z_{l,i}E(Z_{l,i}'Z_{l,i})^{-1}E(Z_{l,i}'y_{i,-1}) \right]}{\left[ E(y_{i,-1}'Z_{l,i})E(Z_{l,i}'Z_{l,i})^{-1}E(Z_{l,i}'y_{i,-1}) \right]^2}
\end{aligned}
$$

Because $E(\Delta y_{i2}^3 u_{i3}) = E(\Delta y_{i2}^2\Delta y_{i3}u_{i4}) = E(\Delta y_{i3}^3 u_{i4}) = 0$ and

$$E(\Delta y_{i2}\Delta y_{i3}^2 u_{i3}) = 2\left(\frac{\alpha-1}{\alpha+1}\right)\sigma_v^4$$

hold, we get

$$
\begin{aligned}
B_4^{lev} &= \frac{1}{8\phi_l^2}\left[ E(\Delta y_{i2}^3 u_{i3}) + E(\Delta y_{i2}\Delta y_{i3}^2 u_{i3}) + E(\Delta y_{i2}^2\Delta y_{i3}u_{i4}) + E(\Delta y_{i3}^3 u_{i4}) \right] \\
&= \frac{\alpha-1}{4\phi_l^2(\alpha+1)}\sigma_v^4.
\end{aligned}
\tag{27}
$$

By adding terms from eq.(24) to eq.(27), the result is obtained. $\square$

## Proof of Theorem 3

The second order bias of $\hat{\alpha}^{sys}$ can be written as

$$Bias(\hat{\alpha}^{sys}) = \frac{B_1^{sys} + B_2^{sys} + B_3^{sys} + B_4^{sys}}{N}$$

where

$$B_1^{sys} = -\frac{E\left[ E(q_{i,-1}'Z_{s,i})E(Z_{s,i}'Z_{s,i})^{-1}Z_{s,i}'Z_{s,i}E(Z_{s,i}'Z_{s,i})^{-1}Z_{s,i}'p_i \right]}{E(q_{i,-1}'Z_{s,i})E(Z_{s,i}'Z_{s,i})^{-1}E(Z_{s,i}'q_{i,-1})}$$

19

$$B_2^{sys} = \frac{trace\left[E(Z_{s,i}'Z_{s,i})^{-1}E(Z_{s,i}'p_iq_{i,-1}'Z_{s,i})\right]}{E(q_{i,-1}'Z_{s,i})E(Z_{s,i}'Z_{s,i})^{-1}E(Z_{s,i}'q_{i,-1})}$$

$$B_3^{sys} = -2\frac{E(q_{i,-1}'Z_{s,i})E(Z_{s,i}'Z_{s,i})^{-1}E(Z_{s,i}'p_iq_{i,-1}'Z_{s,i})E(Z_{s,i}'Z_{s,i})^{-1}E(Z_{s,i}'q_{i,-1})}{\left[E(q_{i,-1}'Z_{s,i})E(Z_{s,i}'Z_{s,i})^{-1}E(Z_{s,i}'q_{i,-1})\right]^2}$$

$$B_4^{sys} = \frac{E\left[E(q_{i,-1}'Z_{s,i})E(Z_{s,i}'Z_{s,i})^{-1}Z_{s,i}'p_i\right.}{\left[E(q_{i,-1}'Z_{s,i})E(Z_{s,i}'Z_{s,i})^{-1}E(Z_{s,i}'q_{i,-1})\right]^2}$$
$$\frac{\times\left. E(q_{i,-1}'Z_{s,i})E(Z_{s,i}'Z_{s,i})^{-1}Z_{s,i}'Z_{s,i}E(Z_{s,i}'Z_{s,i})^{-1}E(Z_{s,i}'q_{i,-1})\right]}{\left[E(q_{i,-1}'Z_{s,i})E(Z_{s,i}'Z_{s,i})^{-1}E(Z_{s,i}'q_{i,-1})\right]^2}.$$

After some algebra, we have

$$B_1^{sys} + B_2^{sys} = \gamma(B_1^{dif} + B_2^{dif}) + (1-\gamma)(B_1^{lev} + B_2^{lev})$$

$$B_3^{sys} + B_4^{sys} = \gamma^2(B_3^{dif} + B_4^{dif}) + (1-\gamma)^2(B_3^{lev} + B_4^{lev})$$
$$-\frac{2}{(\phi_d^2 + \phi_l^2)^2}\pi_d'\left[E(Z_{di}'\Delta u_i y_{i,-1}'Z_{l,i}) + E(Z_{di}'\Delta y_{i,-1}u_i'Z_{l,i})\right]\pi_l$$
$$+\frac{1}{(\phi_d^2 + \phi_l^2)^2}\pi_d'\left[E(Z_{di}'\Delta u_i\pi_l'Z_{l,i}'Z_{l,i}) + E(Z_{di}'Z_{di}\pi_d u_i'Z_{l,i})\right]\pi_l.$$

Hence, we only have to calculate the last two rows in $B_3^{sys} + B_4^{sys}$, that is, $\Psi_3$ and $\Psi_4$. First, $\Psi_3$ can be expressed as

$$\Psi_3 = \pi_d'\left[E(Z_{di}'\Delta u_i y_{i,-1}'Z_{l,i}) + E(Z_{di}'\Delta y_{i,-1}u_i'Z_{l,i})\right]\pi_{l2}$$

$$= \pi_d'\left[\begin{array}{l} E(y_{i1}y_{i2}\Delta y_{i2}\Delta u_{i3}) + E(y_{i1}\Delta y_{i2}^2 u_{i3}) \\ E(y_{i1}y_{i2}\Delta y_{i2}\Delta u_{i4}) + E(y_{i1}\Delta y_{i2}\Delta y_{i3}u_{i3}) \\ E(y_{i2}^2\Delta y_{i2}\Delta u_{i4}) + E(y_{i2}\Delta y_{i2}\Delta y_{i3}u_{i3}) \end{array}\right.$$

$$\left.\begin{array}{r} E(y_{i1}y_{i3}\Delta y_{i3}\Delta u_{i3}) + E(y_{i1}\Delta y_{i2}\Delta y_{i3}u_{i4}) \\ E(y_{i1}y_{i3}\Delta y_{i3}\Delta u_{i4}) + E(y_{i1}\Delta y_{i3}^2 u_{i4}) \\ E(y_{i2}y_{i3}\Delta y_{i3}\Delta u_{i4}) + E(y_{i2}\Delta y_{i3}^2 u_{i4}) \end{array}\right]\pi_l.$$

Note that

$$E(y_{i1}y_{i2}\Delta y_{i2}\Delta u_{i3}) = E(y_{i2}y_{i3}\Delta y_{i3}\Delta u_{i4}) = -\left(\frac{1}{1-\alpha}\right)^2\sigma_\eta^2\sigma_v^2 - \frac{2\alpha-1}{1-\alpha^2}\sigma_v^4$$

$$E(y_{i1}\Delta y_{i2}^2 u_{i3}) = E(y_{i1}\Delta y_{i3}^2 u_{i4}) = E(y_{i2}\Delta y_{i3}^2 u_{i4}) = \frac{2}{1-\alpha^2}\sigma_\eta^2\sigma_v^2$$

$$E(y_{i1}\Delta y_{i2}\Delta y_{i3}u_{i3}) = \frac{-1}{1+\alpha}\sigma_v^2\sigma_\eta^2 - \frac{1}{1+\alpha}\sigma_v^4$$

20

$$E(y_{i2}\Delta y_{i2}\Delta y_{i3}u_{i3}) = \frac{-1}{1+\alpha}\sigma_v^2\sigma_\eta^2 + \frac{1}{1+\alpha}\sigma_v^4$$

$$E(y_{i1}y_{i3}\Delta y_{i3}\Delta u_{i3}) = \frac{2-\alpha}{(1-\alpha)^2}\sigma_\eta^2\sigma_v^2 - \frac{\alpha(2\alpha^2-4\alpha+1)}{1-\alpha^2}\sigma_v^4$$

$$E(y_{i1}y_{i3}\Delta y_{i3}\Delta u_{i4}) = -\left(\frac{1}{1-\alpha}\right)^2\sigma_\eta^2\sigma_v^2 - \frac{\alpha(2\alpha-1)}{1-\alpha^2}\sigma_v^4$$

hold. Using these results, we get

$$\Phi_3 = \pi_d'\begin{bmatrix} \frac{-3\alpha+1}{(1-\alpha)^2(1+\alpha)}\sigma_\eta^2\sigma_v^2 - \frac{2\alpha-1}{1-\alpha^2}\sigma_v^4 & \frac{-2\alpha^2+3\alpha+1}{(1-\alpha)^2(1+\alpha)}\sigma_v^2\sigma_\eta^2 - \frac{\alpha(2\alpha^2-4\alpha+1)}{1-\alpha^2}\sigma_v^4 \\ \frac{-1}{1+\alpha}\sigma_\eta^2\sigma_v^2 - \frac{1}{1+\alpha}\sigma_v^4 & \frac{-3\alpha+1}{(1-\alpha)^2(1+\alpha)}\sigma_v^2\sigma_\eta^2 - \frac{\alpha(2\alpha-1)}{1-\alpha^2}\sigma_v^4 \\ \frac{-1}{1+\alpha}\sigma_\eta^2\sigma_v^2 + \frac{1}{1+\alpha}\sigma_v^4 & \frac{-3\alpha+1}{(1-\alpha)^2(1+\alpha)}\sigma_v^2\sigma_\eta^2 - \frac{2\alpha-1}{1-\alpha^2}\sigma_v^4 \end{bmatrix}\pi_l$$

$$= \begin{bmatrix} \pi_{d,1} & \pi_{d,2} & \pi_{d,3} \end{bmatrix}\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \\ Q_{31} & Q_{32} \end{bmatrix}\begin{bmatrix} \pi_{l,1} \\ \pi_{l,2} \end{bmatrix}. \qquad \square$$

Next, $\Psi_4$ has the following form:

$$\Psi_4 = \pi_d'\left[E(Z_{di}'\Delta u_i\pi_l'Z_{l,i}'Z_{l,i}) + E(Z_{di}'Z_{di}\pi_d u_i'Z_{l,i})\right]\pi_l$$

$$= \pi_d'\left[\begin{pmatrix} \pi_{l,1}E(y_{i1}\Delta y_{i2}^2\Delta u_{i3}) & \pi_{l,2}E(y_{i1}\Delta y_{i3}^2\Delta u_{i3}) \\ \pi_{l,1}E(y_{i1}\Delta y_{i2}^2\Delta u_{i4}) & \pi_{l,2}E(y_{i1}\Delta y_{i3}^2\Delta u_{i4}) \\ \pi_{l,1}E(y_{i2}\Delta y_{i2}^2\Delta u_{i4}) & \pi_{l,2}E(y_{i2}\Delta y_{i3}^2\Delta u_{i4}) \end{pmatrix}\right.$$

$$+ \begin{pmatrix} \pi_{d,1}E(y_{i1}^2\Delta y_{i2}u_{i3}) \\ \pi_{d,2}E(y_{i1}^2\Delta y_{i2}u_{i3}) + \pi_{d,3}E(y_{i1}y_{i2}\Delta y_{i2}u_{i3}) \\ \pi_{d,2}E(y_{i1}y_{i2}\Delta y_{i2}u_{i3}) + \pi_{d,3}E(y_{i2}^2\Delta y_{i2}u_{i3}) \end{pmatrix}$$

$$\left.\begin{pmatrix} \pi_{d,1}E(y_{i1}^2\Delta y_{i2}u_{i3}) \\ \pi_{d,2}E(y_{i1}^2\Delta y_{i2}u_{i3}) + \pi_{d3}E(y_{i1}y_{i2}\Delta y_{i2}u_{i3}) \\ \pi_{d,3}E(y_{i1}y_{i2}\Delta y_{i2}u_{i3}) + \pi_{d3}E(y_{i2}^2\Delta y_{i2}u_{i3}) \end{pmatrix}\right]\pi_l.$$

Because we have the following results:

$$E(y_{i1}\Delta y_{i2}^2\Delta u_{i3}) = E(y_{i2}\Delta y_{i3}^2\Delta u_{i4}) = \frac{2}{1+\alpha}\sigma_v^4$$

$$E(y_{i1}\Delta y_{i3}^2\Delta u_{i3}) = \frac{2\alpha(\alpha-2)}{1+\alpha}\sigma_v^4$$

$$E(y_{i1}\Delta y_{i3}^2\Delta u_{i4}) = \frac{2\alpha}{1+\alpha}\sigma_v^4$$

$$E(y_{i1}^2\Delta y_{i2}u_{i3}) = E(y_{i1}^2\Delta y_{i2}u_{i4}) = E(y_{i2}^2\Delta y_{i3}u_{i4}) = \frac{-2}{1-\alpha^2}\sigma_v^2\sigma_\eta^2$$

$$E(y_{i2}^2\Delta y_{i2}u_{i3}) = \frac{2\alpha}{1-\alpha^2}\sigma_\eta^2\sigma_v^2$$

$$E(y_{i1}^2 \Delta y_{i3} u_{i4}) = \frac{-2\alpha}{1-\alpha^2}\sigma_v^2\sigma_\eta^2$$

$$E(y_{i1}y_{i2}\Delta y_{i3} u_{i4}) = \frac{-1}{1-\alpha}\sigma_\eta^2\sigma_v^2,$$

we obtain

$$\Psi_4 = \pi_d' \begin{bmatrix} \pi_{l,1}\frac{2}{\alpha+1}\sigma_v^4 - \pi_{d,1}\frac{2}{1-\alpha^2}\sigma_v^2\sigma_\eta^2 & \pi_{l,2}\frac{2\alpha(\alpha-2)}{1+\alpha}\sigma_v^4 - \pi_{d,1}\frac{2\alpha}{1-\alpha^2}\sigma_v^2\sigma_\eta^2 \\ -\pi_{d,2}\frac{2}{1-\alpha^2}\sigma_v^2\sigma_\eta^2 & \pi_{l,2}\frac{2\alpha}{1+\alpha}\sigma_v^4 - \pi_{d,2}\frac{2\alpha}{1-\alpha^2}\sigma_v^2\sigma_\eta^2 - \pi_{d,3}\frac{1}{1-\alpha}\sigma_v^2\sigma_\eta^2 \\ \pi_{d,3}\frac{2\alpha}{1-\alpha^2}\sigma_\eta^2\sigma_v^2 & \pi_{l,2}\frac{2}{1+\alpha}\sigma_v^4 - \pi_{d,2}\frac{1}{1-\alpha}\sigma_v^2\sigma_\eta^2 - \pi_{d,3}\frac{2}{1-\alpha^2}\sigma_v^2\sigma_\eta^2 \end{bmatrix} \pi_l$$

$$= \begin{bmatrix} \pi_{d,1} & \pi_{d,2} & \pi_{d,3} \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \\ P_{31} & P_{32} \end{bmatrix} \begin{bmatrix} \pi_{l,1} \\ \pi_{l,2} \end{bmatrix}.$$
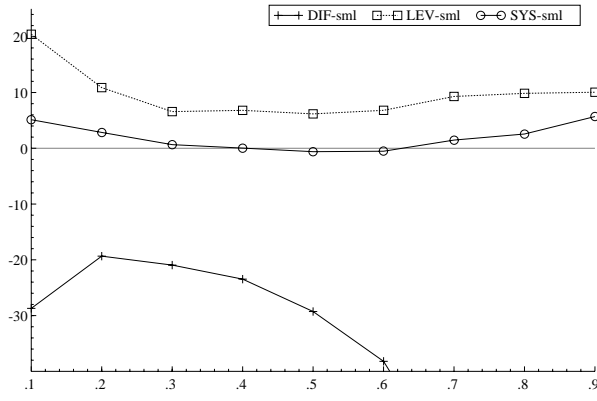
22

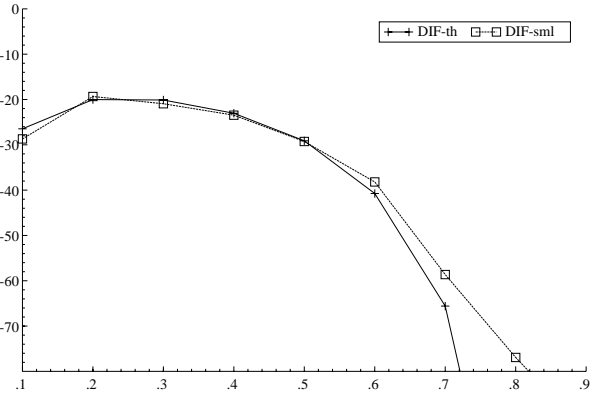Figure 1: Bias of $\widehat{\alpha}$ $(\sigma_\eta^2/\sigma_v^2 = 1)$

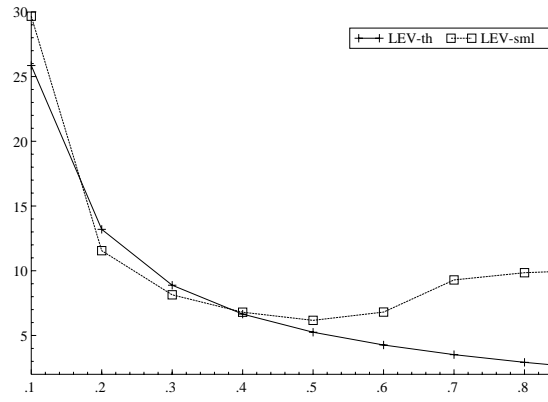Figure 2: Bias of $\widehat{\alpha}^{dif}$ $(\sigma_\eta^2/\sigma_v^2 = 1)$

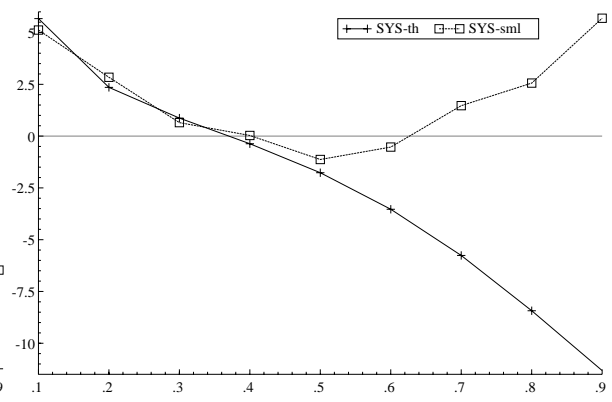Figure 3: Bias of $\widehat{\alpha}^{lev}$ $(\sigma_\eta^2/\sigma_v^2 = 1)$

Figure 4: Bias of $\widehat{\alpha}^{sys}$ $(\sigma_\eta^2/\sigma_v^2 = 1)$

Figure 5: Bias of $\widehat{\alpha}$ $(\sigma_\eta^2/\sigma_v^2 = 4)$

Figure 6: Bias of $\widehat{\alpha}^{dif}$ $(\sigma_\eta^2/\sigma_v^2 = 4)$

23

Figure 7: Bias of $\widehat{\alpha}^{lev}$ $(\sigma_\eta^2/\sigma_v^2 = 4)$



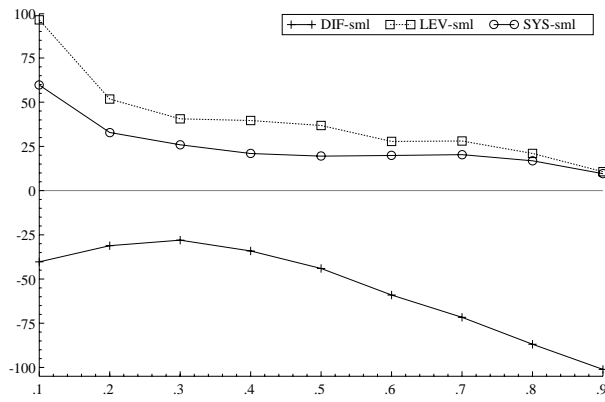Figure 8: Bias of $\widehat{\alpha}^{sys}$ $(\sigma_\eta^2/\sigma_v^2 = 4)$



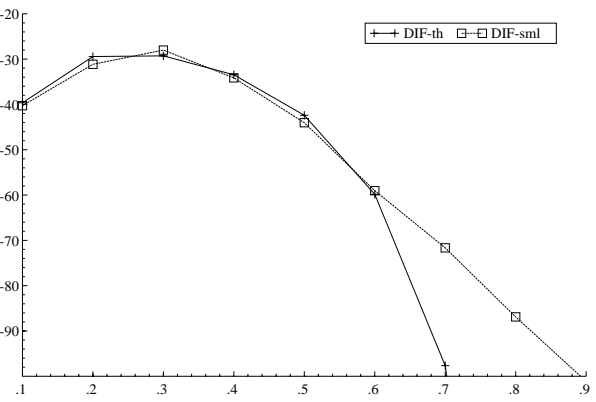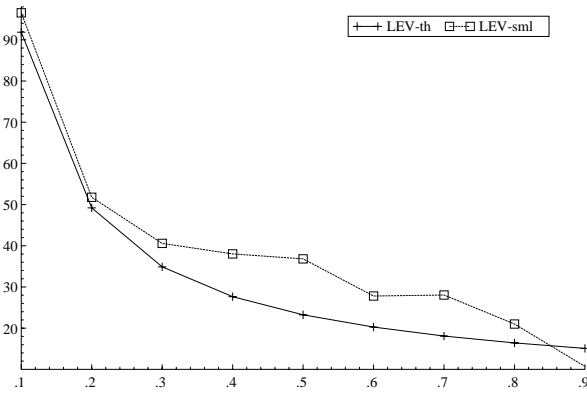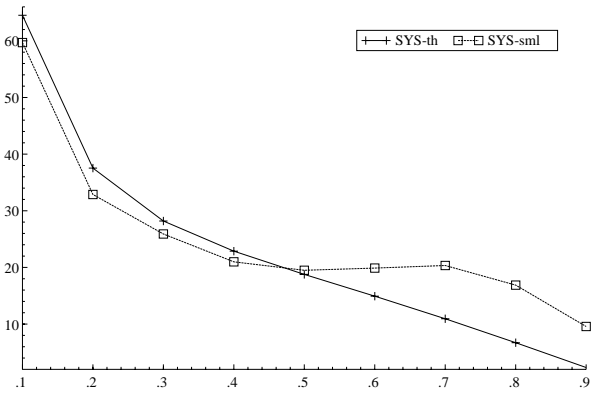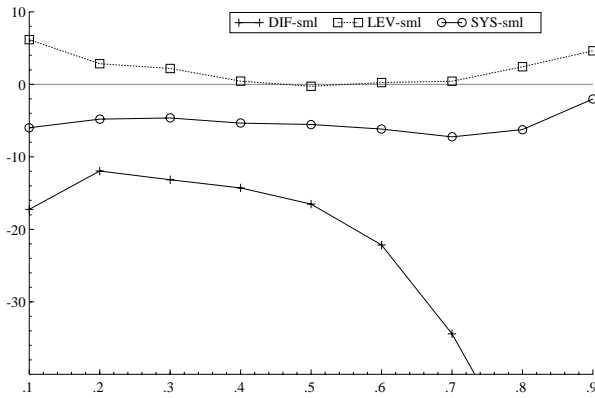Figure 9: Bias of $\widehat{\alpha}$ $(\sigma_\eta^2/\sigma_v^2 = 0.25)$
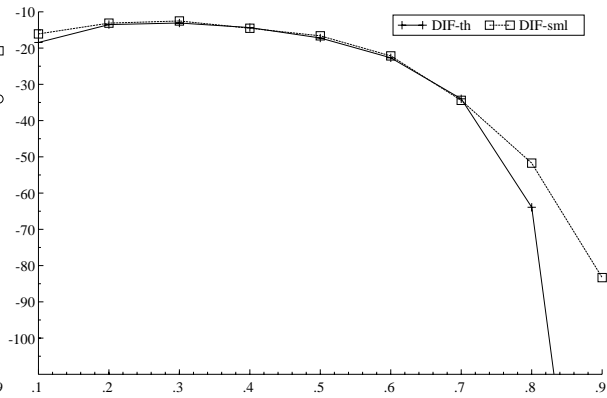


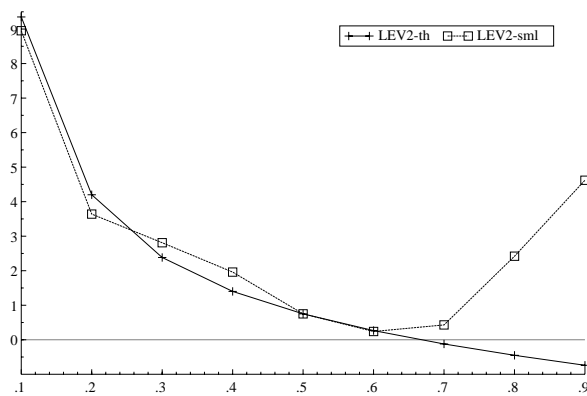Figure 10: Bias of $\widehat{\alpha}^{dif}$ $(\sigma_\eta^2/\sigma_v^2 = 0.25)$



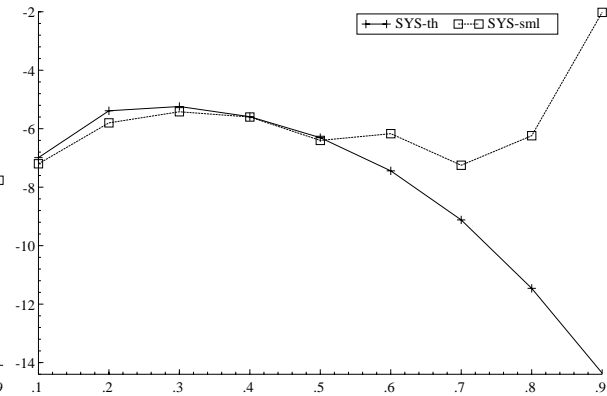Figure 11: Bias of $\widehat{\alpha}^{lev}$ $(\sigma_\eta^2/\sigma_v^2 = 0.25)$



Figure 12: Bias of $\widehat{\alpha}^{sys}$ $(\sigma_\eta^2/\sigma_v^2 = 0.25)$

24

Table 1: Theoretical and simulation values of relative biases (%) when $\sigma_\eta^2/\sigma_v^2 = 1$

|  | GMM(DIF) | | GMM(LEV) | | GMM(SYS) | |
|---|---|---|---|---|---|---|
| $\alpha$ | theory | simulation | theory | simulation | theory | simulation |
| 0.1 | -26.48 | -27.99 | 25.85 | 26.01 | 5.67 | 5.44 |
| 0.2 | -20.02 | -19.79 | 13.20 | 13.42 | 2.35 | 2.26 |
| 0.3 | -20.11 | -20.12 | 8.88 | 8.49 | 0.87 | 0.71 |
| 0.4 | -23.04 | -23.05 | 6.65 | 6.32 | -0.37 | 0.07 |
| 0.5 | -29.11 | -28.00 | 5.25 | 5.19 | -1.77 | -1.25 |
| 0.6 | -40.75 | -38.20 | 4.27 | 6.81 | -3.53 | -0.53 |
| 0.7 | -65.59 | -58.64 | 3.52 | 9.29 | -5.76 | 1.47 |
| 0.8 | -134.23 | -76.93 | 2.92 | 9.85 | -8.43 | 2.56 |
| 0.9 | -488.87 | -93.41 | 2.43 | 10.05 | -11.32 | 5.69 |

Table 2: Theoretical and simulation values of relative biases (%) when $\sigma_\eta^2/\sigma_v^2 = 4$

|  | GMM(DIF) | | GMM(LEV) | | GMM(SYS) | |
|---|---|---|---|---|---|---|
| $\alpha$ | theory | simulation | theory | simulation | theory | simulation |
| 0.1 | -39.65 | -40.28 | 91.85 | 96.55 | 64.52 | 62.48 |
| 0.2 | -29.43 | -31.15 | 49.20 | 51.77 | 37.53 | 35.50 |
| 0.3 | -29.29 | -28.01 | 34.88 | 40.58 | 28.19 | 28.37 |
| 0.4 | -33.46 | -34.14 | 27.65 | 39.65 | 22.87 | 21.20 |
| 0.5 | -42.40 | -44.04 | 23.25 | 36.82 | 18.77 | 20.72 |
| 0.6 | -59.85 | -59.04 | 20.27 | 27.81 | 14.92 | 19.88 |
| 0.7 | -97.65 | -71.66 | 18.09 | 28.05 | 10.94 | 20.34 |
| 0.8 | -203.64 | -86.89 | 16.42 | 21.01 | 6.71 | 16.87 |
| 0.9 | -759.58 | -101.12 | 15.09 | 10.71 | 2.32 | 9.57 |

Table 3: Theoretical and simulation values of relative biases (%) when $\sigma_\eta^2/\sigma_v^2 = 0.25$

| | GMM(DIF) | | GMM(LEV) | | GMM(SYS) | |
| $\alpha$ | Theory | Simulation | Theory | Simulation | Theory | Simulation |
|---|---|---|---|---|---|---|
| 0.1 | -18.46 | -16.09 | 9.35 | 8.95 | -6.99 | -7.20 |
| 0.2 | -13.53 | -13.12 | 4.20 | 3.64 | -5.39 | -5.80 |
| 0.3 | -13.09 | -12.53 | 2.38 | 2.81 | -5.24 | -5.42 |
| 0.4 | -14.35 | -14.53 | 1.40 | 1.96 | -5.59 | -5.60 |
| 0.5 | -17.22 | -16.62 | 0.75 | 0.75 | -6.31 | -6.40 |
| 0.6 | -22.68 | -22.16 | 0.27 | 0.24 | -7.44 | -6.17 |
| 0.7 | -34.00 | -34.42 | -0.12 | 0.43 | -9.12 | -7.25 |
| 0.8 | -63.92 | -51.74 | -0.45 | 2.42 | -11.46 | -6.24 |
| 0.9 | -209.96 | -83.32 | -0.74 | 4.62 | -14.36 | -2.02 |

Table 4: Composition of the bias of the GMM(SYS) estimator

| $\alpha$ | Weighted sum | Correlation | $Bias(\hat{\alpha}^{sys})$ |
|---|---|---|---|
| 0.1 | -4.571 | 10.243 | 5.672 |
| 0.2 | -3.665 | 6.017 | 2.352 |
| 0.3 | -3.673 | 4.543 | 0.869 |
| 0.4 | -4.004 | 3.635 | -0.369 |
| 0.5 | -4.571 | 2.802 | -1.769 |
| 0.6 | -5.366 | 1.834 | -3.531 |
| 0.7 | -6.384 | 0.623 | -5.761 |
| 0.8 | -7.586 | -0.844 | -8.430 |
| 0.9 | -8.856 | -2.462 | -11.318 |

Table 5: Theoretical values of $\gamma$

| $\alpha$ | $\sigma_\eta^2/\sigma_v^2 = 1$ | | $\sigma_\eta^2/\sigma_v^2 = 4$ | | $\sigma_\eta^2/\sigma_v^2 = 0.25$ | |
|---|---|---|---|---|---|---|
| | $\gamma$ | $1-\gamma$ | $\gamma$ | $1-\gamma$ | $\gamma$ | $1-\gamma$ |
| 0.1 | 0.491 | 0.509 | 0.378 | 0.622 | 0.586 | 0.414 |
| 0.2 | 0.437 | 0.563 | 0.323 | 0.677 | 0.545 | 0.455 |
| 0.3 | 0.378 | 0.622 | 0.267 | 0.733 | 0.497 | 0.503 |
| 0.4 | 0.314 | 0.686 | 0.211 | 0.789 | 0.441 | 0.559 |
| 0.5 | 0.245 | 0.755 | 0.156 | 0.844 | 0.374 | 0.626 |
| 0.6 | 0.176 | 0.824 | 0.106 | 0.894 | 0.297 | 0.703 |
| 0.7 | 0.110 | 0.890 | 0.063 | 0.937 | 0.209 | 0.791 |
| 0.8 | 0.053 | 0.947 | 0.029 | 0.971 | 0.117 | 0.883 |
| 0.9 | 0.014 | 0.986 | 0.007 | 0.993 | 0.037 | 0.963 |

Table 6: Bias-corrected GMM estimators ($\sigma_\eta^2/\sigma_v^2 = 4$)

| $\alpha$ | | $\hat{\alpha}^{dif}$ | $\hat{\hat{\alpha}}^{dif}$ | $\hat{\alpha}^{lev}$ | $\hat{\hat{\alpha}}^{lev}$ | $\hat{\alpha}^{sys}$ | $\hat{\hat{\alpha}}^{sys}$ |
|---|---|---|---|---|---|---|---|
| 0.1 | mean | 0.0597 | 0.1008 | 0.1966 | 0.0928 | 0.1597 | 0.1114 |
| | Bias(%) | -40.28 | 0.80 | 96.55 | -7.15 | 59.69 | 11.38 |
| | SD | 0.2637 | 0.3068 | 0.2768 | 0.5016 | 0.1972 | 0.2395 |
| | RMSE | 0.2668 | 0.3068 | 0.2932 | 0.5016 | 0.2061 | 0.2398 |
| 0.2 | mean | 0.1377 | 0.1995 | 0.3035 | 0.1854 | 0.2657 | 0.2083 |
| | Bias(%) | -31.15 | -0.23 | 51.77 | -7.29 | 32.87 | 4.14 |
| | SD | 0.2973 | 0.3964 | 0.2938 | 0.8575 | 0.2115 | 0.2788 |
| | RMSE | 0.3038 | 0.3964 | 0.3115 | 0.8577 | 0.2215 | 0.2790 |
| 0.3 | mean | 0.2167 | 0.3221 | 0.4261 | 0.3031 | 0.3777 | 0.3112 |
| | Bias(%) | -27.77 | 7.36 | 42.04 | 1.03 | 25.90 | 3.75 |
| | SD | 0.3791 | 0.7314 | 0.3060 | 0.9655 | 0.2171 | 0.3171 |
| | RMSE | 0.3881 | 0.7318 | 0.3310 | 0.9655 | 0.2306 | 0.3173 |
| 0.4 | mean | 0.2634 | 0.3910 | 0.5586 | 0.3975 | 0.4840 | 0.3948 |
| | Bias(%) | -34.14 | -2.26 | 39.65 | -0.64 | 20.99 | -1.30 |
| | SD | 0.4467 | 0.9166 | 0.3228 | 1.3210 | 0.2314 | 0.4037 |
| | RMSE | 0.4671 | 0.9166 | 0.3597 | 1.3210 | 0.2462 | 0.4038 |
| 0.5 | mean | 0.2810 | 0.4913 | 0.6941 | 0.4869 | 0.5975 | 0.4827 |
| | Bias(%) | -43.80 | -1.74 | 38.81 | -2.61 | 19.49 | -3.46 |
| | SD | 0.5487 | 2.3790 | 0.3301 | 2.6236 | 0.2262 | 0.5269 |
| | RMSE | 0.5908 | 2.3790 | 0.3829 | 2.6237 | 0.2463 | 0.5272 |