

The Performance of Alternative General Method of Moment Estimators with Application to Employment Panel Data Model

Tegodie Hibstu

Department of Statistics, Addis Ababa University, Ethiopia.

Corresponding author: yitbarekhibstu@gmail.com, tegodie.hibstu@ecs.u.edu.et

Emmanuel Gabreyohannes

Department of Mathematics and Statistics, Ethiopian Civil Service University, Ethiopia.

Eshetu Wencheke

Department of Statistics, Addis Ababa University, Ethiopia.

Abstract

Dynamic panel data models make it possible to address dynamic economic relationships through the inclusion of a lagged dependent variable among the explanatory variables. In the presence of a lagged dependent variable as a regressor, however, the least squares-based estimators may not be consistent. The generalized method of moments (GMM) estimators are the most popular in dynamic panel data estimation. Two issues concerning the GMM estimator are instruments proliferation and the choice of an initial weighting matrix. In this study, we propose alternative estimators that make use of sub-optimal weight matrices in combination with reduced instruments. The small sample performance of these estimators was investigated using a simulation study and an empirical example of panel data of employment for manufacturing firms. The results from Monte Carlo experiments and real data

analysis show that the performance of sub-optimally weighted system GMM estimators that utilize collapsed as well as untransformed instruments seems promising.

Keywords: panel data, GMM, instruments proliferation, reduced instruments set, sub-optimal weight matrices.

1. Introduction

In econometrics, panel data are multi-dimensional data involving repeated observations on the same cross-sectional units (individuals, households, firms, countries, etc.) over time. Panel data provides a means of resolving econometric problems that often arise in empirical studies, specifically the presence of omitted (unobserved) variables that are correlated with explanatory variables. Panel data estimation can be applied either in static panel model or dynamic panel model approaches. In panel data analysis, especially in static panel data models, one of the important issues is whether the unobserved individual heterogeneity is correlated with the regressors or not. If the individual-specific effects are not correlated with the observed regressors, we commonly adopt random effects estimators. In this case, the unobservable individual effects are assumed to be random variables distributed independently of the regressors. Since the error components of the random effects model may be correlated over time for a given individual and/or may be heteroskedastic, we employ the feasible generalized least-square estimation procedures; if the individual-specific effects are correlated with the observed regressors, we employ fixed effects estimators. In this case, we apply transformations to the model equation (often, but not exclusively, first differencing) to remove the individual-specific effects and then apply standard least-square estimation procedures.

If lagged dependent variable terms are included in the usual individual-specific effects panel data model, the model is said to be a dynamic model. In this case, the strict exogeneity of the regressors no longer holds. It is well known that the fixed effects estimator is inconsistent for a panel with a small number of time periods (T) and large cross-sectional units (n). This inconsistency is referred to as dynamic panel bias or Nickel's bias (Nickel,

1981). The standard approach in cases where the explanatory variables are correlated with the error terms is the instrumental variables estimation procedure. In dynamic panel data estimation, Arellano and Bond (1991), Arellano and Bover (1995), and Blundell and Bond (1998) proposed the generalized method of moments (GMM) estimators. One of the attractive properties of GMM estimators is that they are constructed from the sample moment counterparts of the population moment conditions without a complete specification of the distribution of the data.

Several studies in the context of Southeast Asia have utilized panel data analysis. These include investigation of the role of innovation in moderating the influence of internet usage on CO₂ emissions using fixed effects and random effects estimators (Le Hoang et al., 2023), examination of the impact of financial sources on carbon emissions based on pooled mean group estimation for large time dimension panel series (Jannico, 2023), assessment of manufacturing firms' productivity and efficiency from aligning with sustainable development goals using two-stage least-squares, limited-information maximum likelihood and GMM estimators (Amornkitvikai & Pholphirul, 2023), investigation of the impact of various economic, social, and environmental factors on health status employing GMM and panel fully-modified ordinary least-square estimators (Majumder et al., 2022), analysis of the impact of human capital efficiency on firm performance utilizing the difference GMM estimator (Tran & Hong, 2020), and assessment of the relationship between corporate debt maturity and firm performance using the system GMM estimator (Nguyen & Nguyen, 2020). Some of the above empirical studies applied static panel data estimation procedures. In this regard, the use of the dynamic panel data model might provide more insights since it captures the dynamics of the variables of interest. Those authors applying

dynamic panel data model adopted the conventional system GMM estimator and panel cointegration approaches (for long time dimension panel series).

In dynamic panel data estimation, the number of GMM estimators is growing exponentially in the literature. However, it is difficult to make a reasoned choice between many different possible implementations of these estimators. For example, using many instruments may improve the efficiency of GMM estimators, though the asymptotically efficient GMM estimators may suffer from substantial bias when all available moment conditions are employed for small cross-sectional dimension sample sizes, n (Hall, 2005; Bun & Kiviet, 2006; Bun & Windmeijer, 2010; Kiviet et al., 2015). Moreover, the choice of an initial weighting matrix is a crucial issue in GMM estimation procedures.

In addressing these issues, Roodman (2009) tried to address GMM estimation problems due to a large number of instruments. He introduced the commonly used instrument reduction techniques (lag-limited and collapsed instruments set), but he and other authors working on moment reduction techniques did not address the issue of an initial weight matrix in GMM estimation. On the other hand, authors like Kiviet (2007), Jung et al. (2015), and Youssef and Abonazel (2017) tried to address the issue of an initial weight matrix; however, they did not address estimation problems due to instruments proliferation. Hibstu et al. (2023) reported that the two-step system GMM estimator, which employs a sub-optimal weight matrix in combination with collapsed instrument set, outperformed the conventional system GMM estimator in terms of bias and root mean square error.

In this paper, we compare the performance of one-step system GMM estimators with various implementations (in terms of the instruments and the initial weight matrix utilized). These are the system GMM estimator as in Blundell and Bond (1998), the commonly used conventional system GMM

estimator (Bun & Windmeijer, 2010), the sub-optimal system GMM estimator based on all available instruments (e.g., Jung et al., 2015), and the proposed one-step system GMM estimator utilizing the sub-optimal weight matrix in combination with a reduced instruments set. The performance of the estimators was compared in terms of bias and standard deviation.

The next sections of this paper are organized as follows. Section 2 introduces the GMM estimator for the simple dynamic panel model. Estimation problems in connection with instruments proliferation and possible solutions are presented in Section 3. Section 4 discusses various initial weighting matrices. Section 5 presents the simulations results, and Section 6 is devoted to an empirical illustration of the employment panel data model. Conclusions are presented in Section 7.

2. The GMM Estimator

Consider a simple dynamic AR(1) panel data process:

$$\begin{aligned} y_{it} &= \gamma y_{i,t-1} + u_{it}, \quad |\gamma| < 1 \\ u_{it} &= \alpha_i + \varepsilon_{it}, \quad i = 1, \dots, n; t = 1, \dots, T, \end{aligned} \quad (1)$$

where y_{it} is an observable variable for individual i in period t , α_i is the unobserved individual specific effect, and ε_{it} is the idiosyncratic error term. The number of individuals (n) for which data is available is assumed to be large and the number of time periods (T) is small. The assumptions about the dynamic panel model are:

i) The idiosyncratic error terms and individual specific effects are identically and independently distributed: $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$, $\alpha_i \sim iid(0, \sigma_\alpha^2)$,

ii) ε_{it} and α_i are uncorrelated: $E(\alpha_i \varepsilon_{it}) = 0$, and

iii) The initial observation satisfies: $y_{i1} = [\alpha_i / (1 - \gamma)] + \omega_{i1}$, $i = 1, \dots, n$ where

$\omega_{i1} = \sum_{j=0}^{\infty} \gamma^j \varepsilon_{i,1-j}$ is uncorrelated with both α_i and ε_{it} , and ω_{i1} is covariance stationary.

Consider the first difference version of the simple autoregressive panel in (1):

$$\Delta y_{it} = \gamma \Delta y_{i,t-1} + \Delta \varepsilon_{it} : i = 1, \dots, n, t = 1, \dots, T \quad (2)$$

Under the assumptions in (i–iii), the following $(T-2) \times [(T-2)(T-1)/2]$ linear moment conditions are valid:

$$E[y_{i,t-s} \Delta \varepsilon_{it}] = 0, \quad t = 3, \dots, T; 2 \leq s \leq t-1; i = 1, \dots, n \quad (3)$$

The moment conditions in (3) can be written concisely as:

$$E(Z_i^D' \Delta \varepsilon_i) = 0 \quad (4)$$

where $\Delta \varepsilon_i = (\Delta \varepsilon_{i3}, \Delta \varepsilon_{i4}, \dots, \Delta \varepsilon_{iT})'$ and Z_i^D is the $(T-2) \times [(T-2)(T-1)/2]$ matrix of instruments for the lagged difference given by:

$$Z_i^D = \begin{bmatrix} y_{i,1} & 0 & \cdots & 0 \\ 0 & y_{i,1}, y_{i,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & y_{i,1}, \dots, y_{i,T-2} \end{bmatrix} \quad (5)$$

The one-step Arellano and Bond (1991) estimator is given by:

$$\hat{\gamma}^{D1} = \frac{(\Delta y_{-1}' Z^D W_n^D Z^D \Delta y)}{(\Delta y_{-1}' Z^D W^D Z^D \Delta y_{-1})} \quad (6)$$

where $\Delta y = (\Delta y_1', \dots, \Delta y_n')$, $\Delta y_i = (\Delta y_{i3}, \dots, \Delta y_{iT})'$, Δy_{-1} is the lagged version of Δy ,

$Z^D = (Z_1^D, Z_2^D, \dots, Z_n^D)$, $W_n^D = \left(n^{-1} \sum_{i=1}^n Z_i^D D Z_i^D \right)^{-1}$ is the initial weight matrix that minimizes the distance between actual population and sample values, and D is the matrix given by:

$$D = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{bmatrix}_{(T-2) \times (T-2)}$$

However, this estimator does not identify its true value when γ gets close to one and/or the variance ratio of the individual specific effects to that of idiosyncratic error terms ($\rho = \sigma_\alpha^2 / \sigma_\varepsilon^2$) goes to infinity (Blundell & Bond, 1998). As a solution, Blundell and Bond (1998) proposed the system GMM estimator which employs additional moment conditions. The instruments matrix for Blundell and Bond's estimator includes lagged levels as well as lagged differences under the assumption of stationarity on the initial conditions. The moment conditions in system GMM estimation are defined as:

$$E[y_{i,t-s} \Delta \varepsilon_{it}] = 0, \quad t = 3, \dots, T; \quad 2 \leq s \leq (t-1); \quad i = 1, \dots, n \quad (7)$$

$$E[\Delta y_{i,t-1} u_{it}] = 0, \quad t = 3, \dots, T, \quad i = 1, \dots, n \quad (8)$$

The instruments matrix for the system GMM estimator is given by:

$$Z_i^S = \begin{pmatrix} Z_i^D & 0 \\ 0 & Z_i^L \end{pmatrix}, \text{ where } Z_i^D \text{ is as defined in (5) and } Z_i^L = \text{diag}(\Delta y_{i2}, \dots, \Delta y_{iT-1})$$

is the instruments set for the level equation. Using the moment conditions in (7) and (8), we have:

$$E(Z_i^S u_i^S) = 0 \quad (9)$$

where $u_i^s = \begin{pmatrix} \Delta u_i \\ u_i \end{pmatrix}$. The one-step system GMM estimator is then computed as:

$$\hat{\gamma}^{S1} = \frac{(\tilde{y}_{-1}^s)' Z^s W_n^s Z^s \tilde{y}^s}{(\tilde{y}_{-1}^s)' Z^s W_n^s Z^s \tilde{y}_{-1}^s} \quad (10)$$

where $\tilde{y}_{-1}^s = [(\Delta y_{1,-1}, y_{1,-1})', \dots, (\Delta y_{n,-1}, y_{n,-1})']'$, $\tilde{y}^s = [(\Delta y_1, y_1)', \dots, (\Delta y_n, y_n)']'$, $Z^s = (Z_1^s, \dots, Z_n^s)$, and W_n^s is a positive definite weighting matrix that yields a consistent estimator of γ . It is well documented that if the population moment conditions are true, say $E[g(y_i, \gamma)] = 0$, where $g(\cdot)$ denotes a vector of moment functions, and under certain regularity conditions in place, it can be shown that $\lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n g_n(y_i, \gamma) \xrightarrow{p} E[g(y_i, \gamma)]$ and $\lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} \sum_{i=1}^n g_n(y_i, \gamma) \rightarrow N(0, S)$, where S is finite and positive definite. Furthermore, using the law of large numbers on the moment conditions $\sqrt{n}(\hat{\gamma} - \gamma)$ has a limiting normal distribution, that is, $\sqrt{n}(\hat{\gamma} - \gamma) \rightarrow N(0, V_w)$, where:

$$V_w = (G'WG)^{-1} G'WGSWG(G'WG)^{-1}. \quad (11)$$

The asymptotic variance of the GMM estimator in (11) shows that the efficiency of the estimator is affected by the choice of the weight matrix, W . The most efficient GMM estimator is the one that utilizes the inverse of the asymptotic variance of the sample moments as the weight matrix, $S^{-1} = \lim_{n \rightarrow \infty} \left[\text{var} \left(\frac{1}{\sqrt{n}} g_n(y_i, \gamma) \right) \right]^{-1}$ (Windmeijer (2005)). In this case, the asymptotic variance of the estimator becomes $(G'S^{-1}G)^{-1}$. The problem is that the weight matrix ($W = S^{-1}$) depends on the parameter γ , so this estimator is infeasible. In practice, the initial weight matrix is obtained by an inefficient choice of W (for example, the identity matrix) that does not depend on γ .

3. The Issue of Instruments Proliferation and Instruments Reduction Techniques

In dynamic panel data models, the number of moment restrictions increases rapidly with T . It is well documented that, in small and moderate sample sizes (n), using all available instruments may result in substantial bias and/or a serious underestimation of the true variance of the GMM estimates (Arellano & Bond, 1991; Blundell & Bond, 1998; Doran & Schmidt, 2006; Fajeau, 2020). The problem is more serious in estimating the weight matrix since obtaining precise estimates of the matrix elements is difficult unless the cross-sectional dimension (n) of the panel is large relative to the number of instruments utilized (Ziliak, 1997; Alonso-Borrego & Arellano, 1999; Windmeijer, 2005; Roodman, 2009; Fajeau, 2020). Bun and Kiviet (2006) examined the effects of instruments proliferation, and their result showed that the magnitude of small sample bias of the Arellano-Bond (1991) estimator is of order $O(n^{-1})$ when all available instruments sets are employed. On the other hand, the bias of the estimator reduced into $O(n^{-1}T^{-1})$ when they used lag-limited instruments. Therefore, the number of moment conditions one should include in GMM estimation is an important practical issue. Roodman (2009) proposed the use of a subset (only the recent lags) of instruments and collapsing of instruments set. For example, if we only exploit the moment conditions for a certain lag depth, we can rewrite the moment conditions in (4) as follows:

$$E(y_{i,t-s}\Delta\epsilon_{it})=0, \quad 2 \leq s \leq \tau; \quad t=3, \dots, T \quad (12)$$

where τ is the maximum lag depth to be considered. When $\tau = 2$, for instance, the lag-depth truncation technique reduces the instrument count for the difference GMM estimator to $[(T-2)(T-1)/2] - [(T-4)(T-2)/2]$ for $T \geq 4$, and the instruments matrix in (5) is modified into:

$$(Z_i^D)^L = \begin{bmatrix} y_{i,1} & 0 & \cdots & 0 \\ 0 & y_{i,1}, y_{i,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & y_{i,T-3}, y_{i,T-2} \end{bmatrix}.$$

The instruments matrix in (8) is taken as it is since we have considered only the most recent lagged difference as an instrument for the explanatory variable in level in each time period ($t=3, \dots, T$). The collapsed instruments set is achieved by squeezing the number of columns in (5) based on only their lag distance for all periods but not for each period, and we impose the following restrictions on the moment conditions in (5):

$$E \left[\sum_{s=1}^{t-2} y_{is} \Delta \varepsilon_{it} \right] = 0, \quad t=3, \dots, T \quad (13)$$

The instruments matrix for each individual i now becomes:

$$(Z_i^D)^C = \begin{bmatrix} y_{i,1} & 0 & 0 & \cdots & 0 \\ y_{i2} & y_{i1} & 0 & \cdots & 0 \\ y_{i3} & y_{i2} & y_{i1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ y_{i,T-2} & y_{i,T-3} & y_{i,T-4} & \cdots & y_{i1} \end{bmatrix}.$$

Here, the number of instruments in (5) is reduced into $(T-2)$.

4. Suboptimal Initial Weights Matrix

The small sample properties of GMM estimators depend on the choice of the initial weight matrix. For the first difference GMM estimators, Kiviet (2007) shows that an optimal initial weighting matrix can be easily driven. However, the optimal weight matrix for the one-step system GMM estimator is not easily driven due to its dependence on the nuisance parameter. In practice, the initial weight matrix is constructed either based on convenience

(e.g., identity matrix) or efficiency criteria (Kiviet, 2007; Jung & Kwon, 2007; Jung et al., 2015; Youssef & Abonazel, 2017). The initial weight matrix in Blundell and Bond (1998) is given by:

$$W_n^{BB} = \left[n^{-1} \sum_{i=1}^n Z_i^S {}' H_I Z_i^S \right]^{-1}, \quad H_I^{BB} = \begin{pmatrix} I_2 & 0 \\ 0 & I_{T-2} \end{pmatrix} \quad (14)$$

This yields the simple generalized instrumental variables estimator, which is far from optimal if the idiosyncratic error terms (ε_{it}) are heteroskedastic and/or $\sigma_\alpha^2 > 0$. Windmeijer (2000) applied an operational weight matrix for the one-step GMM estimator given by:

$$W_n^W = \left[n^{-1} \sum_{i=1}^n Z_i^S {}' H_I^W Z_i^S \right]^{-1}, \quad H_I^W = \begin{pmatrix} D & C \\ C' & I_{T-2} \end{pmatrix} \quad (15)$$

$$\text{where } C = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \ddots & \vdots \\ 0 & -1 & 1 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix}_{(T-2) \times (T-2)} \quad \text{and } D \text{ is as defined in (6).}$$

Here, the variance-covariance structure of the error terms (corresponding to the level equation) in the system GMM estimation procedure is not considered. The initial weight matrix setup for the conventional one-step system GMM estimation is given by:

$$W_n^C = \left[n^{-1} \sum_{i=1}^n Z_i^S {}' H_I^C Z_i^S \right]^{-1}, \quad H_I^C = \begin{pmatrix} D & 0 \\ 0 & I_{T-2} \end{pmatrix} \quad (16)$$

System GMM estimators based on initial weights matrices in (14–16) perform well for small values σ_α^2 . However, there exists the potential for efficiency loss when σ_α^2 gets large (Windmeijer, 2000; Youssef & Abonazel, 2017; Jung et al., 2015). Consider the simple dynamic panel model in (1); under

the assumptions stated in Section 2 and assuming that the series is covariance stationary, it can be shown that:

$$E(y_{it} \alpha_i) = \sigma_{y\alpha} = \frac{\sigma_\alpha^2}{1-\gamma} \forall_{i,t} \quad (17)$$

$$\text{var}(y_{it}) = \frac{\sigma_\alpha^2}{(1-\gamma)^2} + \frac{\sigma_\varepsilon^2}{1-\gamma^2} \quad (18)$$

$$E[y_{i,t-1} y_{it}] = \frac{\sigma_\alpha^2}{(1-\gamma)^2} + \frac{\gamma \sigma_\varepsilon^2}{1-\gamma^2} = \sigma_y^2 - \frac{\sigma_\varepsilon^2}{1+\gamma} \quad (19)$$

Kiviet (2007) and Jung & Kwon (2007) suggest the use of a suboptimal weight matrix given by:

$$W_n^J = \left[n^{-1} \sum_{i=1}^n Z_i^S H_J Z_i^S \right]^{-1}, \quad H_J = \begin{pmatrix} D & 0 \\ 0 & J_{T-2} \end{pmatrix},$$

$$J_{T-2} = \begin{pmatrix} 1+\rho & \rho & \rho & \cdots & \rho \\ \rho & 1+\rho & \rho & \cdots & \rho \\ \rho & \rho & 1+\rho & \cdots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \cdots & 1+\rho \end{pmatrix}, \quad \rho = \frac{\sigma_\alpha^2}{\sigma_\varepsilon^2} \quad (20)$$

Since the variance ratio (ρ) is unknown in practice, Jung and Kwon (2007) suggest an estimate of the optimal weight matrix:

$$\hat{H}_J = \begin{pmatrix} D & 0 \\ 0 & \hat{J}_{T-2} \end{pmatrix} \quad (21)$$

where the variance ratio (ρ) in (20) is replaced by the estimated variance ratio $\hat{\rho} = \hat{\sigma}_\alpha^2 / \hat{\sigma}_\varepsilon^2$ in (21). In this study, the performance of sub-optimal system GMM estimators utilizing reduced instruments (collapsed and lag-limited instruments set) is compared against the conventional system GMM estimators and the sub-optimal system GMM estimator utilizing all available instruments in terms of bias and standard deviation. Appendix II of this paper provides

derivations of alternative initial weighing matrices using various instrument sets (untransformed, lag-limited, and partially collapsed) in combination with initial weighting matrices setup as in (16) and suboptimal weight matrices in (21). Under these setups, we have different system GMM estimators since each initial weighting matrix yields a unique system GMM estimator.

5. Simulation Design and the Results

5.1 Simulation Design

In this section, we investigate the performance of system GMM estimators, which utilize various initial weight matrices in combination with different instruments sets in terms of bias and standard deviation. These are the system GMM estimator as in Blundell and Bond (1998), the commonly used conventional system GMM estimator based on an initial weight matrix as in equation (16), the system GMM estimator using sub-optimal weight and utilizing all available instruments (see Jung et al., 2015; Youssef & Abonazel, 2017), and the system GMM estimator utilizing sub-optimal weight matrix and reduced instruments set (suggested by Hibstu et al., 2023).

The simulation design was based on the following:

1. The individual effects (α_i) were generated as independent normally distributed random variables with mean zero and variance σ_α^2 , and independent across individuals i .
2. The disturbances (ε_{it}) were generated as independently normally distributed random variables with mean zero and variance σ_ε^2 , and independent across individuals i and time t .

3. We generate the initial conditions (y_{i1}) as: $y_{i1} = [\alpha_i / (1 - \gamma)] + \omega_{i1}$, $i = 1, \dots, n$, where $\omega_{i1} = \sum_{j=0}^{\infty} \gamma^j \omega_{i,1-j}$ is independent of both α_i and ε_{it} , and is covariance stationary.
4. The values of the variance ratio $\rho = \sigma_{\alpha}^2 / \sigma_{\varepsilon}^2$ considered are 0.5, 1, 5, and 10.
5. n is fixed at 100; $T = 4, 10$ and $\gamma = 0.4, 0.9$.
6. For all experiments, we ran 5000 replications, and the results of all separate experiments were obtained by precisely the same series of random numbers.

Three different estimation procedures were considered based on the initial weighting matrices as well as instruments set applied in order to compare the small sample performance of system GMM estimators. These are the one-step sub-optimally weighted system GMM estimators (denoted by SYS_Sub) using an initial weight matrix setup as in (21) in combination with untransformed (all available instruments set), lag-limited, and collapsed instruments set; the commonly used conventional one-step system GMM estimators (denoted by SYS_Conv.) based on an initial weight matrix setup as in (16) in combination with untransformed, lag-limited, and collapsed instruments; and the one-step system GMM estimators (denoted by SYS_BB) based on an initial weight matrix as in Blundell and Bond in combination with untransformed, lag-limited, and collapsed instruments. The simulation results are presented in Tables 1 and 2 (Appendix I) for $T = 4$ and $T = 10$, respectively. Monte Carlo experiments were performed using R4.0.5 software, and comparisons of the estimators were made in terms of mean absolute bias (MAB) and the standard deviation (std. dev.).

5.2 Simulation Results

Figures 1 to 3 pertain to the MAB for one stage system GMM estimators at $T = 4$ and $T = 10$. The properties of small sample estimators for selected values of γ and various levels of estimated variance ratios (ρ) are discussed below.

5.2.1 Comparison of Estimators Based on Mean Absolute Bias

For $T = 4$ and $\gamma = 0.4$ (Figure 1), there is no marked difference in the MAB among all estimators for small values of ρ ($\rho = 0.5, 1$). For moderate and large values of ρ , however, the sub-optimally weighted GMM estimator based on collapsed instruments is the least biased. In contrast, the conventional one-step system GMM estimator and that based on an initial weight matrix, as in Blundell and Bond (1998), are the most biased. Moreover, the MAB of all estimators keeps on increasing as the variance ratio increases. For $T = 4$ and $\gamma = 0.9$ (Figure 2), there is no marked difference in terms of bias across the board regardless of the values of the variance ratios. For $T = 10$ and $\gamma = 0.4$ (Figure 3), the sub-optimally weighted GMM estimator that employs collapsed instruments is the least biased, and the difference against competing estimators is more pronounced for moderate and large values of ρ ($\rho = 5, 10$). The estimator based on suboptimal initial weights matrix and untransformed instruments is also successful for $\rho = 0.5$. The reverse is true for the conventional one-step system GMM estimator, and that is based on an initial weight matrix as in Blundell and Bond (1998) for all values of the variance ratios.

5.2.2 Comparison of Estimators Using Standard Deviation

Figure 4 presents the standard deviation of one stage system GMM estimators at $T = 10$ and $\gamma = 0.4$. We can observe that, for a small value of the

variance ratio ($\rho = 0.5$), there is no marked difference in the performance of system GMM estimators (save the sub-optimally weighted estimator based on lag-limited instruments). For moderate and large values of ρ ($\rho = 0.5$ and $\rho = 10$), however, the sub-optimally weighted GMM estimator using all available (or untransformed) instruments is more efficient, followed by those with reduced instruments sets (collapsed and lag-limited). In contrast, the conventional one-step system GMM estimator and that based on an initial weight matrix, as in Blundell and Bond (1998), have relatively large standard deviations, especially when the value of ρ gets large. For $T = 10$ and $\gamma = 0.9$, there is no notable difference across all estimators in terms of precision. We also observe that the dispersion of all estimators decreases as the values of ρ get larger. This might be explained by the underestimation of the true variance of the parameters as the value of γ approaches one (Blundell & Bond, 1998; Doran & Schmidt, 2006; Bun & Windmeijer, 2010).

In general terms, the comparative analysis of estimators using mean absolute bias and standard deviation shows that sub-optimally weighted GMM estimators utilizing collapsed instruments and untransformed instruments fared well, in particular for moderate and large values of the variance ratio (ρ). For $T = 10$ and $\gamma = 0.4$, for instance, the former estimator registered an 8.4% and 23.5% reduction in MAB over the latter for $\rho = 5$ and $\rho = 10$, respectively. On the other hand, the sub-optimally weighted GMM estimator utilizing all available instruments exhibited a 4.2% and 6.3% reduction in standard deviation over the estimator utilizing collapsed instruments for the same pair of variance ratios. Thus, we can conclude that, in relative terms, the substantial reduction in MAB of the sub-optimally weighted GMM estimator based on

collapsed instruments more than compensated for the loss in precision (or imprecision).

One-step system GMM estimator was based on an initial weight matrix as in Blundell-Bond (1998) (**SYS_BB**), conventional one-step system GMM estimator (**SYS_Conv.**), and system GMM estimators based suboptimal initial weights matrix setup as in (21) in combination with untransformed, lag-limited, and collapsed instruments (denoted by, **SYS_Subf**, **SYS_Subl**, and **SYS_Subcol**, respectively) at $\gamma = \{0.4, 0.9\}$ and $T = \{4, 10\}$.

Figure 1. Mean absolute bias for $T = 4$ and $\gamma = 0.4$

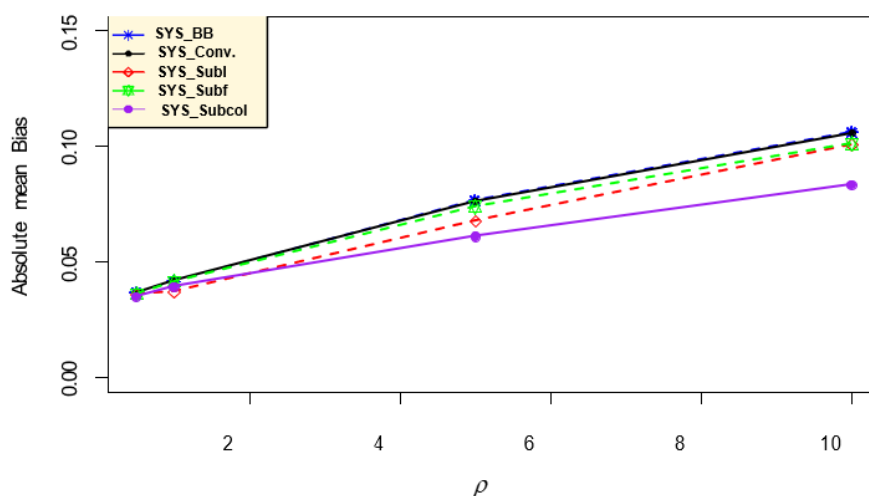


Figure 2. Mean absolute bias for $T = 4$ and $\gamma = 0.4$

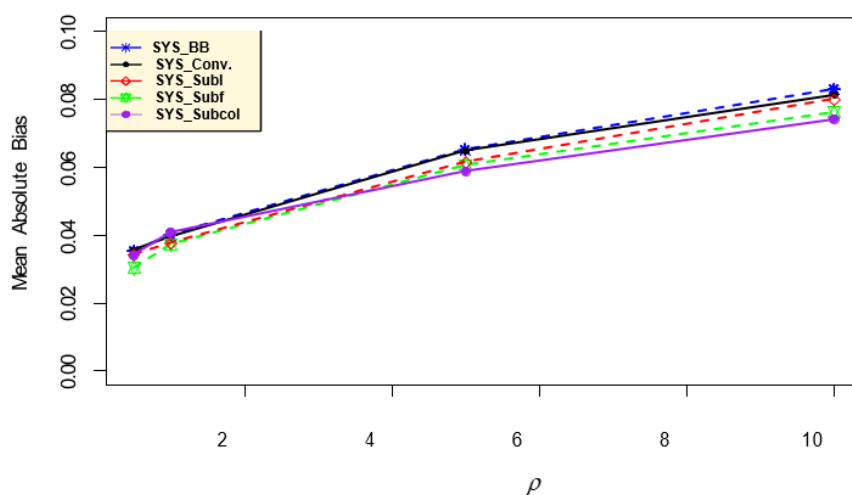
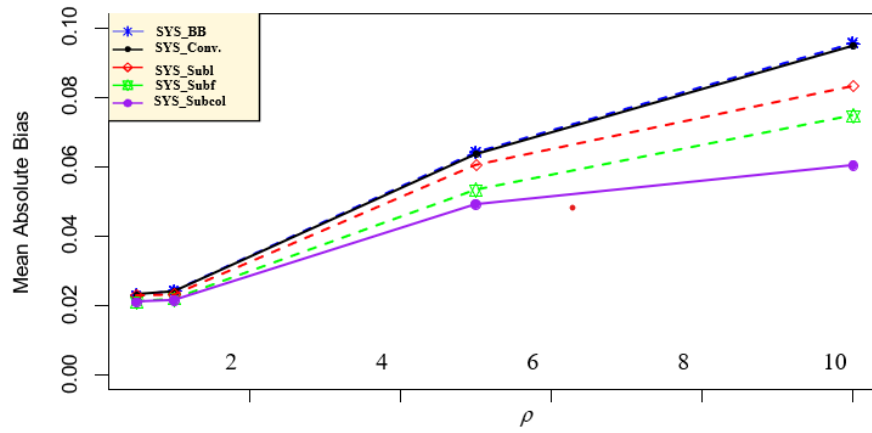
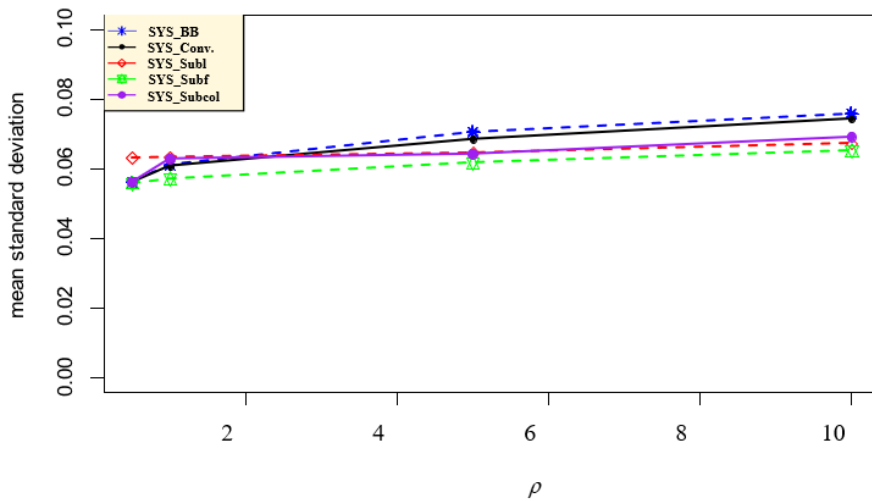


Figure 3. Mean absolute bias for $T = 10$ and $\gamma = 0.4$ Figure 4. Standard deviation for $T = 10$ and $\gamma = 0.4$ 

6. Empirical illustration

In this section, we illustrate the performance of alternative GMM estimators using a real dataset obtained from a series of Large and Medium Scale Manufacturing Industries (LMSMI) surveys conducted in Ethiopia by the Ethiopian Statistical Service (ESS) from 2010 to 2019. The ESS defines large- and medium-scale manufacturing industries as all manufacturing establishments that employ ten or more people and use power-driven machines. This definition is based on the International Standard Industrial Classification (ISIC Revision-3.1). The panel

dataset for our comparative analysis comprised 284 establishments and 2840 observations (balanced panel). Real wage and real working capital were computed by deflating the nominal values using the price deflator series obtained from the Ministry of Planning and Development.

We considered the labor demand (employment) equation of the form:

$$y_{it} = \beta_w w_{it} + \beta_k k_{it} + \delta_t + \alpha_i + \varepsilon_{it} ; i = 1, \dots, n; t = 2, \dots, T$$

$$\varepsilon_{it} = \phi \varepsilon_{i,t-1} + e_{it}, |\phi| < 1; m_{it}, e_{it} \sim MA(0) \quad (22)$$

where y_{it} is log employment of firm i in year t , w_{it} is log real wage, k_{it} is log real working capital, δ_t is a year-specific intercept, α_i is firm specific time-invariant unobserved effects, ε_{it} is a random shock, and m_{it} refers to serially uncorrelated measurement errors. The explanatory variables we considered (w_{it}, k_{it}) were potentially correlated with the individual-specific effects, the random shocks, and measurement errors, so the instrument set for each variable of interest was constructed accordingly. The model had dynamic (common factor) representation defined as:

$$y_{it} = \phi y_{i,t-1} + \beta_w w_{it} - \phi \beta_w w_{i,t-1} + \beta_k k_{it} - \phi \beta_k k_{i,t-1} + (\delta_t - \phi \delta_{t-1}) + \alpha_i (1 - \phi) + e_{it} + m_{it} - \phi m_{i,t-1} \quad (23)$$

Equation (23) can be expressed as:

$$y_{it} = \pi_1 y_{i,t-1} + \pi_2 w_{it} + \pi_3 w_{i,t-1} + \pi_4 k_{it} + \pi_5 k_{i,t-1} + \tilde{\delta}_t + \tilde{\alpha}_i + \eta_{it} \quad (24)$$

subject to two non-linear (common factor) restrictions, $\pi_3 = -\pi_1 \pi_2$ and $\pi_5 = -\pi_1 \pi_4$, where $\tilde{\delta}_t = \delta_t - \phi \delta_{t-1}$, $\tilde{\alpha}_i = \alpha_i (1 - \phi)$, and $\eta_{it} = e_{it} + m_{it} - \phi m_{i,t-1}$. Consequently, both the unrestricted parameter vector $\pi = (\pi_1, \dots, \pi_5)$ and the restricted ones (β_w, β_k, ϕ) can be estimated. Note that η_{it} is serially uncorrelated if there are no measurement errors or $\eta_{it} \sim MA(1)$ otherwise.

During the implementation of various GMM estimation procedures, the validity of the moment restrictions was tested first. We began by using lagged levels dated $(t-2)$ and earlier as instruments for the differenced equation and differenced lags dated $(t-1)$ as instruments for the levels equation. However, the Sargan test of overidentifying restrictions rejects the validity of instruments; this may arise from the presence of measurement errors. Hence, we excluded lagged levels dated $(t-2)$ and differenced lags dated $(t-1)$ from the instruments set and apply $(t-3)$ and earlier level lags as instruments for difference equations and differenced lags dated $(t-2)$ for level equations. The validity of instruments under this specification was confirmed by the Sargan test. The common factor restrictions were accepted for various implementations of system GMM estimators, while they were rejected by OLS, FE, and first difference GMM estimators.

Table 3 (Appendix I) presents the results. The coefficient on the lagged dependent variable is higher for sub-optimally weighted system GMM estimators. Specifically, sub-optimally weighted system GMM estimators utilizing untransformed instruments (denoted by SYS_SubU) and collapsed instruments (denoted by SYS_Subcol) had a larger coefficient for the lagged dependent variable. The standard deviation of these estimates is smaller compared to those from the conventional system GMM estimator and sub-optimally weighted system GMM estimator employing lag-limited instruments. In addition, the coefficient of the lagged dependent variable is higher than within the group estimate and below the OLS levels estimate as expected. One unexpected result is that the elasticity of employment with respect to real wage is positive and significant across all estimators. This is probably attributed to the decline in the real wage over time for LMSMI workers. Our finding is consistent with those of Hailu et al. (2020), who argued that this is likely a reflection of a fixed nominal public wage in the presence of a very high inflation rate. Consequently, a very low real wage rate can be

considered as one of the reasons for unemployment (especially, voluntary unemployment) in the manufacturing sector of Ethiopia. The result may suggest the importance of revising the wage rate (at least a minimum wage rate) in the sector. The elasticity of employment with respect to working capital is positive. However, the coefficients for all estimators are very small, with relatively larger sizes for conventional and proposed estimators. In general, sub-optimal weight system GMM estimators utilizing either untransformed or collapsed instruments may yield reasonable estimates.

7. Conclusion

In this study, we investigated the performance of one-step system GMM estimators with various initial weight matrices as well as instruments. The simulation studies show that the sub-optimally weighted GMM estimator based on collapsed instruments is the least biased for moderate and large values of the variance ratio, and this is more so as the time dimension increases and for relatively smaller autoregressive coefficients (e.g., at $T = 10$ and $\gamma = 0.4$). The reverse is true for the conventional one-step system GMM estimator and that based on an initial weight matrix, as in Blundell and Bond (1998). When the autoregressive coefficient gets close to one, however, there is no marked difference in the mean absolute bias across the board, regardless of the values of the variance ratio.

In terms of precision, there is no notable difference across all estimators for small values of the variance ratio as well as when γ gets close to one. For moderate and large values of the variance ratio, however, the sub-optimally weighted GMM estimator using all available instruments is the most efficient, followed by utilizing the collapsed instruments set.

In general terms, the performance of the sub-optimally weighted system GMM estimators, which utilize collapsed as well as untransformed instruments, seems promising. Furthermore, as the value of T increases, the bias reduction from the former estimator outweighs the efficiency gain from the latter estimator in relative terms. In this regard, utilizing these estimators (instead of the conventional system GMM estimators) might provide more insights in panel data analysis of various issues involving the economies of Southeast Asia and across the globe.

Acknowledgments

We authors would like to thank the Chief Editor (Vasileios Zikos, PhD) and anonymous referees for their thoughtful comments and efforts toward improving our manuscript.

References

- Amornkitvikai, Y., & Pholphirul, P. (2023). Business productivity and efficiency from aligning with sustainable development goals: Empirical evidence from ASEAN manufacturing firms. *Business Strategy and Development*, 6(2), 189-204.
- Arellano, M., & Bond, R. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277-297.
- Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-component models. *Journal of Econometrics*, 68, 29-51.
- Bigsten, A., Gebreeyesus, M., & Söderbom, M. (2016). Tariffs and firm performance in Ethiopia. *The Journal of Development Studies*, 52(7), 986-1001.
- Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87, 115 -143.
- Blundell, R., & Bond, S. (2000). GMM estimation with persistent panel data: An application to production functions. *Econometric Reviews*, 19(3), 321-340.
- Blundell, R., Bond, S., & Windmeijer, F. (2000). Estimation in dynamic panel data models: Improving on the performance of the standard GMM estimator. *Non-stationary Panels, Panel Cointegration and Dynamic Panels*, 15, 53-91.
- Bun, M. J. G., & Kiviet, J. F. (2006). The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models. *Journal of Econometrics*, 132, 409-444.
- Bun, M.J .G., & Windmeijer, F. (2010). The weak instrument problem of the system GMM estimator in dynamic panel data models. *Econometrics Journal*, 13, 95-126.

- Bun, M. J. G., & Poldermans, R. W. (2015). *Weak identification robust inference in dynamic panel data models*. Tinbergen Institute and Amsterdam School of Economics, University of Amsterdam.
- Cabañero, J. G. (2023). Do financing sources affect CO2 emissions? The case of growing ASEAN. *Southeast Asian Journal of Economics*, 11(1), 159-187.
- Doran, H. E., & Schmidt, P. (2006). GMM estimators with improved finite sample properties using principal components of the weighting matrix, with an application to the dynamic panel data model. *Journal of Econometrics*, 133, 387-409.
- Fajeau, M. (2020). Too much finance or too many weak instruments? *International Economics*, 165(3), 14-16.
- Hailu, K. B., Gebreeyesus, M., & Tekleselassie, T. (2020). *Ethiopia productivity report*. National Graduate Institute for Policy Studies.
- Hall, A. R. (2005). *Generalized method of moments*. Oxford University Press.
- Hibstu, T., Gabreyohannes.E., & Wencheko, E. (2023). Alternative general method of moments estimators in dynamic panel data models. *Communications in Statistics-Simulation and Computation*. 1-18.
- Hoang, V. L., & Hoang, M. N. (2023). Impacts of internet usage on CO2 emissions under the moderation of innovative activities: The case of Southeast Asian countries. *Southeast Asian Journal of Economics*, 11(2), 83-105.
- Jung, H., & Kwon, H. U. (2007). *An alternative system GMM estimation in dynamic panel models*. Discussion Paper Series No. 217. Hitotsubashi University Repository.
- Jung, H., Kwon, H. U., & Jeon, G. (2015). An alternative system GMM estimation in dynamic panel models. *Journal of Economic Theory and Econometrics*, 26(2).

- Kiviet, J. F. (2007). On the optimal weighting matrix for the GMM system estimator in dynamic panel data models. Discussion Paper No. 2007/02 (2007b). Tinbergen Institute.
- Koenker, R., & Machado, J. A. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, 94(448), 1296-1310.
- Majumder, S. C., Dilanchiev, A., & Rahman, H. (2022). Examination of a health production function: Evidence from South Asian countries. *Southeast Asian Journal of Economics*, 10(3), 75-102.
- Nguyen, L., & Nguyen, C. (2020). Financial constraints, corporate debt maturity and firm performance: The case of firms in Southeast Asian countries. *Afro-Asian Journal of Finance and Accounting*, 10(1), 48-59.
- Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71(1).
- Tran, N. P., & Hong, V. D. (2020). Human capital efficiency and firm performance across sectors in an emerging market. *Cogent Business & Management*, 7(1).
- Windmeijer, F. (2000). Efficiency comparisons for a system GMM estimator in dynamic panel data models. *Innovations in Multivariate Statistical Analysis*, 36.
- Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM. *Journal of Econometrics*, 126, 25-51
- Youssef, A. H., & Abonazel, M. R.(2017). *Alternative GMM estimators for first-order autoregressive panel model: An improving efficiency approach*. MPRA Paper 68674, University Library of Munich, Germany.

Appendices

Appendix I

Table 1. Mean absolute bias and standard deviation of alternative one-step GMM estimators based on original (untransformed), lag-limited, and collapsed instruments from top to bottom respectively ($T = 4$)

	ρ	SYS_BB	SYS_Conv.	SYS_Sub
$\gamma = 0.4$	0.5	0.0207(0.0806)	0.0205(0.0806)	0.0203(0.0791)
	1	0.0216(0.0814)	0.0213(0.0811)	0.0210(0.0803)
	5	0.0565(0.1082)	0.0542(0.1060)	0.0440(0.0922)
	10	0.0789(0.1190)	0.0757(0.1032)	0.0612(0.1000)
Untransformed IVs				
$\gamma = 0.9$	0.5	0.0351(0.1009)	0.0355(0.1006)	0.0303(0.1003)
	1	0.0400(0.1003)	0.0394(0.0942)	0.0373(0.0950)
	5	0.0651(0.0871)	0.0649(0.0875)	0.0607(0.0883)
	10	0.0829(0.0723)	0.0813(0.0826)	0.0763(0.0839)
$\gamma = 0.4$	0.5	0.0210(0.0863)	0.0209(0.0861)	0.0206(0.0859)
	1	0.0217(0.0870)	0.0217(0.0867)	0.0213(0.0863)
	5	0.0568(0.0882)	0.0533(0.1100)	0.0478(0.0939)
	10	0.0785(0.1038)	0.0741(0.1031)	0.0628(0.1198)
Lag-limited IVs				
$\gamma = 0.9$	0.5	0.0342(0.0995)	0.0358(0.0993)	0.0342(0.1080)
	1	0.0414(0.0981)	0.0415(0.0983)	0.0377(0.0979)
	5	0.0666(0.0848)	0.0663(0.0849)	0.0615(0.0886)
	10	0.0885(0.0825)	0.0885(0.0824)	0.0800(0.0865)
$\gamma = 0.4$	0.5	0.0205(0.0874)	0.0203(0.0862)	0.0199(0.0867)
	1	0.0211(0.0876)	0.0210(0.0871)	0.0208(0.0969)
	5	0.0555(0.0978)	0.0537(0.0953)	0.0406(0.1095)
	10	0.0745(0.1210)	0.0708(0.1140)	0.0609(0.1183)
Collapsed IVs				
$\gamma = 0.9$	0.5	0.0368(0.0952)	0.0371(0.0956)	0.0341(0.1069)
	1	0.0414(0.0937)	0.0415(0.0937)	0.0407(0.0980)
	5	0.0660(0.0902)	0.0661(0.0908)	0.0588(0.0869)

	10	0.0883(0.0881)	0.0866(0.0888)	0.0741(0.0854)
--	----	----------------	----------------	----------------

Note: In Tables 1 and 2, the standard deviation is presented with in parentheses. We applied bootstrap approach replicates so that the standard deviation is an estimate of the standard error of the coefficients.

Table 2. Mean absolute bias and standard deviation of alternative one-step GMM estimators based on original (untransformed), lag-limited, and collapsed instruments from top to bottom respectively ($T = 10$)

	ρ	SYS_BB	SYS_Conv.	SYS_Sub
$\gamma = 0.4$	0.5	0.0229(0.0560)	0.0231(0.0560)	0.0211(0.0558)
	1	0.0242(0.0611)	0.0497(0.0607)	0.0219(0.0571)
	5	0.0641(0.0705)	0.0635(0.0684)	0.0532(0.0617)
	10	0.0955(0.0758)	0.0948 (0.0743)	0.0746(0.0651)
Untransformed IVs				
$\gamma = 0.9$	0.5	0.0351(0.0516)	0.0355(0.0513)	0.0303(0.0522)
	1	0.0400(0.0516)	0.0394(0.0512)	0.0373(0.0520)
	5	0.0651(0.0456)	0.0649(0.0452)	0.0607(0.0465)
	10	0.0829(0.0429)	0.0813(0.0438)	0.0763(0.0441)
$\gamma = 0.4$	0.5	0.0224(0.0508)	0.0224(0.0630)	0.0221(0.0631)
	1	0.0227(0.0667)	0.0226(0.0665)	0.0222(0.0633)
	5	0.0673(0.0694)	0.0661(0.0690)	0.0605(0.0645)
	10	0.0903(0.0738)	0.0895(0.0731)	0.0832(0.0672)
Lag-limited IVs				
$\gamma = 0.9$	0.5	0.0331(0.0556)	0.0320(0.0553)	0.0342(0.0547)
	1	0.0597(0.0552)	0.0581(0.0551)	0.0377(0.0547)
	5	0.0914(0.0469)	0.0879(0.0475)	0.0615(0.0488)
	10	0.0924(0.0449)	0.0934(0.0448)	0.0800(0.0473)
$\gamma = 0.4$	0.5	0.0217(0.0544)	0.0215(0.0537)	0.0209(0.0561)
	1	0.0220(0.0684)	0.0219(0.0576)	0.0214(0.0629)
	5	0.0523(0.0770)	0.0518(0.0733)	0.0491(0.0643)
	10	0.0632(0.0894)	0.0631(0.0783)	0.0604(0.0692)
Collapsed IVs				
$\gamma = 0.9$	0.5	0.0382(0.0504)	0.0387(0.0504)	0.0341(0.0510)
	1	0.0619(0.0503)	0.0612(0.0501)	0.0407(0.0508)
	5	0.0882(0.0466)	0.0882(0.0468)	0.0588(0.0471)
	10	0.0930(0.0460)	0.0930(0.0479)	0.0741(0.0463)

Table 3. The estimates from empirical example using employment equation

	OLS	FE	DIF	SYS_Conv	SYS_SubU	SYS_SubI	SYS_Subcol
$y_{i,t-1}$	0.8520 (0.0128)	0.0432 (0.0238)	0.0864 (0.1083)	0.7802 (0.0637)	0.7882 (0.0516)	0.7648 (0.0761)	0.7865 (0.0608)
w_{it}	0.5092 (0.0107)	0.2568 (0.0294)	0.4310 (0.0894)	0.5340 (0.0530)	0.5291 (0.0508)	0.4809 (0.0666)	0.5062 (0.0529)
$w_{i,t-1}$	-0.1736 (0.0114)	-0.2004 (0.0194)	-0.0210 (0.0745)	-0.4202 (0.0531)	-0.4542 (0.0499)	-0.4170 (0.0682)	-0.4488 (0.0524)
k_{it}	0.0518 (0.0082)	0.0108 (0.0073)	0.0219 (0.0889)	0.0308 (0.0441)	0.0317 (0.0415)	0.0445 (0.0584)	0.0480 (0.0551)
$k_{i,t-1}$	-0.0046 (0.0081)	-0.0167 (0.0071)	0.0190 (0.0610)	-0.0112 (0.0469)	-0.0184 (0.0428)	-0.0389 (0.0480)	-0.0429 (0.0461)
m_1	-	-	-9.320706	-6.5911	-6.6201	-7.8140	-10.0087
m_2	-	-	-1.5674	-2.3671	-2.3085	-2.2760	-1.8944
Sargan	-	-	0.3075	0.1187	0.1920	0.2733	0.4600
β_w	0.5589 (0.0358)	0.4529 (0.0388)	0.4647 (0.0623)	0.5030 (0.0595)	0.5151 (0.0583)	0.0637 (0.0432)	0.0770 (0.0418)
β_w	0.0532 (0.0326)	0.0110 (0.0469)	0.0493 (0.0508)	0.0754 (0.0381)	0.0791 (0.0375)	0.0637 (0.0432)	0.0770 (0.0418)
ϕ	0.5144 (0.0342)	-0.0442 (0.0399)	-0.3868 (0.0488)	0.3048 (0.0332)	0.2933 (0.0317)	0.3001 (0.0391)	0.3200 (.0346)
Comfac.	0.0012	0.0092	0.0187	0.1098	0.1164	0.2153	0.1676

Note: Comfac. = p-values of common factor restrictions; the under the null common factor restrictions are valid.

Appendix II

Derivation of Alternative Initial Weighting Matrices

Consider $T = 4$ for simplicity and let the assumptions under Blundell and Bond (1988) hold. Under the orthogonal conditions in (7) and (8), there are four over-identifying moment conditions:

$$\begin{aligned} E[y_{i1}(\Delta y_{i3} - \gamma \Delta y_{i2})] &= 0; E[y_{i2}(\Delta y_{i3} - \gamma \Delta y_{i2})] = 0, \\ E[\Delta y_{i2}(y_{i3} - \gamma y_{i2})] &= 0, \text{ and } E[\Delta y_{i3}(y_{i3} - \gamma y_{i2})] = 0. \end{aligned}$$

- 1) Untransformed instruments set in combination with the initial weight matrix set up in (14) and (21) yields (5×5) the initial weight matrices for Blundell-Bond (1998) and suboptimal system GMM estimators, W_n^{BB} and W_n^{Ju} , respectively.
- 2) Collapsed instruments set in combination with initial weight matrices set up in (21) gives (4×4) the suboptimal initial weight matrix for estimator based on collapsed instruments, W_n^{Jc} .
- 3) Lagged limited instruments set in combination with initial weight matrices set up (21) yields the suboptimal initial weight matrix for estimator based on lag-limited instruments, W_n^{Jl} .

Under certain regularity conditions, the asymptotic variance of sample moments and varying initial weights matrix are derived as follows:

- I) Using untransformed instruments, the asymptotic covariance matrix for sample moments is given as:

$$S^U = \lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{i=1}^n Z_i^S' u_i u_i' Z_i^S \right) = E \left[\begin{pmatrix} y_{i1} & 0 & 0 & 0 \\ 0 & y_{i1} & 0 & 0 \\ 0 & y_{i2} & 0 & 0 \\ 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & \Delta y_{i3} \end{pmatrix} (u_i u_i') \begin{pmatrix} y_{i1} & 0 & 0 & 0 & 0 \\ 0 & y_{i1} & y_{i2} & 0 & 0 \\ 0 & 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & 0 & \Delta y_{i3} \end{pmatrix} \right]$$

$$= \sigma_\varepsilon^2 \begin{bmatrix} 2\sigma_y^2 & -\sigma_y^2 & -\delta & -(1-\gamma)\sigma_y^2 & 0 \\ -\sigma_y^2 & 2\sigma_y^2 & 2\delta & (1-\gamma)\sigma_y^2 & -\frac{\gamma\sigma_\varepsilon^2}{1+\gamma} \\ -\delta & 2\delta & 2\sigma_y^2 & -\frac{\sigma_\varepsilon^2}{1+\gamma} & -\frac{\sigma_\varepsilon^2}{1+\gamma} \\ -(1-\gamma)\sigma_y^2 & (1-\gamma)\sigma_y^2 & -\frac{\sigma_\varepsilon^2}{1+\gamma} & \frac{2(1+\rho)\sigma_\varepsilon^2}{1+\gamma} & \delta - \frac{\gamma\sigma_\varepsilon^2}{1+\gamma} \\ 0 & -\frac{\gamma\sigma_y^2}{1+\gamma} & -\frac{\sigma_\varepsilon^2}{1+\gamma} & \delta - \frac{\gamma\sigma_\varepsilon^2}{1+\gamma} & \frac{2(1+\rho)\sigma_\varepsilon^2}{1+\gamma} \end{bmatrix}$$

where $u_i = (\Delta\varepsilon_{i3}, \Delta\varepsilon_{i4}, \alpha_i + \varepsilon_{i3}, \alpha_i + \varepsilon_{i4})'$ is the vector of error terms in system

GMM estimation procedures and $\delta = E(y_{i,t-1}y_{it}) = \sigma_y^2 - \frac{\sigma_\varepsilon^2}{1+\gamma}$.

a) The initial weighting matrices, as in Blundell and Bond (1998), which is based on weight matrix as in (14) is computed as:

$$W_n^{BB} = \left[\lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{i=1}^n Z_i^S' Z_i^S \right) \right]^{-1} = E \left[\begin{pmatrix} y_{i1} & 0 & 0 & 0 \\ 0 & y_{i1} & 0 & 0 \\ 0 & y_{i2} & 0 & 0 \\ 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & \Delta y_{i3} \end{pmatrix} \begin{pmatrix} y_{i1} & 0 & 0 & 0 & 0 \\ 0 & y_{i1} & y_{i2} & 0 & 0 \\ 0 & 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & 0 & \Delta y_{i3} \end{pmatrix} \right]^{-1}$$

$$= \begin{bmatrix} \sigma_y^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_y^2 & \delta & 0 & 0 \\ 0 & \delta & \sigma_y^2 & 0 & 0 \\ 0 & 0 & 0 & \frac{2\sigma_\varepsilon^2}{1+\gamma} & 0 \\ 0 & 0 & 0 & 0 & \frac{2\sigma_\varepsilon^2}{1+\gamma} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{\sigma_y^2} & 0 & 0 & 0 & 0 \\ 0 & \frac{-\sigma_y^2}{\delta^2 - \sigma_y^4} & \frac{\delta}{\delta^2 - \sigma_y^4} & 0 & 0 \\ 0 & \frac{\delta}{\delta^2 - \sigma_y^4} & \frac{-\sigma_y^2}{\delta^2 - \sigma_y^4} & 0 & 0 \\ 0 & 0 & 0 & \frac{1+\gamma}{2\sigma_\varepsilon^2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1+\gamma}{2\sigma_\varepsilon^2} \end{bmatrix}$$

and,

- b) The initial weighting matrix in (21) for sub-optimally weighted system GMM estimator (using standard instruments set) is computed as:

$$\begin{aligned}
 W_n^{Ju} &= \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n Z_i^S{}' H_J Z_i^S \right)^{-1} = E \left[\left(\begin{bmatrix} y_{i1} & 0 & 0 & 0 \\ 0 & y_{i1} & 0 & 0 \\ 0 & y_{i2} & 0 & 0 \\ 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & \Delta y_{i3} \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 1+\rho & \rho \\ 0 & 0 & \rho & 1+\rho \end{bmatrix} \begin{bmatrix} y_{i1} & 0 & 0 & 0 & 0 \\ 0 & y_{i1} & y_{i2} & 0 & 0 \\ 0 & 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & 0 & \Delta y_{i3} \end{bmatrix} \right) \right]^{-1} \\
 &= \begin{bmatrix} 2\sigma_y^2 & -\sigma_y^2 & -\delta & 0 & 0 \\ -\sigma_y^2 & 2\sigma_y^2 & 2\delta & 0 & 0 \\ -\delta & 2\delta & 2\sigma_y^2 & 0 & 0 \\ 0 & 0 & 0 & 2(1-\gamma)(\rho+1)\sigma_y^2 & 2\rho\gamma(1-\gamma)\sigma_y^2 \\ 0 & 0 & 0 & 2\rho\gamma(1-\gamma)\sigma_y^2 & 2(1-\gamma)(1+\rho)\sigma_y^2 \end{bmatrix}^{-1} \\
 &= \begin{bmatrix} \frac{2}{3\sigma_y^2} & \frac{2(\sigma_y^4 - \delta^2)}{3\sigma_y^2(2\sigma_y^4 - \delta^2)} & \frac{\delta}{3(2\sigma_y^4 - \delta^2)} & 0 & 0 \\ \frac{1}{3\sigma_y^2} & \frac{4\sigma_y^4 - \delta^2}{3\sigma_y^2(2\sigma_y^4 - \delta^2)} & \frac{-\delta}{3(2\sigma_y^4 - \delta^2)} & 0 & 0 \\ 0 & \frac{-\delta}{2\sigma_y^4 - \delta^2} & \frac{\sigma_y^2}{2\sigma_y^4 - \delta^2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1+\rho}{2(1-\gamma)[(1+\rho)^2 - (\gamma\rho)^2]\sigma_y^2} & \frac{-\gamma\rho}{2(1-\gamma)[(1+\rho)^2 - (\gamma\rho)^2]\sigma_y^2} \\ 0 & 0 & 0 & \frac{-\gamma\rho}{2(1-\gamma)[(1+\rho)^2 - (\gamma\rho)^2]\sigma_y^2} & \frac{(1+\rho)}{2(1-\gamma)[(1+\rho)^2 - (\gamma\rho)^2]\sigma_y^2} \end{bmatrix}
 \end{aligned}$$

II) Using collapsed instruments, the asymptotic covariance matrix for sample moments is computed as:

$$S^C = \left(p \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (Z_i^S)^C (Z_i^S)^C' u_i u_i' \right) = E \left[\begin{pmatrix} y_{i1} & y_{i2} & 0 & 0 \\ 0 & y_{i1} & 0 & 0 \\ 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & \Delta y_{i3} \end{pmatrix} (u_i u_i') \begin{pmatrix} y_{i1} & 0 & 0 & 0 \\ y_{i2} & y_{i1} & 0 & 0 \\ 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & \Delta y_{i3} \end{pmatrix} \right]$$

$$= \begin{bmatrix} 4\sigma_y^2 - 2\delta & 2\delta - \sigma^2 & -\frac{2\sigma_\varepsilon^2}{1+\gamma} & -\frac{\sigma_\varepsilon^2}{1+\gamma} \\ 2\delta - \sigma_y^2 & 2\sigma_y^2 & -(1-\gamma)\sigma_y^2 & -\frac{\gamma\sigma_\varepsilon^2}{1+\gamma} \\ -\frac{2\sigma_\varepsilon^2}{1+\gamma} & -(1-\gamma)\sigma_y^2 & \frac{2(1+\rho)\sigma_\varepsilon^2}{1+\gamma} & \frac{\rho(2\gamma - \gamma^2 - 1)}{(1-\gamma^2)} \\ -\frac{\sigma_\varepsilon^2}{1+\gamma} & -\frac{\gamma\sigma_\varepsilon^2}{1+\gamma} & \frac{\rho(2\gamma - \gamma^2 - 1)}{(1-\gamma^2)} & \frac{2(1+\rho)\sigma_\varepsilon^2}{1+\gamma} \end{bmatrix}$$

c) The corresponding sub-optimal initial weighting matrix is computed as:

$$W_n^{JC} = \left(p \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (Z_i^S)^C H_J (Z_i^S)^C \right)^{-1} = E \left[\begin{pmatrix} y_{i1} & y_{i2} & 0 & 0 \\ 0 & y_{i1} & 0 & 0 \\ 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & \Delta y_{i3} \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 1+\rho & \rho \\ 0 & 0 & \rho & 1+\rho \end{pmatrix} \begin{pmatrix} y_{i1} & 0 & 0 & 0 \\ y_{i2} & y_{i1} & 0 & 0 \\ 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & \Delta y_{i3} \end{pmatrix} \right)^{-1}$$

$$= \begin{bmatrix} \frac{2\sigma_y^2}{(7\sigma_y^4 - 4\delta^2)} & \frac{\sigma_y^2 - 2\delta}{7\sigma_y^4 - 4\delta^2} & 0 & 0 \\ \frac{\sigma_y^2 - 2\delta}{7\sigma_y^4 - 4\delta^2} & \frac{4\sigma_y^4 - 2\delta}{7\sigma_y^4 - 4\delta^2} & 0 & 0 \\ 0 & 0 & \frac{2(1-\gamma)}{(1+\gamma)C} & -\frac{B}{(1+\gamma)(\rho+1)C} \\ 0 & 0 & -\frac{B}{(1+\gamma)(\rho+1)C} & \frac{(\rho+1)}{C} \end{bmatrix}$$

where $A = (1-\gamma)(\rho+1)$, $B = \rho(2\gamma - \gamma^2 - 1)$, and $C = (2A - B^2)(1-\gamma)\sigma_\varepsilon^2$.

III) Using lagged limited instruments set ($\tau = 1$), the asymptotic covariance matrix for sample moments is computed as:

$$S^L = \left(p \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (Z_i^S)^L u_i u_i' (Z_i^S)^L \right)^{-1} = E \left[\begin{pmatrix} y_{i1} & 0 & 0 & 0 \\ 0 & y_{i2} & 0 & 0 \\ 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & \Delta y_{i3} \end{pmatrix} (u_i u_i') \begin{pmatrix} y_{i1} & 0 & 0 & 0 \\ 0 & y_{i1} & 0 & 0 \\ 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & \Delta y_{i3} \end{pmatrix} \right]$$

$$= \begin{bmatrix} 2\sigma_y^2 & -\delta & -(1-\gamma)\sigma_y^2 & 0 \\ -\delta & 2\sigma_y^2 & \frac{-\sigma_\varepsilon^2}{1+\gamma} & -\frac{\sigma_\varepsilon^2}{1+\gamma} \\ -(1-\gamma)\sigma_\varepsilon^2 & \frac{-\sigma_\varepsilon^2}{1+\gamma} & \frac{2(1+\rho)\sigma_\varepsilon^2}{1+\gamma} & \frac{\rho(2\gamma-\gamma^2-1)}{(1-\gamma^2)} \\ 0 & -\frac{\sigma_\varepsilon^2}{1+\gamma} & \frac{\rho(2\gamma-\gamma^2-1)}{(1-\gamma^2)} & \frac{2(1+\rho)\sigma_\varepsilon^2}{1+\gamma} \end{bmatrix}$$

d) The corresponding suboptimal initial weighting matrix is computed as:

$$W_n^{Jl} = \left(p \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N H_{Sli}' J_l H_{Sli} \right)^{-1} = E \left[\begin{pmatrix} y_{i1} & 0 & 0 & 0 \\ 0 & y_{i2} & 0 & 0 \\ 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & \Delta y_{i3} \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 1+\rho & \rho \\ 0 & 0 & \rho & 1+\rho \end{pmatrix} \begin{pmatrix} y_{i1} & 0 & 0 & 0 \\ 0 & y_{i2} & 0 & 0 \\ 0 & 0 & \Delta y_{i2} & 0 \\ 0 & 0 & 0 & \Delta y_{i3} \end{pmatrix} \right]^{-1}$$

$$= \begin{bmatrix} \frac{2\sigma_y^2}{4\sigma_y^4 - \delta^2} & \frac{\delta}{4\sigma_y^4 - \delta^2} & 0 & 0 \\ \frac{\delta}{4\sigma_y^2 - \delta^2} & \frac{2\sigma_y^2}{4\sigma_y^2 - \delta^2} & 0 & 0 \\ 0 & 0 & \frac{(\rho+1)^2}{2(1-\gamma)[(\rho+1)^2 - (\gamma\rho)^2]\sigma_y^2} & \frac{-\gamma\rho}{2(1-\gamma)[(\rho+1)^2 - (\gamma\rho)^2]\sigma_y^2} \\ 0 & 0 & \frac{-\gamma\rho}{2(1-\gamma)[(\rho+1)^2 - (\gamma\rho)^2]\sigma_y^2} & \frac{(1+\rho)^2}{2(1-\gamma)[(\rho+1)^2 - (\gamma\rho)^2]\sigma_y^2} \end{bmatrix}$$