Model Estimation with AR(1) Endogenous Unobservable

1) Model:

The observables are y_{it} and x_{it}. The goal is to identify β in the relationship,

$$y_{it} = x_{it}'\beta + w_{it}'\gamma + \epsilon_{it} + \nu_{it}, \quad \nu_{it} \sim \text{iid}, \quad \nu_{it} \text{ is serially independent} \tag{1}$$
$$x_{it} = \gamma\epsilon_{it} + (1-\gamma)\, u_{it}, \quad u_{it} \sim \text{iid}, \quad u_{it} \text{ is serially dependent} \tag{2}$$
$$\epsilon_{it} = \rho\epsilon_{it-1} + \eta_{it}, \quad \eta_{it} \sim \text{iid}, \quad \eta_{it} \text{ is serially independent} \tag{3}$$

where w_{it} includes a constant.

2) Identification:

The usual argument for identification is based on the quasi-difference expression:

$$(y_{it} - \rho y_{it-1}) = (x_{it} - \rho x_{it-1})'\,\beta + (w_{it} - \rho w_{it-1})'\,\delta + \eta_{it} \tag{4}$$

Panel IV Approach: One approach is to rearrange Equation (4) as a panel regression:

$$y_{it} = y_{it-1}(\rho) + x_{it}'(\beta) + x_{it-1}'(-\rho\beta) + w_{it}'(\delta) + w_{it-1}'(-\rho\delta) + \eta_{it} \tag{5}$$

The only source of endogeneity in this regression is that x_{it} depends on η_{it}. This implies that β is identified by a regression of y_{it} on x_{it}, controlling for $(y_{it-1}, x_{it-1}, w_{it}, w_{it-1})$, and instrumented by $z_{it} = (x_{it-2})$ or $z_{it} = (x_{it-2}, y_{it-2})$ or $z_{it} = (x_{it-2}, y_{it-2}, w_{it-2})$.

GMM Approach: For any guess $\left(\hat{\beta}, \hat{\delta}, \hat{\rho}\right)$, we can define the guess of η_{it}:

$$\hat{\eta}_{it}\left(\hat{\beta}, \hat{\delta}, \hat{\rho}\right) \equiv (y_{it} - \hat{\rho} y_{it-1}) - (x_{it} - \hat{\rho} x_{it-1})'\,\hat{\beta} - (w_{it} - \rho w_{it-1})'\,\hat{\delta} \tag{6}$$

Then,

$$(\beta, \delta, \rho) \quad \text{solves} \quad \mathbb{E}\left[(w_{it}, z_{it})\;\hat{\eta}_{it}\left(\hat{\beta}, \hat{\delta}, \hat{\rho}\right)\right] = 0 \tag{7}$$

where z_{it} includes variables that are independent of η_{it}. The standard choices of instruments are $z_{it} = (x_{it-1}, y_{it-1})$ and $z_{it} = (x_{it-1}, x_{it-2})$. Note that the number of instruments z_{it} must include at least one more than the number of enodgenous variables in x_{it}. See the Appendix for further implementation details.

3) Simulation Exercise:

In order to compare the estimation approaches, Figure 1 simulates the model defined above. It sets $\beta = (0.5, -0.2)$, $\delta = 1$, $\rho = 0.5$, $\gamma = 0.5$, $\eta_{it} \sim \mathcal{N}(0,1)$, $u_{it} = u_{it-1} + \mathcal{N}(0,1)$, and $\nu_{it} = 0$. The length of the panel is $T = 3$, which is the minimum required. For various choices of N, it draws 10 random samples from the model, applies the estimators, and presents the box-plot of the distribution of estimates.

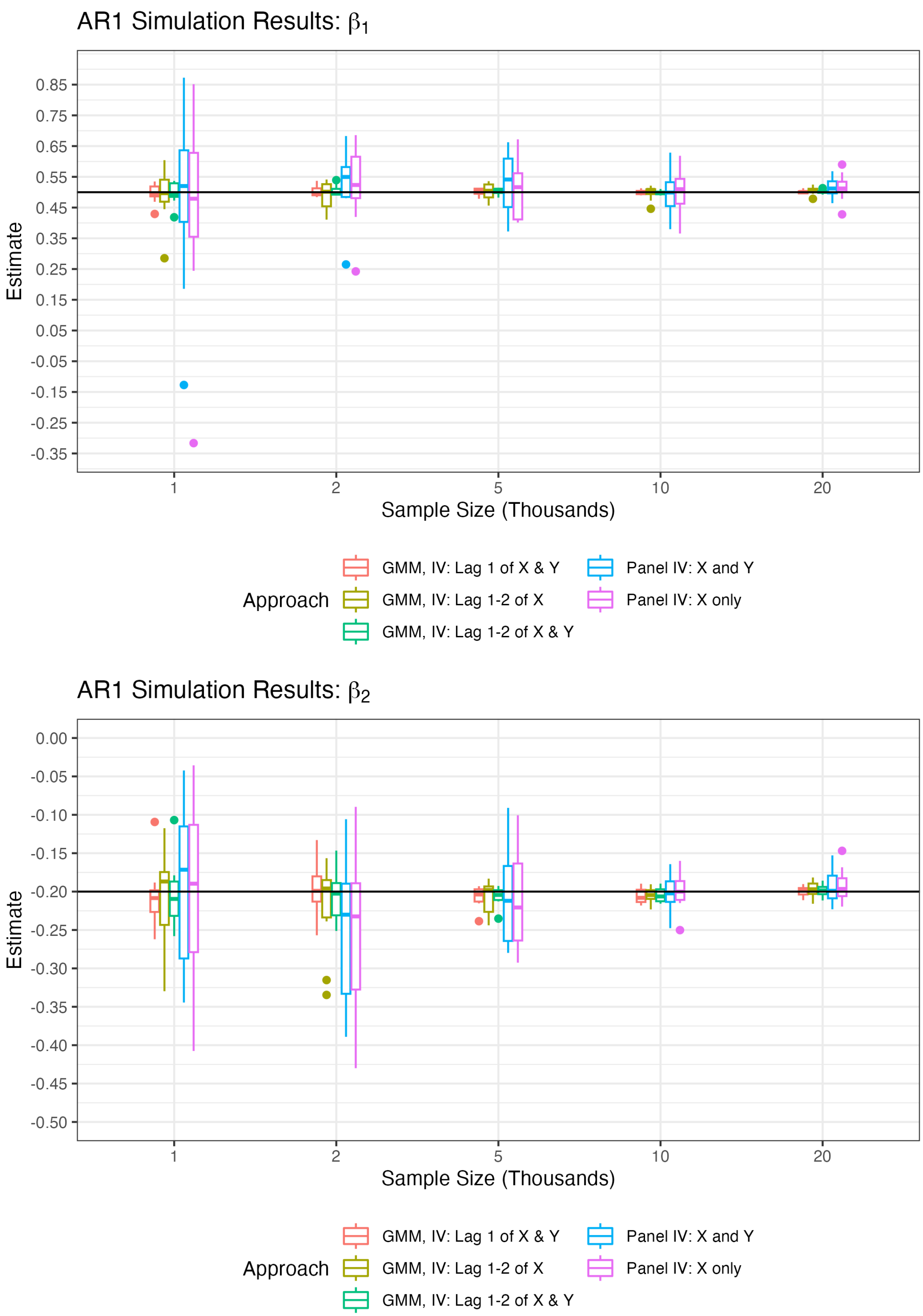

Figure 1: Simulation Exercise

Appendix: GMM Implementation Algorithm

Denote $g_{it}\left(\hat{\beta},\hat{\delta},\hat{\rho}\right)\equiv\left(w_{it},z_{it}\right)\ \hat{\eta}_{it}\left(\hat{\beta},\hat{\delta},\hat{\rho}\right)$.

1. Set the weighting matrix to $W=I$. Solve $(\beta,\delta)=\min_{\hat{\beta},\hat{\delta}}\mathbb{E}\left[g_{it}\left(\hat{\beta},\hat{\delta},0\right)\right]'W\mathbb{E}\left[g_{it}\left(\hat{\beta},\hat{\delta},0\right)\right]$. This gives an initial solution given $\rho=0$.
2. Solve $(\beta,\delta,\rho)=\min_{\hat{\beta},\hat{\delta},\hat{\rho}}\mathbb{E}\left[g_{it}\left(\hat{\beta},\hat{\delta},\hat{\rho}\right)\right]'W\mathbb{E}\left[g_{it}\left(\hat{\beta},\hat{\delta},\hat{\rho}\right)\right]$, taking the most recent solution as the initialization point.
3. Set the weighting matrix to $W=\hat{\Omega}^{-1}$, where $\hat{\Omega}=\mathbb{E}\left[g_{it}\left(\hat{\beta},\hat{\delta},\hat{\rho}\right)g_{it}\left(\hat{\beta},\hat{\delta},\hat{\rho}\right)'\right]$, evaluated at the most recent solution.
4. Repeats steps 2 then 3, until numerical convergence is achieved. The solutions are $(\beta,\delta,\rho),\Omega$.
5. Compute $G\equiv\mathbb{E}\left[\nabla g_{it}\left(\beta,\delta,\rho\right)\right]$, where G is a $(|w|+|z|)\times(|\beta|+|\delta|+|\rho|)$ dimensional matrix with elements,
$$\begin{aligned}\frac{\partial}{\partial\beta^{(j)}}g_{it}&=-\left(w_{it},z_{it}\right)\ \left(x_{it}^{(j)}-\rho x_{it-1}^{(j)}\right)\\ \frac{\partial}{\partial\rho}g_{it}&=-y_{it-1}+x_{it-1}'\hat{\beta}+w_{it-1}'\hat{\delta}\\ \frac{\partial}{\partial\delta^{(1)}}g_{it}&=-\left(w_{it},z_{it}\right)\ (1-\rho)\\ \frac{\partial}{\partial\delta^{(j)}}g_{it}&=-\left(w_{it},z_{it}\right)\ \left(w_{it}^{(j)}-\rho w_{it-1}^{(j)}\right),\ j>1\end{aligned}$$
6. Compute $\mathrm{SE}\left(\beta,\delta,\rho\right)=\mathrm{diag}\left(\mathrm{Var}\left(\beta,\delta,\rho\right)\right)^{1/2}$, where $\mathrm{Var}\left(\beta,\delta,\rho\right)=\left(G'\Omega^{-1}G\right)^{-1}/(NT)$.

Appendix: Faster GMM Implementation

Recall from equation (4):

$$y_{it}=y_{it-1}(\rho)+x_{it}'(\beta)+x_{it-1}'(-\rho\beta)+w_{it}'(\delta)+w_{it-1}'(-\rho\delta)+\eta_{it}$$

We can form a moment using the covariance between y_{it} and any instrument z_{it}:

$$\begin{aligned}0=&(1)\mathrm{Cov}\left(z_{it},y_{it}\right)+(-\rho)\mathrm{Cov}\left(z_{it},y_{it-1}\right)\\ &+\sum_{k=1,\ldots,|x_{it}|}(-\beta^{(k)})\mathrm{Cov}\left(z_{it},x_{it}^{(k)}\right)+\sum_{k=1,\ldots,|x_{it}|}(\rho\beta^{(k)})\mathrm{Cov}\left(z_{it},x_{it-1}^{(k)}\right)\\ &+\sum_{m=2,\ldots,|w_{it}|}(-\delta^{(m)})\mathrm{Cov}\left(z_{it},w_{it}^{(m)}\right)+\sum_{m=2,\ldots,|w_{it}|}(\rho\delta^{(m)})\mathrm{Cov}\left(z_{it},w_{it-1}^{(k)}\right)\end{aligned}$$

where δ_1 does not appear because it corresponds to the intercept. This is one equation in $|x_{it}| + |w_{it}|$ unknowns.

With $J \geq |x_{it}| + |w_{it}|$ instruments, we can solve this set of equations for ρ, β, δ (except for the intercept δ_1). The covariances only have to be computed once; they do not have to be updated for each parameter guess, which means we can solve for the parameters very quickly. We can then solve for δ_1 such that the unconditional expectation of y_{it} equals the model prediction from the right-hand side. Thus, we can provide a moment-matching estimate of all parameters very quickly, and use it to initialize the GMM estimator at a near-optimal solution.