Generalized method of moments estimation of linear dynamic panel data models

with an introduction to the `xtdpdgmm` Stata package

Sebastian Kripfganz

University of Exeter Business School, Department of Economics, Exeter, UK
www.kripfganz.de

Carpathian Stata Conference, 27th July 2026, Cluj-Napoca, Romania

University of Exeter Business School

Linear dynamic panel data model

- This talk considers panel data settings with N subjects and T time periods, where $T \ll N$ ("short-T panels" or "micro panels").
- While the presented methods also apply to static models, the focus will be on dynamic models with a lagged dependent variable:

$$y_{it} = \lambda y_{i,t-1} + \mathbf{x}'_{it}\beta + \underbrace{\alpha_i + \varepsilon_{it}}_{\text{error components}}$$

  where α_i is a time-invariant unobserved subject-specific effect and ε_{it} is an idiosyncratic error term.
- By construction, the lagged dependent variable $y_{i,t-1}$ is correlated with the unobserved effects α_i. Thus, ordinary least-squares estimation is inconsistent/biased.

Inconsistent estimation methods

- "Fixed-effects" estimation is still biased (when T is small; Nickell, 1981) because the deviations of $y_{i,t-1}$ from subject-specific averages over time $\bar{y}_{i,-1}$ are still correlated with the error term:

$$y_{it} - \bar{y}_i = \lambda(y_{i,t-1} - \bar{y}_{i,-1}) + (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'\beta + (\varepsilon_{it} - \bar{\varepsilon}_i)$$

- For similar reasons, the first-difference estimator is biased as well:

$$y_{it} - y_{i,t-1} = \lambda(y_{i,t-1} - y_{i,t-2}) + (\mathbf{x}_{it} - \mathbf{x}_{i,t-1})'\beta + (\varepsilon_{it} - \varepsilon_{i,t-1})$$

Consistent estimation methods

- If all regressors $\mathbf{x}_{it}$ are strictly exogenous – i.e., uncorrelated with the shocks ε_{is} for all time periods s and t – then consistent and efficient quasi-maximum likelihood (QML) and bias-corrected (BC) estimators can be constructed (Bhargava and Sargan, 1983; Hsiao, Pesaran, and Tahmiscioglu, 2002; Breitung, Kripfganz, and Hayakawa, 2022).
  - Stata packages `xtdpdqml` (Kripfganz, 2016) and `xtdpdbc` (Kripfganz and Breitung, 2022)
- In many applications, one wants to allow for the regressors $\mathbf{x}_{it}$ to be predetermined or endogenous – i.e., correlated with past (and contemporaneous) shocks (e.g., dynamic or simultaneous feedback from the outcome variable). In this case, generalized method of moments (GMM) estimators with a suitable choice of instrumental variables (IVs) are predominantly used.
  - Stata package `xtdpdgmm` (Kripfganz, 2019)

Instruments in first-differenced model

- Recall the first-differenced model:

$$\Delta y_{it} = \lambda \Delta y_{i,t-1} + \Delta \mathbf{x}'_{it}\beta + \Delta\varepsilon_{it}$$

- Under the assumption that ε_{it} is serially uncorrelated, lags of the dependent variable $y_{i,t-2}, y_{i,t-3}, \ldots$ are valid instruments for $\Delta y_{i,t-1}$ (Anderson and Hsiao, 1981; Arellano and Bond, 1991). Equivalently, the following moment conditions can be exploited:

$$E[y_{i,t-s}\Delta\varepsilon_{it}] = 0 \quad \text{for } s \geq 2$$

- Instead of using standard instruments, a much larger number of "GMM-type" instruments can be obtained by interacting them with a full set of time dummies.
  - Intuitively, this creates separate instruments for each time period, while standard instruments impose homogeneity of the first-stage coefficients across time periods.

(Too) many GMM-type instruments

- There are $\frac{(T-1)(T-2)}{2}$ of these GMM-type instruments. Their number thus increases quadratically in T, which can quickly lead to first-stage overfitting. Consequences include biased coefficient and standard error estimates, and weakened specification tests (Roodman, 2009).
- Remedies include "collapsing" into standard IVs and/or "curtailing" of the maximum lag order for the IVs (Kiviet, 2020).

Instruments for other regressors

- For the regressors $\mathbf{x}_{it}$, standard or GMM-type instruments are constructed similarly depending on their classification:
  - Strictly exogenous regressors satisfy $E[\mathbf{x}_{i,t-s}\Delta\varepsilon_{it}] = 0$ for all s (but in practice usually restricted to $s \geq 0$), leading to a large number of potential instruments.
  - Predetermined regressors satisfy $E[\mathbf{x}_{i,t-s}\Delta\varepsilon_{it}] = 0$ for $s \geq 1$. In other words, valid instruments are $\mathbf{x}_{i,t-1}, \mathbf{x}_{i,t-2}, \ldots$.
  - Endogenous regressors satisfy $E[\mathbf{x}_{i,t-s}\Delta\varepsilon_{it}] = 0$ for $s \geq 2$, implying one instrument less per variable (and time period) than for predetermined regressors. Valid instruments are $\mathbf{x}_{i,t-2}, \mathbf{x}_{i,t-3}, \ldots$.

Two-step, iterated, and continuously-updating GMM estimation

- Due to the large number of instruments, the model is (highly) overidentified. The two-stage least-squares (2SLS) estimator is generally inefficient. An efficient "two-step GMM" estimator can be obtained as follows:
  1. Estimate the model by 2SLS or GMM with an inefficient weighting matrix.
  2. Re-estimate the model by GMM with an "optimal" weighting matrix constructed from the residuals of the initial step.
- An asymptotically equivalent "iterated GMM" estimator iterates this procedure until convergence of the weighting matrix (Hansen, Heaton, and Yaron, 1996). This yields finite-sample robustness to the choice of the initial weighting matrix.
- Alternatively, the "continuously-updating GMM" estimator obtains the optimal weighting matrix simultaneously with the coefficient estimates (Hansen, Heaton, and Yaron, 1996). This can be computationally more demanding.
- Standard errors need to be corrected (in finite samples) to account for the first-step estimation error in the optimal weighting matrix (Windmeijer, 2005; Hwang, Kang, and Lee, 2022; Hansen and Lee, 2021).

Difference GMM estimation in Stata

- The so-called "difference GMM" estimator can be implemented in Stata with the `xtdpdgmm` command, here exemplarily with collapsing and curtailing (maximum lag order 4):

```
. xtdpdgmm y L.y x_exo x_pre x_endo, model(diff) gmmiv(L.n, lagrange(1 4))
> gmmiv(x_exo, lagrange(0 4)) gmmiv(x_pre, lagrange(1 4)) gmmiv(x_endo, lagrange(2 4))
> collapse teffects nolevel twostep vce(robust, dc)
```

- An equivalent syntax using the convenience tool `xtdpdgmmfe` is:

```
. xtdpdgmmfe y x_exo x_pre x_endo, lags(1) exogenous(x_exo) predetermined(x_pre)
> endogenous(x_endo) serial(0) curtail(4) collapse teffects nonl twostep vce(robust, dc)
```

- Iterated and continuously-updating GMM estimators are easily obtained by replacing option `twostep` with either `igmm` or `cugmm`.

Overidentification tests

- While the joint validity of all instruments is untestable, we can test the overidentifying restrictions implied by the additional instruments after assuming that as many instruments are valid as there are regressors (Hansen, 1982).
- A rejection might be indicative of various model misspecifications:
  - The classification of regressors into strictly exogenous, predetermined, or endogenous variables might be incorrect.
  - The dynamic structure of the model might be incomplete, causing the idiosyncratic errors ε_{it} to be serially correlated.
  - Relevant explanatory variables might be omitted.
- Overidentification tests with too many instruments might incorrectly signal validity of the overidentifying restrictions when in fact the model is misspecified.

Overidentification tests

- Instead of (or in addition to) testing all overidentifying restrictions jointly, we can test subsets with incremental overidentification tests (Eichenbaum, Hansen, and Singleton, 1988). Effectively, they compare two models with and without the suspicious instruments; a significant difference indicates invalidity, similar (and asymptotically equivalent) to a Hausman-type specification test.
- These incremental tests are especially useful to assess the regressor classification (Kiviet, 2020).

Serial-correlation tests

- A key assumption is absence of serial correlation in the idiosyncratic error component ε_{it}.
- Widely applied tests check for serial correlation (of order 2 or higher) in the first-differenced errors $\Delta\varepsilon_{it}$ (Arellano and Bond, 1991; Yamagata, 2008). Those tests have low power against some alternatives, especially if there is high autocorrelation near to a random walk.
- A recently proposed portmanteau test has superior power when T is very small (Jochmans, 2020), but suffers from a curse of dimensionality when T is growing (similar to the too-many-instruments problem). This can bite even for moderately small T when N is not very large.
- New tests based on a collapsing strategy and seasonal differencing (instead of first differencing) outperform the existing tests under most scenarios, and they retain power when T increases or the shocks are very persistent (Kripfganz, 2024; Kripfganz, Hosseinkouchack, and Demetrescu, 2026).

Identification failure

- Difference GMM estimators can suffer from weak instruments if the first-differenced regressors are poorly correlated with their past levels, resulting in biases and large standard errors.
- In the extreme case, identification might break down. While not routinely applied in the empirical practice yet, underidentification tests can shed light on this issue (Sanderson and Windmeijer, 2016; Windmeijer, 2024).
- In principle, inference can be made robust to weak identification. However, the results will often be uninformative due to very wide confidence intervals. The first-best solution would always be to find strong instruments.

Specification tests in Stata

- These specification tests can be carried out in Stata with dedicated post-estimation commands:

```
. xtdpdgmm y L.y x, model(diff) collapse gmm(L.n, lagrange(1 4))
> gmm(x, lagrange(1 4)) gmm(x, lagrange(0 0)) twostep vce(robust, dc) overid

. estat overid        // overidentification test

. estat overid, difference        // incremental overidentification tests

. underid, jgmm2s sw        // underidentification tests

. xtdpdserial        // serial-correlation tests
```

Nonlinear moment conditions

- Throughout, we maintained the assumption of serially uncorrelated idiosyncratic errors ε_{it}. This allows to exploit nonlinear moment conditions (Ahn and Schmidt, 1995):

$$E[(\alpha_i + \varepsilon_{iT})\Delta\varepsilon_{it}] = 0 \quad \text{for } t < T$$

- These moment conditions can improve identification and efficiency (at the only cost of slightly higher computational complexity) for the difference GMM estimator.
- Alternative nonlinear moment conditions can further improve efficiency under a homoskedasticity assumption (Ahn and Schmidt, 1995) or if none of the regressors $\mathbf{x}_{it}$ are classified as endogenous (Chudik and Pesaran, 2022).

Nonlinear GMM estimation in Stata

- The difference GMM estimator with additional nonlinear moment conditions can be implemented in Stata with the `xtdpdgmm` command and options `nl(noserial)`, `nl(iid)`, or `nl(predetermined)`. With `xtdpdgmmfe`, they are added by default depending on the specified assumptions:

```
. xtdpdgmm y L.y x_exo x_pre x_endo, model(diff) collapse gmmiv(L.n, lagrange(1 4))
> gmmiv(x_exo, lagrange(0 4)) gmmiv(x_pre, lagrange(1 4)) gmmiv(x_endo, lagrange(2 4))
> nl(noserial) twostep vce(robust, dc)

. xtdpdgmmfe y x_exo x_pre x_endo, lags(1) exogenous(x_exo) predetermined(x_pre)
> endogenous(x_endo) serial(0) curtail(4) collapse twostep vce(robust, dc)
```

Instruments for the untransformed level model

- A potential cause of weak identification can be a high persistence of the dependent variable (or a regressor). Additional instruments might re-establish identification strength.
- Such instruments can be found in the original model in levels:

$$y_{it} = \lambda y_{i,t-1} + \mathbf{x}_{it}'\beta + \alpha_i + \varepsilon_{it}$$

- Under the assumption of joint mean stationarity of y_{it} and $\mathbf{x}_{it}$, (lagged) first differences $\Delta y_{i,t-1}$ and $\Delta \mathbf{x}_{i,t(-1)}$ qualify as valid instruments, maintaining the assumption that ε_{it} is serially uncorrelated (Blundell and Bond, 1998).
- Mean stationarity implies that $\Delta y_{i,t-1}$ and $\Delta \mathbf{x}_{i,t(-1)}$ are uncorrelated with α_i, and therefore

$$E[\Delta y_{i,t-1}(\alpha_i + \varepsilon_{it})] = 0, \quad E[\Delta \mathbf{x}_{i,t(-1)}(\alpha_i + \varepsilon_{it})] = 0$$

Instruments for the untransformed level model

- Higher-order lags for these instruments – as well as nonlinear moment conditions – are usually not included because they are redundant when combined with all GMM-type instruments for the first-differenced model.
- For strictly exogenous or predetermined regressors, contemporaneous differences $\Delta \mathbf{x}_{it}$ qualify as instruments. For endogenous regressors, lagged differences $\Delta \mathbf{x}_{i,t-1}$ need to be used.
- The mean stationarity assumption is sometimes called an initial condition because it implies that the initial deviations of y_{it} from its subject-specific long-run equilibrium (steady state) should be unrelated to the long-run equilibrium itself.
- This requirement is insufficiently recognized in the empirical practice. In many (macroeconomic) applications, it is hard to justify.

System GMM estimation

- The combination of instruments for the first-differenced and the level model creates a "system" of regression equations. The resulting estimator is thus called "system GMM" estimator. To be asymptotically efficient, it again needs to be implemented as a two-step, iterated, or continuously-updating GMM estimator.
- In addition to the usual specification tests, the incremental overidentification test comparing the system GMM to the difference GMM estimator is especially useful to assess the validity of the mean stationarity assumption.

Deviations from within-subject means

- This system can be extended by adding instruments for other model transformations.
- To maximize the instrument strengths for strictly exogenous regressors, instruments $\mathbf{x}_{it}$ can be added for the model in deviations from within-group means (akin to a fixed-effects estimation):

$$\bar{\Delta} y_{it} = \lambda \bar{\Delta} y_{i,t-1} + \bar{\Delta} \mathbf{x}_{it}' \beta + \bar{\Delta} \varepsilon_{it}$$

where

$$\bar{\Delta} y_{it} = y_{it} - \frac{1}{T} \sum_{s=1}^{T} y_{is} \quad \text{etc.}$$

Forward-orthogonal deviations

- Instead of first differencing, the model can be transformed into deviations from their forward mean (Arellano and Bover, 1995):

$$\overrightarrow{\Delta} y_{it} = \lambda \overrightarrow{\Delta} y_{i,t-1} + \overrightarrow{\Delta} \mathbf{x}_{it}' \beta + \overrightarrow{\Delta} \varepsilon_{it}$$

where

$$\overrightarrow{\Delta} y_{it} = y_{it} - \frac{1}{T-t+1} \sum_{s=t}^{T} y_{is} \quad \text{etc.}$$

- For the lagged dependent variable, valid instruments are $y_{i,t-1}, y_{i,t-2}, \ldots$, and similarly for the regressors $\mathbf{x}_{it}$.
    - Note that the smallest lag is one less than with first differencing. However, the total number of instruments remains the same because the transformed model is not defined for period $t = T$.

Difference GMM estimation in Stata

- The system GMM estimator can be implemented in Stata with the `xtdpdgmm` command (or the `xtdpdgmmfe` convenience tool), here again with collapsing and curtailing:

```
. xtdpdgmm y L.y x_exo x_pre x_endo, model(diff) collapse gmmiv(L.n, lagrange(1 4))
> gmmiv(x_exo, lagrange(0 4)) gmmiv(x_pre, lagrange(1 4)) gmmiv(x_endo, lagrange(2 4))
> gmmiv(L.y x_exo x_pre L.x_endo, difference model(level) lagrange(0 0)) twostep vce(robust, dc)

. xtdpdgmmfe y x_exo x_pre x_endo, lags(1) exogenous(x_exo) predetermined(x_pre)
> endogenous(x_endo) serial(0) curtail(4) collapse twostep nonl stationary vce(robust, dc)
```

- Replacing first differences by forward-orthogonal deviations and deviations from within-subject means is achieved as follows:

```
. xtdpdgmm y L.y x_exo x_pre x_endo, model(fodev) collapse gmmiv(L.n, lagrange(0 3))
> gmmiv(x_exo, lagrange(0 3)) gmmiv(x_pre, lagrange(0 3)) gmmiv(x_endo, lagrange(1 3))
> gmmiv(x_exo, lagrange(0 0) model(mdev))
> gmmiv(L.y x_exo x_pre L.x_endo, difference model(level) lagrange(0 0)) twostep vce(robust, dc)

. xtdpdgmmfe y x_exo x_pre x_endo, lags(1) exogenous(x_exo) predetermined(x_pre)
> endogenous(x_endo) serial(0) curtail(4) collapse orthogonal twostep nonl stationary
> vce(robust, dc)
```

Concluding remarks

- GMM estimators offer a flexible approach to estimating dynamic panel data models with possibly predetermined or endogenous regressors.
- Consistency relies on possibly strong assumptions that should not be taken for granted (serially uncorrelated errors, mean stationarity for system GMM).
- Specification tests are routinely used to assess the reliability of the results.
- A lack of robustness to seemingly innocuous changes in the estimator's specification is a key challenge in the empirical practice.
- There are many interesting details and relevant aspects that I could not discuss today; see the Stata help files for a description of the available options.

References

- Ahn, S. C., and P. Schmidt (1995). Efficient estimation of models for dynamic panel data. *Journal of Econometrics 68* (1): 5–27.
- Anderson, T. W., and C. Hsiao (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association 76* (375), 598–606.
- 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* (2), 277–297.
- Arellano, M., and O. Bover (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics 68* (1), 29–51.
- Bhargava, A., and J. D. Sargan (1983). Estimating dynamic random effects models from panel data covering short time periods. *Econometrica 51* (6), 1635–1659.
- Blundell, R., and S. R. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics 87* (1), 115–143.
- Breitung, J., S. Kripfganz, and K. Hayakawa (2022). Bias-corrected method of moments estimators for dynamic panel data models. *Econometrics and Statistics 24*, 116–132.
- Chudik, A., and M. H. Pesaran (2022). An augmented Anderson-Hsiao estimator for dynamic short-T panels. *Econometric Reviews 41* (4), 416–447.
- Eichenbaum, M. S., L. P. Hansen, and K. J. Singleton (1988). A time series analysis of representative agent models of consumption and leisure choice under uncertainty. *Quarterly Journal of Economics 103* (1), 51–78.
- Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50* (4), 1029–1054.
- Hansen, L. P., J. Heaton, and A. Yaron (1996). Finite-sample properties of some alternative GMM estimators. *Journal of Business & Economic Statistics 14* (3), 262–280.
- Hansen, B. E., and S. Lee (2021). Inference for iterated GMM under misspecification. *Econometrica 89* (3), 1419–1447.
- Hsiao, C., M. H. Pesaran, and A. K. Tahmiscioglu (2002). Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. *Journal of Econometrics 109* (1), 107–150.

References

- Hwang, J., B. Kang, and S. Lee (2022). A doubly corrected robust variance estimator for linear GMM. *Journal of Econometrics 229* (2), 276–298.
- Jochmans, K. (2020). Testing for correlation in error-component models. *Journal of Applied Econometrics 35* (7), 860–878.
- Kiviet, J. F. (2020). Microeconometric dynamic panel data methods: Model specification and selection issues. *Econometrics and Statistics 13*, 16–45.
- Kripfganz, S. (2016). Quasi-maximum likelihood estimation of linear dynamic short-T panel-data models. *Stata Journal 16* (4), 1013–1038.
- Kripfganz, S. (2019). Generalized method of moments estimation of linear dynamic panel data models. *Proceedings of the 2019 London Stata Conference*.
- Kripfganz, S. (2024). Robust testing for serial correlation in linear panel-data models. *Proceedings of the 2024 London Stata Conference*.
- Kripfganz, S., and J. Breitung (2022). Bias-corrected estimation of linear dynamic panel data models. *Proceedings of the 2022 London Stata Conference*.
- Kripfganz, S., M. Hosseinkouchack, and M. Demetrescu (2026). Serial correlation testing in error component models with moderately small T. *Department of Economics Discussion Papers 26/03*, University of Exeter.
- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica 49* (6), 1417–1426.
- Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics 71* (1), 135–158.
- Sanderson, E., and F. Windmeijer (2016). A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics 190* (2), 212–221.
- Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics 126* (1), 25–51.
- Windmeijer, F. (2024). Testing underidentification in linear models, with applications to dynamic panel and asset pricing models. *Journal of Econometrics 240* (2), 105104.
- Yamagata, T. (2008). A joint serial correlation test for linear panel data models. *Journal of Econometrics 146* (1), 135–145.