Method and Formulas for `xtpraisk`: Prais–Winsten Regression with AR(k) Errors and Panel-Corrected Standard Errors

Ariel Linden
University of California, San Francisco
Department of Medicine
Division of Clinical Informatics & Digital Transformation (DoC-IT)
ariel.linden@ucsf.edu

1 Introduction

Panel (time-series cross-section) data present estimation challenges beyond those of a single time series. Errors may be serially correlated within each panel, heteroskedastic across panels, and contemporaneously correlated between panels at the same time point. Ignoring these complications leads to incorrect standard errors and invalid inference.

The standard approach for panel data with serially correlated errors is Prais–Winsten regression with panel-corrected standard errors (PCSEs), as implemented in Stata's official `xtpcse` command [Beck and Katz, 1995]. However, `xtpcse` accommodates only first-order autoregressive AR(1) errors. When panel time series exhibit higher-order serial dependence, researchers have lacked a panel-aware Prais–Winsten estimator for AR(k) errors.

The `xtpraisk` command fills this gap. It extends the higher-order Prais–Winsten estimator of `praisk` [Linden, 2026a,b, Vougas, 2021] to the panel setting, combining:

1. The exact Prais–Winsten GLS transformation for AR(k) errors described by Vougas [2021] and Galbraith and Galbraith [1974];

2. Panel-by-panel Yule–Walker estimation of the AR parameter vector $\mathbf{p}$, pooled across panels as a weighted mean matching the `correlation(ar1)` behavior of `xtpcse` [Beck and Katz, 1995]; and

3. Panel-corrected standard errors (PCSEs) for the regression coefficients, accounting for cross-panel contemporaneous correlation [Beck and Katz, 1995].

For AR(1), `xtpraisk` produces results identical to `xtpcse, correlation(ar1)`. For AR($k > 1$), it extends this framework to higher-order processes not covered by `xtpcse`.

The remainder of this document describes the statistical methods underlying `xtpraisk`. Section 2 presents the statistical model. Section 3 describes the single-pass estimation algorithm. Section 4 covers Yule–Walker estimation in the panel setting. Section 5 describes the Prais–Winsten GLS transformation. Section 6 covers the PCSE variance estimator. Section 7 discusses stationarity and the iteration display. Section 8 covers residual autocorrelation diagnostics. Section 9 provides guidance on when to use `praisk` versus `xtpraisk`. Section 10 concludes.

2 The Statistical Model

2.1 Overview

The model consists of a linear regression equation with autoregressive errors, fit to panel data with N units (panels) indexed $i = 1, \dots, N$ and T_i time periods per panel indexed $t = 1, \dots, T_i$.

The following notation is used throughout. k denotes the AR lag order. $\mathbf{p}$ denotes the $k \times 1$ vector of AR parameters $(p_1, \dots, p_k)'$; when $k = 1$ the single element is written as the scalar p. q denotes the number of regression coefficients including the constant. $\mathbf{L}$ denotes the $T_i \times T_i$ Prais–Winsten transformation matrix for panel i, and $\mathbf{L}_0$ its $k \times k$ initialization block.

2.2 Model

Let y_{it} denote the dependent variable for unit i at time t and $\mathbf{x}_{it}$ a $(q \times 1)$ vector of regressors. The model is:

$$y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + u_{it} \tag{1}$$

$$u_{it} = p_1 u_{i,t-1} + p_2 u_{i,t-2} + \cdots + p_k u_{i,t-k} + \varepsilon_{it} \tag{2}$$

where $\varepsilon_{it} \sim \text{iid}(0, \sigma_\varepsilon^2)$. The regression coefficients $\boldsymbol{\beta}$ and the AR parameter vector $\mathbf{p}$ are common to all panels (pooled). The regressors $\mathbf{x}_{it}$ are treated as strictly exogenous; the consistency of the Yule–Walker estimator of $\mathbf{p}$ relies on this assumption [Hamilton, 1994, Judge et al., 1985]. Although `xtpraisk` accepts time-series operators in *depvar* and *indepvars*, including lagged dependent variables as regressors violates strict exogeneity and the theoretical guarantees of the estimator no longer apply.

2.3 Panel Error Structure

Stacking observations within panel i as $\mathbf{y}_i$, $\mathbf{X}_i$, $\mathbf{u}_i$, and $\boldsymbol{\varepsilon}_i$, the model can be written panel by panel as:

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{u}_i, \quad i = 1, \ldots, N. \tag{3}$$

Following Beck and Katz [1995], the full $NT \times NT$ disturbance covariance matrix (for balanced panels) is assumed to be:

$$E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}'] = \boldsymbol{\Omega} = \boldsymbol{\Sigma}_{N\times N} \otimes \mathbf{I}_{T\times T} \tag{4}$$

where $\boldsymbol{\Sigma}$ is the $N \times N$ matrix of contemporaneous panel-by-panel covariances, with typical element $\sigma_{ij} = E[\epsilon_{it}\epsilon_{jt}]$.

The AR(k) serial correlation within each panel is eliminated first via the Prais–Winsten transformation; the PCSE sandwich then accounts for the remaining cross-panel covariance structure $\boldsymbol{\Sigma}$.

3 Single-Pass Estimation Algorithm

`xtpraisk` uses a single-pass estimation procedure matching the approach of `xtpcse` [Beck and Katz, 1995]:

1. **OLS.** Estimate $\boldsymbol{\beta}$ by OLS to obtain residuals $\hat{\mathbf{u}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{\text{OLS}}$.
2. **Rho estimation.** Estimate the AR parameter vector $\mathbf{p}$ from the OLS residuals using panel-by-panel Yule–Walker equations (Section 4).
3. **Prais–Winsten transformation.** Apply the exact Prais–Winsten GLS transformation using $\hat{\mathbf{p}}$ (Section 5) to produce transformed data $(\mathbf{y}_r, \mathbf{X}_r)$.
4. **GLS.** Compute GLS estimates:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}_r'\mathbf{X}_r)^{-1}\mathbf{X}_r'\mathbf{y}_r. \tag{5}$$

5. **PCSE.** Compute post-transformation residuals $\mathbf{e} = \mathbf{y}_r - \mathbf{X}_r\hat{\boldsymbol{\beta}}$ and estimate the PCSE variance-covariance matrix (Section 6).

All estimates are conditional on the estimated value of $\hat{\mathbf{p}}$; see `[TS] prais`.

This single-pass design matches `xtpcse`'s approach and avoids the overconfidence problems of full iterative FGLS documented by Beck and Katz [1995]. The iteration display shows rho at iteration 0 (= 0) and at iteration 1 (the Yule–Walker estimate from OLS residuals).

4 Panel-by-Panel Yule–Walker Estimation

4.1 Overview

The AR(k) parameter vector $\mathbf{p}$ is estimated from OLS residuals using a panel-by-panel approach that matches the `correlation(ar1)` behavior of `xtpcse` exactly for $k = 1$ and extends naturally to $k > 1$.

4.2 Within-Panel Yule–Walker System

For each panel i with $n_{s,i}$ usable observations and residuals $\hat{u}_{it}$, the $k \times k$ Yule–Walker normal equations $\mathbf{A}_i \mathbf{p}_i = \mathbf{b}_i$ are:

$$b_i[j] = \sum_{t=j+1}^{n_{s,i}} \hat{u}_{it}\hat{u}_{i,t-j}, \quad j = 1, \ldots, k \tag{6}$$

$$A_i[l, j] = \sum_{t=\max(l,j)+1}^{n_{s,i}} \hat{u}_{i,t-l}\hat{u}_{i,t-j}, \quad l, j = 1, \ldots, k. \tag{7}$$

The system is solved via LU decomposition to obtain the panel-specific estimate $\hat{\mathbf{p}}_i$.

4.3 Bounding

For AR(1), each panel-specific $\hat{p}_i$ is bounded to $[-1, 1]$ before pooling, matching `xtpcse`'s behavior. For AR($k > 1$), individual coefficients may legitimately exceed 1 in absolute value [Hamilton, 1994] and no element-wise bounding is applied.

4.4 Weighted Mean Across Panels

The common AR parameter vector is estimated as a weighted mean of the panel-specific estimates:

$$\hat{\mathbf{p}} = \frac{\sum_{i=1}^{N} w_i \hat{\mathbf{p}}_i}{\sum_{i=1}^{N} w_i} \tag{8}$$

where $w_i = n_{s,i} - 1$ by default (matching `xtpcse`'s default weighting by $T_i - 1$), or $w_i = n_{s,i}$ when the `np1` option is specified (matching `xtpcse`'s `np1` option). For balanced panels all weights are equal and the formula reduces to a simple arithmetic mean, identical to `xtpcse`'s $\hat{\rho} = (\hat{\rho}_1 + \cdots + \hat{\rho}_N)/N$.

Panels with $n_{s,i} \leq k$ cannot contribute a valid Yule–Walker estimate and are excluded from the mean.

4.5 Standard Errors for $\hat{\mathbf{p}}$

The asymptotic covariance of the Yule–Walker estimator is [Brockwell and Davis, 1991]:

$$\widehat{\mathrm{Var}}(\hat{\mathbf{p}}) = \hat{\sigma}_\varepsilon^2 \bar{\mathbf{A}}^{-1} \tag{9}$$

where $\bar{\mathbf{A}} = (\sum_i w_i \mathbf{A}_i)/\sum_i w_i$ is the weighted mean cross-product matrix and $\hat{\sigma}_\varepsilon^2$ is estimated from the AR residuals.

5 The Exact Prais–Winsten GLS Transformation

5.1 Overview

The Prais–Winsten transformation is applied independently to each panel (and to each contiguous segment within a panel if gaps are present). A *contiguous segment* is a maximal run of observations within the same panel with no gaps in the time index. The transformation is restarted at the beginning of every segment, so the first k observations of each panel receive the full Prais–Winsten initialization.

5.2 Cochrane–Orcutt Rows

For observations $t = k+1, \ldots, T_i$ within a segment, the standard AR(k) filter is applied:

$$\tilde{u}_{it} = u_{it} - p_1 u_{i,t-1} - \cdots - p_k u_{i,t-k}. \tag{10}$$

The Cochrane–Orcutt estimator [Cochrane and Orcutt, 1949] discards the first k observations of each segment. The Prais–Winsten estimator retains them by using the exact covariance structure.

5.3 Prais–Winsten Initialization

The covariance matrix of $(u_{i1}, \ldots, u_{ik})'$ is $\sigma_\varepsilon^2 V_k$, where V_k is the $k \times k$ autocovariance matrix of the AR(k) process. The top-left $k \times k$ block of the transformation matrix is set to the reversed Cholesky factor of V_k^{-1}:

$$\mathbf{L}_0 = \text{chol}(V_k^{-1})[k:1,\, k:1]. \tag{11}$$

The inverse autocovariance matrix V_k^{-1} is computed analytically using the closed-form expression of Galbraith and Galbraith [1974]. This approach is identical to that used in the Vougas [2021] MATLAB code and in Stata's official `prais` command for AR(1).

For AR(1), $V_k^{-1} = (1 - p^2)$ is a scalar, and the first transformed observation is $\tilde{y}_{i1} = \sqrt{1 - p^2}\, y_{i1}$, the classical Prais–Winsten transformation [Prais and Winsten, 1954].

5.4 Short Segments

A segment of length $n_s \leq k$ cannot contribute Cochrane–Orcutt rows. For such segments, only the Prais–Winsten initialization block applies: the transformation uses the top-left $n_s \times n_s$ submatrix of $\mathbf{L}_0$. For $n_s = 1$ and $k = 1$ this reduces to $\sqrt{1 - p^2} \cdot y$, identical to `prais`'s treatment of the first observation after a gap. This ensures `xtpraisk` handles gaps and unbalanced panels without error, matching the behavior of `prais` and `xtpcse`.

6 Panel-Corrected Standard Errors

6.1 Overview

Following Beck and Katz [1995], the default variance-covariance matrix of $\hat{\boldsymbol{\beta}}$ uses panel-corrected standard errors (PCSEs). After the Prais–Winsten transformation, the PCSE formula is:

$$\widehat{V}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}_r'\mathbf{X}_r)^{-1}\, \mathbf{X}_r'\hat{\boldsymbol{\Omega}}\mathbf{X}_r\, (\mathbf{X}_r'\mathbf{X}_r)^{-1} \tag{12}$$

where $\hat{\boldsymbol{\Omega}} = \hat{\boldsymbol{\Sigma}} \otimes \mathbf{I}_T$ and $\hat{\boldsymbol{\Sigma}}$ is the estimated contemporaneous panel covariance matrix.

6.2 Estimation of $\boldsymbol{\Sigma}$

The (i, j) element of $\boldsymbol{\Sigma}$ is estimated from the post-transformation residuals $\mathbf{e}_i$ and $\mathbf{e}_j$:

$$\hat{\sigma}_{ij} = \frac{\mathbf{e}_i'\mathbf{e}_j}{T_{ij}} \tag{13}$$

where T_{ij} is the number of time periods common to panels i and j. For balanced panels $T_{ij} = T$ for all i, j. This estimator mirrors `xtpcse`'s `getSigma` subroutine.

6.3 $X'\hat{\Omega}X$ Accumulation

The cross-product matrix $\mathbf{X}_r'\hat{\mathbf{\Omega}}\mathbf{X}_r$ is computed as:

$$\mathbf{X}_r'\hat{\mathbf{\Omega}}\mathbf{X}_r = \sum_{i=1}^{N}\sum_{j=1}^{N} \hat{\sigma}_{ij}\,\mathbf{X}_{r,i}'\,\mathbf{X}_{r,j} \tag{14}$$

where the summation is restricted to the T_{ij} matched time periods. This mirrors `xtpcse`'s `glsaccum` step.

6.4 Normalization

By default, standard errors are normalized by N (matching `xtpcse`). When the `nmk` option is specified, normalization is by $N - k$ (matching `praisk`), where k is the number of estimated coefficients.

6.5 Robust and Cluster-Robust Standard Errors

When `vce(robust)` is specified, the HC1 heteroskedasticity-consistent sandwich estimator is computed on the GLS-transformed data. When `vce(cluster` *clustvar*`)` is specified, the cluster-robust sandwich estimator [Liang and Zeger, 1986] is used. In both cases the PCSE cross-panel covariance structure is not used; these options are appropriate when the researcher wishes to account for heteroskedasticity or clustering rather than cross-panel correlation.

7 Stationarity and the Iteration Display

7.1 Stationarity Conditions

For AR(1), stationarity requires $|p| < 1$. For AR($k > 1$), stationarity requires all eigenvalues of the $k \times k$ companion matrix C to have modulus strictly less than 1 [Hamilton, 1994]:

$$C = \begin{bmatrix} p_1 & p_2 & \cdots & p_k \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}. \tag{15}$$

Individual coefficients p_j may exceed 1 in absolute value while the process remains stationary; this is expected behavior and not a cause for concern. No stationarity bounding is applied, matching `prais` and `xtpcse`.

7.2 Iteration Display

For AR(1), `xtpraisk` displays $\hat{p}$ at iteration 0 ($= 0$, the starting value) and iteration 1 (the Yule–Walker estimate from OLS residuals). For AR($k > 1$), the maximum eigenvalue modulus $\max_j |\lambda_j|$ of the companion matrix is displayed instead, since individual coefficients outside $[-1, 1]$ are routinely legitimate and would be misleading to display.

8 Residual Autocorrelation Diagnostics

After estimation, `xtpraisk` displays a table of residual autocorrelations at lags $1, \ldots, k$ for two series:

1. *Untransformed residuals*: OLS residuals $\hat{u}_{it} = y_{it} - \mathbf{x}'_{it}\hat{\boldsymbol{\beta}}_{\text{OLS}}$, computed before AR filtering. These establish the baseline level of serial correlation.

2. *Transformed (innovation) residuals*: GLS innovation residuals $e_{it} = y_{r,it} - \mathbf{x}'_{r,it}\hat{\boldsymbol{\beta}}$, computed after the Prais–Winsten transformation. Under correct AR(k) specification these should be approximately iid, so their autocorrelations should be near zero.

The lag-j sample autocorrelation for panel data is the pooled within-panel estimator:

$$\hat{\rho}_{\text{pooled}}(j) = \frac{\sum_{i=1}^{N} \sum_{t=j+1}^{T_i} (r_{it} - \bar{r})(r_{i,t-j} - \bar{r})}{\sum_{i=1}^{N} \sum_{t=1}^{T_i} (r_{it} - \bar{r})^2} \tag{16}$$

where cross-panel pairs never contribute, as the Stata lag operator returns missing for the first observation of each panel. This estimator is consistent for the common within-panel autocorrelation under the pooling assumption.

A successful AR(k) transformation is indicated by transformed autocorrelations close to zero at all displayed lags. For single-panel inspection, the innovation residuals may be saved with `predict ue, ue` and examined with `ac` applied to a single panel (e.g., `ac ue if panelvar == 1`), since `ac` does not support multiple panels.

9 When to Use `praisk` vs. `xtpraisk`

`praisk` and `xtpraisk` share the same AR(k) Prais–Winsten estimation core but differ in how they handle the panel structure and how they compute standard errors. The appropriate choice depends on the nature of the data and the error structure assumed.

9.1 Use `praisk` when:

- The data consist of a **single time series** (no panel structure). `praisk` requires only `tsset timevar` and is designed for this setting.

- The data have a **panel structure but the primary concern is serial correlation**, not cross-panel contemporaneous correlation. `praisk` handles panel data and gaps correctly, producing standard errors based on the GLS-transformed data that are valid when panels are independent.

- **Iterated FGLS** is desired. `praisk` alternates between Yule–Walker estimation and GLS regression until convergence, providing fully iterated estimates. `xtpraisk` uses a single pass.

- The analysis involves **model comparison via AIC/BIC**: `praisk` stores the exact Gaussian log likelihood (including the Prais–Winsten initialization correction), enabling valid `estat ic` comparisons across AR orders.

- **Robust or cluster-robust standard errors** are needed alongside the GLS transformation. Both commands support `vce(robust)` and `vce(cluster)`, but `praisk` applies these directly without the PCSE sandwich layer.

9.2 Use `xtpraisk` when:

- The data have a **panel structure with contemporaneous cross-panel correlation**. When errors for different panels at the same time point are correlated (e.g., countries in the same year), PCSEs account for this structure while OLS-based standard errors from `praisk` do not.

- The data exhibit **panel heteroskedasticity**. The PCSE sandwich in `xtpraisk` is robust to differing error variances across panels, as the $\hat{\Sigma}$ matrix captures both variances ($\hat{\sigma}_{ii}$) and covariances ($\hat{\sigma}_{ij}$).

- The researcher follows the **Beck and Katz (1995)recommendation**: retain OLS or Prais–Winsten parameter estimates but replace standard errors with PCSEs. `xtpraisk` implements exactly this strategy for AR(k) errors.

- The dataset is a **time-series cross-section (TSCS) dataset** typical in comparative political economy or macroeconomics (10–20 panels, 20–50 time periods). This is the setting for which Beck and Katz [1995] demonstrate PCSEs perform well.

9.3 Summary

Situation	`praisk`	`xtpraisk`
Single time series	✓	
Panel data, independent panels	✓	✓
Panel data, cross-panel correlation		✓
Panel heteroskedasticity		✓
Iterated FGLS	✓	
Single-pass (matching `xtpcse`)		✓
AIC/BIC model comparison	✓	
AR($k > 1$) panel data	✓	✓

10 Conclusions

`xtpraisk` extends the Prais–Winsten AR(k) estimator of `praisk` to the panel data setting, producing panel-corrected standard errors that account for contemporaneous cross-panel correlation. For AR(1), `xtpraisk` replicates `xtpcse` exactly; for AR($k > 1$), it fills a gap in the available panel time-series estimators.

The command follows the strategy recommended by Beck and Katz [1995]: Prais–Winsten parameter estimates (rather than full FGLS) combined with PCSEs for the variance-covariance matrix. Simulation evidence on the relative performance of this approach versus panel Newey–West methods such as `xtscc` [Driscoll and Kraay, 1998, Hoechle, 2007] under AR(k) panel errors remains an open question and is a natural direction for future work.

References

Nathaniel L. Beck and Jonathan N. Katz. What to do (and not to do) with time-series cross-section data. *American Political Science Review*, 89(3):634–647, 1995.

Peter J. Brockwell and Richard A. Davis. *Time Series: Theory and Methods*. Springer, New York, 2nd edition, 1991.

Donald Cochrane and Guy H. Orcutt. Application of least squares regression to relationships containing auto-correlated error terms. *Journal of the American Statistical Association*, 44(245):32–61, 1949.

John C. Driscoll and Aart C. Kraay. Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, 80(4):549–560, 1998.

Robin F. Galbraith and John I. Galbraith. On the inverses of some patterned matrices arising in the theory of stationary time series. *Journal of Applied Probability*, 11(1):63–71, 1974.

James D. Hamilton. *Time Series Analysis*. Princeton University Press, Princeton, NJ, 1994.

Daniel Hoechle. Robust standard errors for panel regressions with cross-sectional dependence. *Stata Journal*, 7(3):281–312, 2007.

George G. Judge, William E. Griffiths, R. Carter Hill, Helmut Lütkepohl, and Tsoung-Chao Lee. *The Theory and Practice of Econometrics*. Wiley, New York, 2nd edition, 1985.

Kung-Yee Liang and Scott L. Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22, 1986.

Ariel Linden. Adjustment for autocorrelation in multiple-group (controlled) interrupted time series analysis and its effect on power: A simulation study of the Newey-West and Prais-Winsten methods. *Preprint. Research Square*, 2026a. `https://doi.org/10.21203/rs.3.rs-8865851/v1`.

Ariel Linden. Multiple-group (controlled) interrupted time series analysis with higher-order autoregressive errors: A simulation study comparing Newey-West and Prais-Winsten methods. *Preprint. arXiv:2603.24814*, 2026b. https://arxiv.org/abs/2603.24814.

Sigbert J. Prais and Christopher B. Winsten. Trend estimators and serial correlation. Technical Report 383, Cowles Commission, Chicago, 1954.

Dimitrios V. Vougas. Prais-Winsten algorithm for regression with second or higher order autoregressive errors. *Econometrics*, 9(3):32, 2021.