Advanced Econometrics: SURE and FIML Methods

Prepared by

Aditya Korekallu Srinivasa

Scientist (Senior Scale)

ICAR-Indian Agricultural Research Institute
New Delhi

April 25, 2025

Contents

1 Seemingly Unrelated Regression Equations (SURE)

1.1 Motivation

Many empirical settings, especially in agricultural economics, involve modeling several related outcomes (e.g., crop yields, input demands, household expenditures) that may have different regressors but share unobserved shocks (e.g., weather, policy). SURE, introduced by Zellner (1962), efficiently estimates such systems by exploiting contemporaneous correlation among the errors.

1.2 The SURE System

Suppose we have G equations for n observations:

$$\begin{aligned} y_{1i} &= x'_{1i}\beta_1 + u_{1i} \\ y_{2i} &= x'_{2i}\beta_2 + u_{2i} \\ &\vdots \\ y_{Gi} &= x'_{Gi}\beta_G + u_{Gi} \end{aligned}$$

where u_{gi} are error terms, possibly correlated across g for each i.

1.3 Error Covariance Structure

Let $u_i = (u_{1i}, \ldots, u_{Gi})'$. Assume:

$$E[u_i] = 0, \quad E[u_i u_i'] = \Omega, \quad E[u_i u_j'] = 0 \ (i \neq j)$$

where Ω is a $G \times G$ positive definite matrix.

1.4 When to Use SURE

- When you have multiple equations with different regressors.
- When errors are correlated across equations for each observation.
- When there is no endogeneity (all regressors are exogenous).

1.5 OLS vs. SURE

- If errors are uncorrelated across equations, OLS applied to each equation is efficient. - If errors are correlated, SURE (GLS) is more efficient than OLS.

1.6 SURE Estimator (GLS Formulation)

Stack the system for all n observations:

$$Y = X\beta + u$$

where Y is $nG \times 1$, X is block-diagonal, β is stacked, u is $nG \times 1$.

The GLS estimator is:

$$\hat{\beta}_{SURE} = (X'(\Omega^{-1} \otimes I_n)X)^{-1} X'(\Omega^{-1} \otimes I_n)Y$$

where $\otimes$ is the Kronecker product.

1.7 Estimation Steps

1. Estimate each equation by OLS, obtain residuals.
2. Estimate Ω using the residuals.
3. Use $\hat{\Omega}$ in the GLS formula above to obtain the SURE estimator.

1.8 Agricultural Example: Joint Input Demand

Suppose a farm household chooses fertilizer and pesticide use:

$$\begin{aligned}
\text{Fert}_i &= \alpha_0 + \alpha_1 \text{PriceF}_i + \alpha_2 \text{Income}_i + u_{1i} \\
\text{Pest}_i &= \beta_0 + \beta_1 \text{PriceP}_i + \beta_2 \text{Income}_i + u_{2i}
\end{aligned}$$

Unobserved factors (e.g., risk preferences, weather) may affect both u_{1i} and u_{2i}, leading to correlated errors. SURE allows efficient estimation and joint hypothesis testing (e.g., whether income effects are equal).

1.9 Software

- **R:** `systemfit` package, method="SUR".
- **Stata:** `suest` or `reg3, sure`.
- **Python:** `linearmodels.system.SUR`.

1.10 Advantages and Limitations

- **Advantages:** More efficient than OLS if errors are correlated; allows joint testing.
- **Limitations:** Cannot handle endogenous regressors; requires correct specification of error covariance.

2 Full Information Maximum Likelihood (FIML)

2.1 Motivation

FIML is a system estimation method for simultaneous equation models (SEMs) that exploits all available information in the system. It is asymptotically the most efficient estimator under correct specification.

2.2 The FIML Approach

Consider a system of G simultaneous equations:

$$y_{gi} = x'_{gi}\beta_g + \sum_{h \neq g} \gamma_{gh} y_{hi} + u_{gi}$$

for $g = 1, \ldots, G$ and $i = 1, \ldots, n$.

Define the structural form as:

$$BY_i = \Gamma X_i + u_i$$

where B is a $G \times G$ matrix of coefficients on endogenous variables, Γ is $G \times K$, X_i is $K \times 1$.

2.3 Likelihood Function

Assume $u_i \sim N(0, \Omega)$, Ω known up to parameters.

The joint likelihood for all n observations is:

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{G/2}|\Omega|^{1/2}} \exp\left\{-\frac{1}{2}u_i'\Omega^{-1}u_i\right\}$$

where $u_i = BY_i - \Gamma X_i$.

2.4 FIML Estimation Steps

1. Write the log-likelihood function in terms of the structural parameters (B, Γ, Ω).
2. Maximize the log-likelihood jointly with respect to all structural parameters.
3. Use numerical optimization (e.g., Newton-Raphson) to obtain estimates.

2.5 Properties

- FIML is consistent and asymptotically efficient if the system is correctly specified.
- More efficient than 3SLS if all equations are correctly specified.
- Sensitive to misspecification in any equation.

2.6 Comparison with Other Methods

- **2SLS**: Single equation, less efficient.
- **3SLS**: System method, efficient under correct Ω.
- **FIML**: Most efficient, but requires correct specification of all equations and distributional assumptions.

2.7 Agricultural Example: Output Supply and Input Demand

Suppose a system for a farm household:

$$\text{Output}_i = \alpha_0 + \alpha_1 \text{Fertilizer}_i + \alpha_2 \text{Labor}_i + u_{1i}$$
$$\text{Fertilizer}_i = \beta_0 + \beta_1 \text{Output}_i + \beta_2 \text{PriceF}_i + u_{2i}$$

Both equations are jointly determined. FIML estimates all parameters jointly, using the full system likelihood.

2.8 Software

- **R:** `systemfit` package, method="FIML".
- **Stata:** `reg3, fiml`.

2.9 Advantages and Limitations

- **Advantages:** Asymptotically most efficient; allows full system hypothesis testing.
- **Limitations:** Sensitive to misspecification; computationally intensive; requires normality.

3 Comparison Table: SURE, 2SLS, 3SLS, FIML

Method	Endogeneity	Error Correlation	System Estimation	Distributional Assump
SURE	No	Yes	Yes	None
2SLS	Yes	No	No	None
3SLS	Yes	Yes	Yes	None
FIML	Yes	Yes	Yes	Normality

4 Agricultural Economics Papers Using SURE and FIML

4.1 SURE Example 1: Input Demand Systems

Source: Abdulai, A., Huffman, W. E. (2000). "Structural adjustment and economic efficiency of rice farmers in Northern Ghana." *Economic Development and Cultural Change*, 48(3), 503–520.

Application: Joint estimation of input demand equations (fertilizer, labor, land) using SURE to account for correlated errors due to unobserved farm-specific shocks.

4.2 SURE Example 2: Household Expenditure

Source: Deaton, A. (1988). "Quality, quantity, and spatial variation of price." *American Economic Review*, 78(3), 418–430.

Application: SURE used to jointly estimate equations for different food expenditure categories, allowing for correlated errors across equations.

4.3 FIML Example: Supply Response and Acreage Allocation

Source: Askari, H., Cummings, J. T. (1977). "Estimating agricultural supply response with the Nerlove model: A survey." *International Economic Review*, 18(2), 257–292.

Application: FIML used to estimate a system of supply response and acreage allocation equations, exploiting full information for efficiency.

4.4 FIML Example 2: Agricultural Household Models

Source: Taylor, J. E., Adelman, I. (2003). "Agricultural household models: Genesis, evolution, and extensions." *Review of Economics of the Household*, 1(1-2), 33–58.

Application: FIML used to estimate simultaneous equations for production, consumption, and labor allocation in agricultural households.

5 Practice Questions

1. Explain the SURE model and its efficiency advantage over OLS.
2. Describe the steps in estimating a SURE system.
3. What are the main differences between SURE and 3SLS?
4. Outline the FIML estimation procedure for a simultaneous equations model.
5. Compare the efficiency and assumptions of FIML and 3SLS.
6. Discuss an agricultural application where SURE or FIML would be preferred to single-equation estimation.
7. What are the limitations of SURE and FIML in empirical research?

Glossary

- **SURE**: Seemingly Unrelated Regression Equations, a system estimation method for equations with correlated errors.
- **FIML**: Full Information Maximum Likelihood, a system estimation method for simultaneous equations using the full likelihood.
- **GLS**: Generalized Least Squares, efficient estimation in the presence of correlated errors.
- **System Estimation**: Joint estimation of multiple equations.
- **Endogeneity**: Correlation between regressor and error term.
- **Efficiency**: Achieving the smallest possible variance of estimators.
- **Identification**: Ability to uniquely estimate model parameters.

References

- Gujarati, D. N., & Porter, D. C. (2010). *Basic Econometrics* (5th ed.). McGraw-Hill.
- Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach* (5th ed.). Cengage Learning.

- Zellner, A. (1962). "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias." *Journal of the American Statistical Association*, 57(298), 348–368.

- Abdulai, A., Huffman, W. E. (2000). "Structural adjustment and economic efficiency of rice farmers in Northern Ghana." *Economic Development and Cultural Change*, 48(3), 503–520.

- Deaton, A. (1988). "Quality, quantity, and spatial variation of price." *American Economic Review*, 78(3), 418–430.

- Askari, H., Cummings, J. T. (1977). "Estimating agricultural supply response with the Nerlove model: A survey." *International Economic Review*, 18(2), 257–292.

- Taylor, J. E., Adelman, I. (2003). "Agricultural household models: Genesis, evolution, and extensions." *Review of Economics of the Household*, 1(1-2), 33–58.