# systemfit
## Simultaneous Equation Systems in R

Arne Henningsen (University of Kiel, Germany)
Jeff D. Hamann (Forest Informatics, Inc., Corvallis, USA)

useR!, Vienna, June 16, 2006



## Motivation

Many theoretical models consist of more than one equation

- contemporaneous correlation of disturbance terms (likely)
- simultaneous estimation of all equations as "Seemingly Unrelated Regression" (SUR) leads to efficient results

Theoretically derived cross-equation parameter restrictions

- simultaneous estimation of all equations required

Endogeneity of some variables

- estimation using "Two-Stage Least Squares" (2SLS) or "Three-Stage Least Squares" (3SLS) required

$\Rightarrow$ All these models can be estimated by `systemfit`



## Outline

- Introduction
- Features of systemfit
- Example
- Plans for the Future



## Estimation Methods

- Ordinary Least Squares (OLS)
- Two-Stage Least Squares (2SLS)
- Seemingly Unrelated Regression (SUR)
- Three-Stage Least Squares (3SLS)
- ...

## Estimation Control

- imposition of linear restrictions
- instrumental variables
- iteration of FGLS estimation
- formulas for the residual covariance matrix
- formulas for 3SLS estimation
- ...



## Other Tools

- `systemfitClassic`: wrapper function for (classical) panel-like data in long format
- testing linear hypotheses using the F-, Wald-, and LR-statistic
- Hausman test for the consistency of the 3SLS estimator



## Example: Commands

- from Kmenta (1986): Elements of Econometrics, p. 685
- specification of the equation system:

```
eqDemand <- consump ~ price + income
eqSupply <- consump ~ price + farmPrice + trend
eqSystem <- list(demand=eqDemand, supply=eqSupply)
```

- estimation using method "SUR":

```
fitsur <- systemfit("SUR", eqSystem, data=Kmenta)
```

- printing summary results:

```
summary( fitsur )
```



## Results of the Entire System

```
systemfit results
method: SUR

         N DF      SSR      MSE    RMSE       R2    Adj R2
demand 20 17  65.6829 3.86370 1.96563 0.755019 0.726198
supply 20 16 104.0584 6.50365 2.55023 0.611888 0.539117

[...]

The correlations of the residuals
          demand   supply
demand 1.000000 0.982348
supply 0.982348 1.000000

The determinant of the residual covariance matrix: 0.879285
OLS R-squared value of the system: 0.683453
McElroy's R-squared value for the system: 0.788722
```

## Results of a Single Equation

```
SUR estimates for 'demand' (equation 1)
Model Formula: consump ~ price + income

             Estimate Std. Error   t value Pr(>|t|)
(Intercept) 99.332894   7.514452 13.218913        0 ***
price       -0.275486   0.088509 -3.112513 0.006332  **
income        0.29855   0.041945  7.117605    2e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '

Residual standard error: 1.96563 on 17 degrees of freedom
Number of observations: 20 Degrees of Freedom: 17
SSR: 65.682902 MSE: 3.8637 Root MSE: 1.96563
Multiple R-Squared: 0.755019 Adjusted R-Squared: 0.726198
```



## Plans for the Future

- estimation with unbalanced data sets
- estimation methods: LIML, FIML, and GMM
- fitting equation systems with serially correlated and heteroscedastic disturbances
- spatial econometric methods
- simplify specification of parameter restrictions
- improving the function `nlsystemfit` to estimate systems of non-linear equations
- ...



## User Interface: Arguments

Arguments of `systemfit`:

- method
- eqns
- eqnlabels
- inst
- data
- R.restr
- q.restr
- TX
- maxiter
- tol
- rcovformula
- centerResiduals
- formula3sls
- probdfsys
- single.eq.sigma
- solvetol
- saveMemory
- (more in the future)

**Too many?**



## Arguments

Reducing arguments?

- method
- eqns
- inst
- data
- R.restr
- q.restr
- TX
- **control** (like in `optim`)

However: This would break existing code!

# The End . . .

*Thank you for your attention!*