Simultaneous Equations in Young Stands

Was it worth it?

Jeff D. Hamann

jeff_hamann@hamanndonald.com

The Topics

- Simultaneous equation background
- Development of young tree model
- Model results and performance
- R and systemfit
- Conclusions

Individual Tree Growth Model Types

- Establishment Models
  - SYSTEM-1 (Ritchie and Powers, 1993)
  - CONIFERS (Ritchie, 2001)
  - RVMM (Opalach et al., 1990)
- Established Models
  - ORGANON (Hann et al., 1993)
  - FVS (Wykoff et al., 1982)
  - CACTOS/CRPTOS (Wensel et al., 1986)

Traditional OLS Approach

- Each Model Consists of Several Equations
  - Height Growth
  - Diameter Growth
  - Crown Recession
  - Mortality
- Equations are fit individually (different datasets?)
- Equations are applied individually

OLS Formulas

$$\hat{\beta}_{OLS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

$$\hat{\beta}_{OLS} = (\mathbf{X}'(\hat{\mathbf{\Omega}}^{-1} \bigotimes \mathbf{I})\mathbf{X})^{-1}(\mathbf{X}'(\hat{\mathbf{\Omega}}^{-1} \bigotimes \mathbf{I})\mathbf{Y})$$

$$\text{Asympt. Var-Cov}(\hat{\beta}_{OLS}) = (\mathbf{X}'\hat{(}\mathbf{\Omega}^{-1} \bigotimes \mathbf{I})\mathbf{X})^{-1}$$

$$\mathbf{\Omega} = \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} & \dots & \hat{\sigma}_{1M} \\ \hat{\sigma}_{21} & \hat{\sigma}_{22} & \dots & \hat{\sigma}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}_{M1} & \hat{\sigma}_{M2} & \dots & \hat{\sigma}_{MM} \end{bmatrix}$$

$$\hat{\sigma}_{ij} = \begin{cases} 0 & i \neq j \\ \frac{\epsilon_i'\epsilon_j}{n-p} & \text{otherwise} \end{cases}$$

Assumptions of OLS Approach

- Residuals are $N(0, s^2)$
- Residuals are Not Serially Correlated
- Independent Observations
- Independent Variables are Measured Without Error

These can be problematic in multiple equation models.

Solutions to OLS Deficiences

- Single Equation Methods
  - Ignore Deficiencies
  - Two-Stage Least Squares
- Simultaneous Equation Methods
  - Seemingly Unrelated Regressions
  - Three-Stage Least Squares

Two-Stage Least Squares (2SLS)

- Removes simultaneity bias from equation [a]
- Algorithm
  - Regress all instrumental variables on dependent variables
  - Substitute the regressed observations in data and refit equation

[a] Thiel, 1953

Seemingly-Unrelated Regressions

- Corrects for contemporaneous correlation [a]
- Algorithm
  - Compute the OLS estimates
  - Build a variance-covariance matrix for residuals
  - Stack the equations and fit

---

[a] Zellner, 1963

Three-Stage Least Squares (3SLS)

- Combine 2SLS and SUR [a]
- Algorithm
  - Perform 2SLS
  - Perform SUR on 2SLS results

[a] Zellner and Thiel, 1962

Parameter Estimates

$$
\begin{aligned}
\hat{\beta}_{OLS} &= \left[\mathbf{X}'\left(\hat{\mathbf{\Omega}}^{-1}\bigotimes\mathbf{I}\right)\mathbf{X}\right]^{-1}\left[\mathbf{X}'\left(\hat{\mathbf{\Omega}}^{-1}\bigotimes\mathbf{I}\right)\mathbf{Y}\right] \\
\hat{\beta}_{j,2SLS} &= \left[\mathbf{X_j'}\left[\mathbf{Z}(\mathbf{Z}'\mathbf{Z})\right]^{-\mathbf{1}}\mathbf{X_j}\right]\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-\mathbf{1}}\mathbf{Z}'y_j \\
\hat{\beta}_{SUR} &= \left[\mathbf{X}'\left(\hat{\mathbf{\Omega}}^{-1}\bigotimes\mathbf{I}\right)\mathbf{X}\right]^{-1}\left[\mathbf{X}'\left(\hat{\mathbf{\Omega}}^{-1}\bigotimes\mathbf{I}\right)\mathbf{Y}\right] \\
\hat{\beta}_{3SLS} &= \left[\hat{\mathbf{Z}}'\left(\mathbf{\Omega}^{-1}\bigotimes\mathbf{I}\right)\hat{\mathbf{Z}}\right]^{-\mathbf{1}}\left[\hat{\mathbf{Z}}'\left(\mathbf{\Omega}^{-1}\bigotimes\mathbf{I}\right)\mathbf{Y}\right]
\end{aligned}
$$

where **Z** is a matrix of all the instrumental variables

Variance-Covariance Matrix

$$\text{Asympt. Var-Cov}(\hat{\beta}_{SUR}) = \left[\mathbf{X}'\left(\hat{\mathbf{\Omega}}^{-\mathbf{1}}\bigotimes\mathbf{I}\right)\mathbf{X}\right]^{-1}\left[\mathbf{X}'\left(\hat{\mathbf{\Omega}}^{-\mathbf{1}}\bigotimes\mathbf{I}\right)\mathbf{Y}\right]$$

where

$$\hat{\Omega} = \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} & \cdots & \hat{\sigma}_{1M} \\ \hat{\sigma}_{21} & \hat{\sigma}_{22} & \cdots & \hat{\sigma}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}_{M1} & \hat{\sigma}_{M2} & \cdots & \hat{\sigma}_{MM} \end{bmatrix}$$

and

$$\hat{\sigma}_{ij} = \frac{\epsilon_i'\epsilon_j}{[(N-K_i)(N-K_j)]^{1/2}}$$

Previous Applications

- 2SLS $\rightarrow$ Furnival & Wilson (1971) - stand level
- SUR $\rightarrow$ Rose and Lynch (2001) - basal area growth
- 3SLS $\rightarrow$ Hasenauer et al. (1998) - nonlinear

This work was applied to young stand models.

Plots and Data

- 109 Plots Total
  - 81 from USFS
  - 16 from BLM
  - 7 Fruit Growers
  - 5 Roseburg Forest Products
- 51 California/58 Oregon
- 428 Douglas-fir Records

Data Ranges

Variable	Min	Mean	Max
Basal Diameter, in	0.10	2.04	17.90
Height, ft.	2.01	9.14	53.40
Crown Ratio	0.18	0.81	0.98
WHC, in	2.78	6.59	11.25
Elevation, ft.	1800	3377	6200

The Model

$$
\begin{aligned}
\ln(BINC) &= a_0 + a_1 \ln(D6) + a_2 \ln(HINC) + a_3 BAT \\
\ln(HINC) &= b_0 + b_1 \ln(HT) + b_2 HT^2 + b_3 WHC + b_4 C \\
C &= c_0 + c_1 \ln(D6) + c_2 HT^2 + c_3 BAT
\end{aligned}
$$

where,

$$C = \ln\left(\frac{1}{CR} - 1\right)$$

Projection Results

- Basal Diameter Increment
- Total Height Increment
- Crown Ratio
- Assumptions
  - Single Tree Record
  - Initial Size (BD=0.1, THT=1.0)
  - Dominant Tree (BAIN = 0.0)
  - Medium Productivity for Data (WHC = 7.0)
  - No Mortality
- Corrected for Log-Bias using Baskerville Correction

Predicted Relationships...

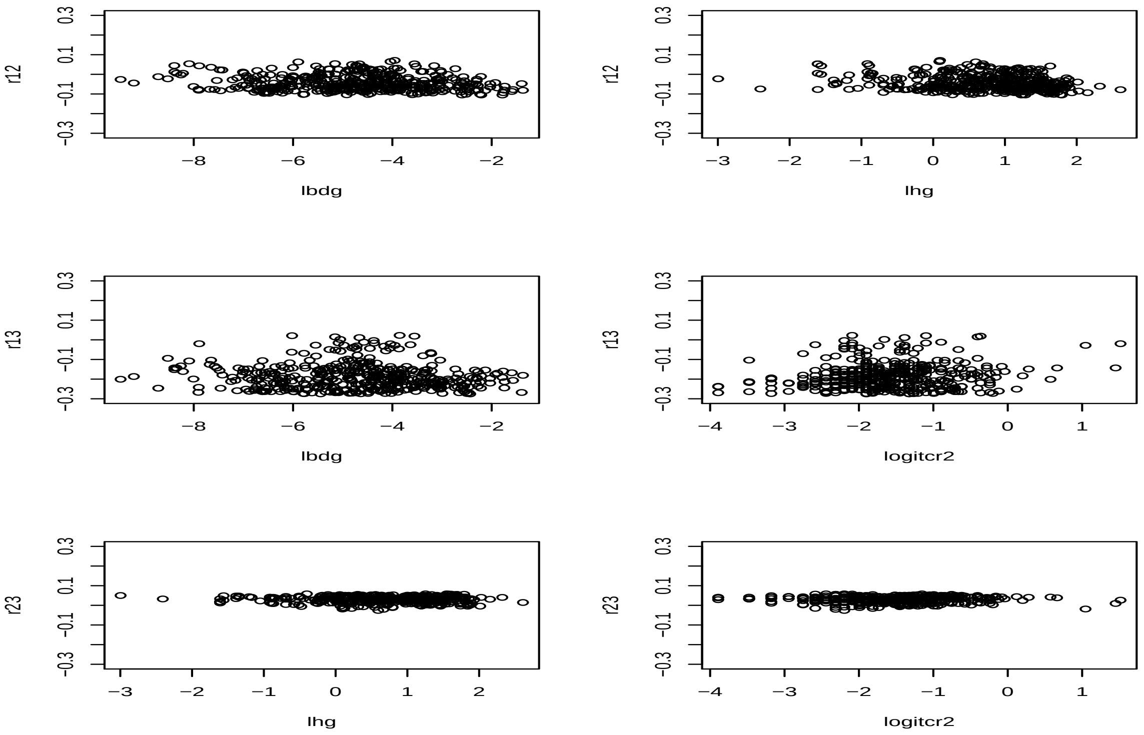
- Are the predictions correlated?
- Is there an improvement in prediction?
- Examine the standard errors of the predictions.

Cross Equation Correlation

$$r_{ijk} = \frac{x'_{ik} C_{ij} x_{jk}}{\sqrt{(x'_{ik} C_{ii} x_{ik})(x'_{jk} C_{jj} x_{jk})}}$$

where r_{ijk} is the correlation between the predicted values of equations i and j for the k^{th} observation. C_{ij} is the cross-equation variance-covariance matrix between equations i and j.

Correlations



Improvement in Prediction Efficiency

$$\eta = \frac{se_{i,k}(2SLS)}{se_{i,k}(3SLS)}$$

where

$$se_{ik} = \sqrt{x'_{ik}\hat{C}_{ii}x_{ik}}$$

and x'_{ik} is the row vector of covariates for a single tree observation[a].

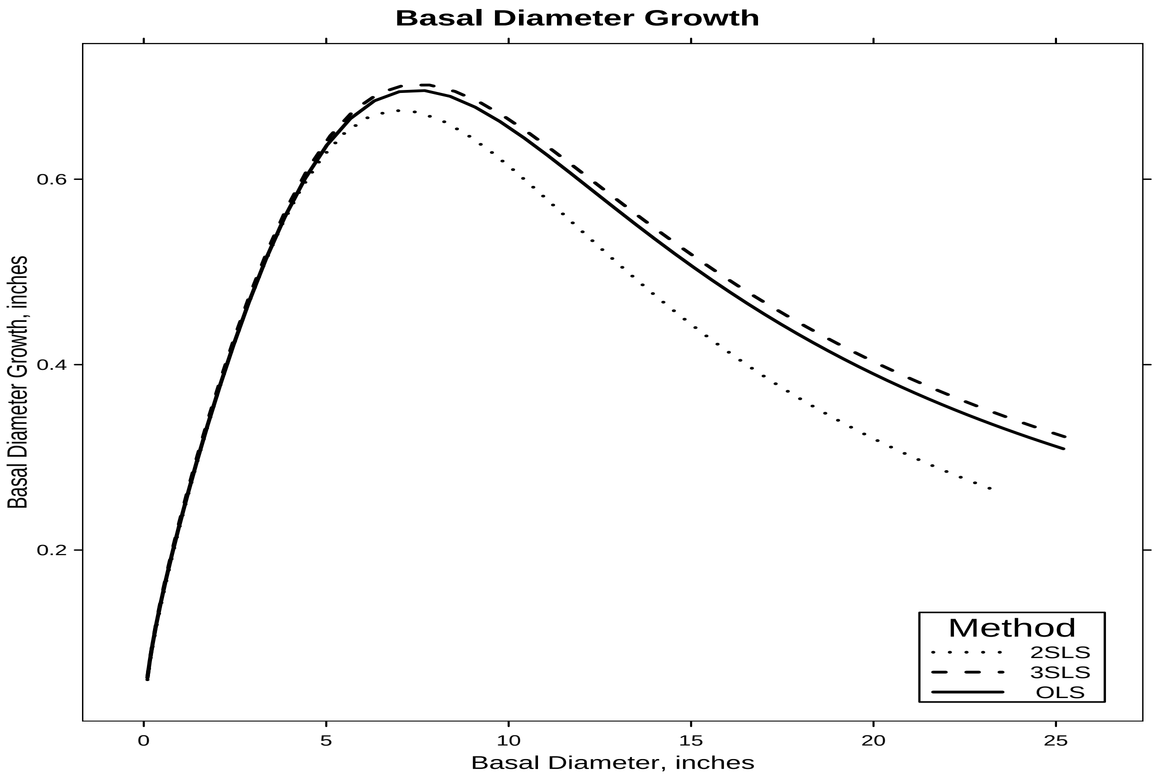
[a] Hasenauer et al., 2001

Improvement Results

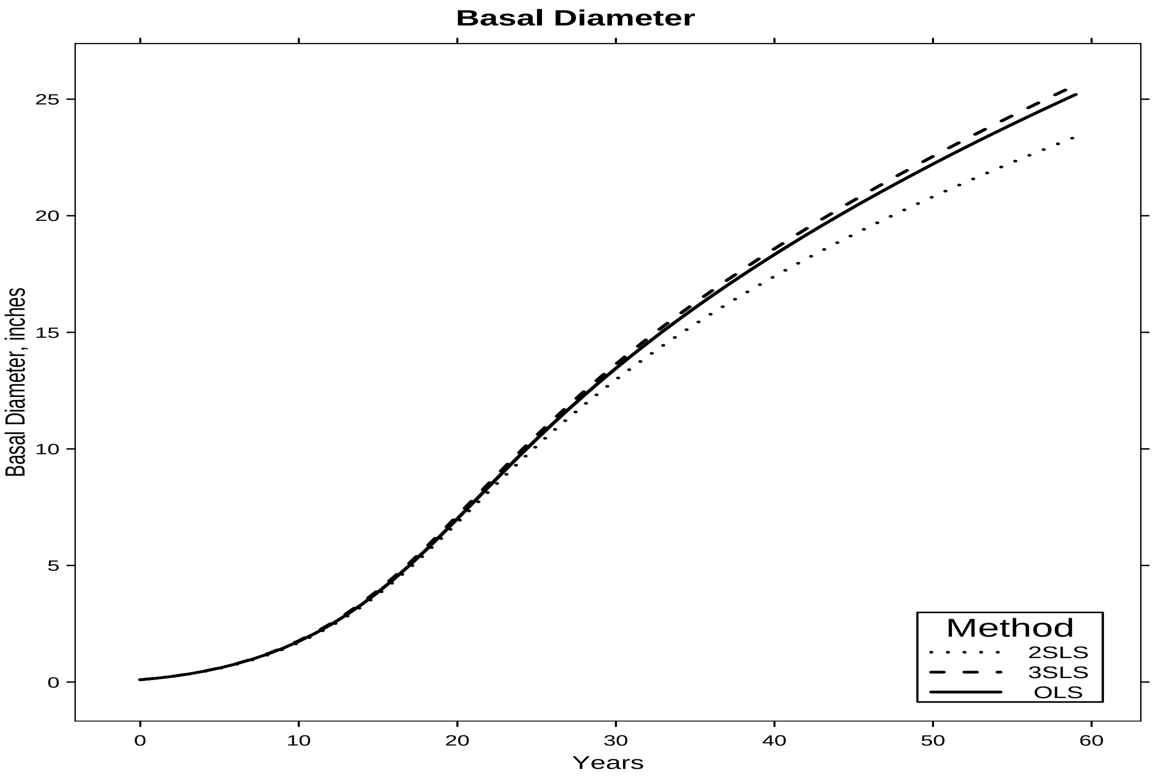
	BINC	HINC	LOGITCR2
Min.	1.000	1.000	1.000
Mean	1.016	1.003	1.005
Max.	1.034	1.006	1.028

1.00 = no improvement

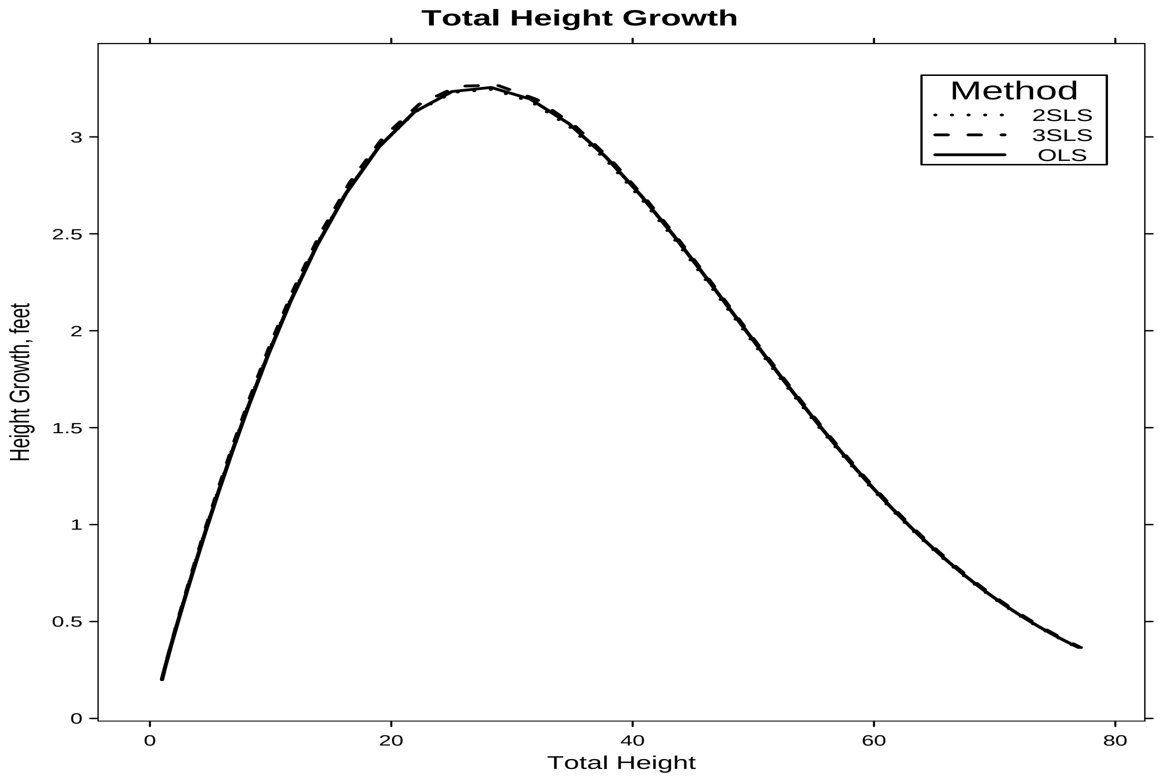
Basal Diameter Growth



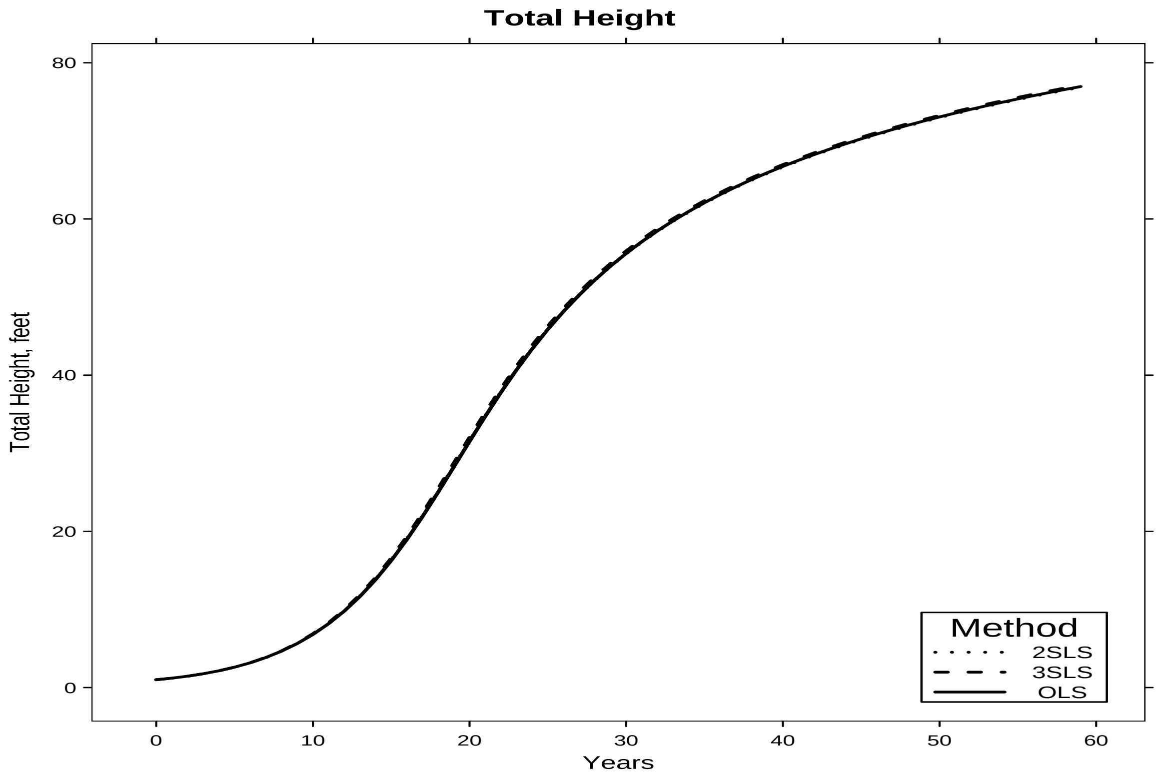
Basal Diameter



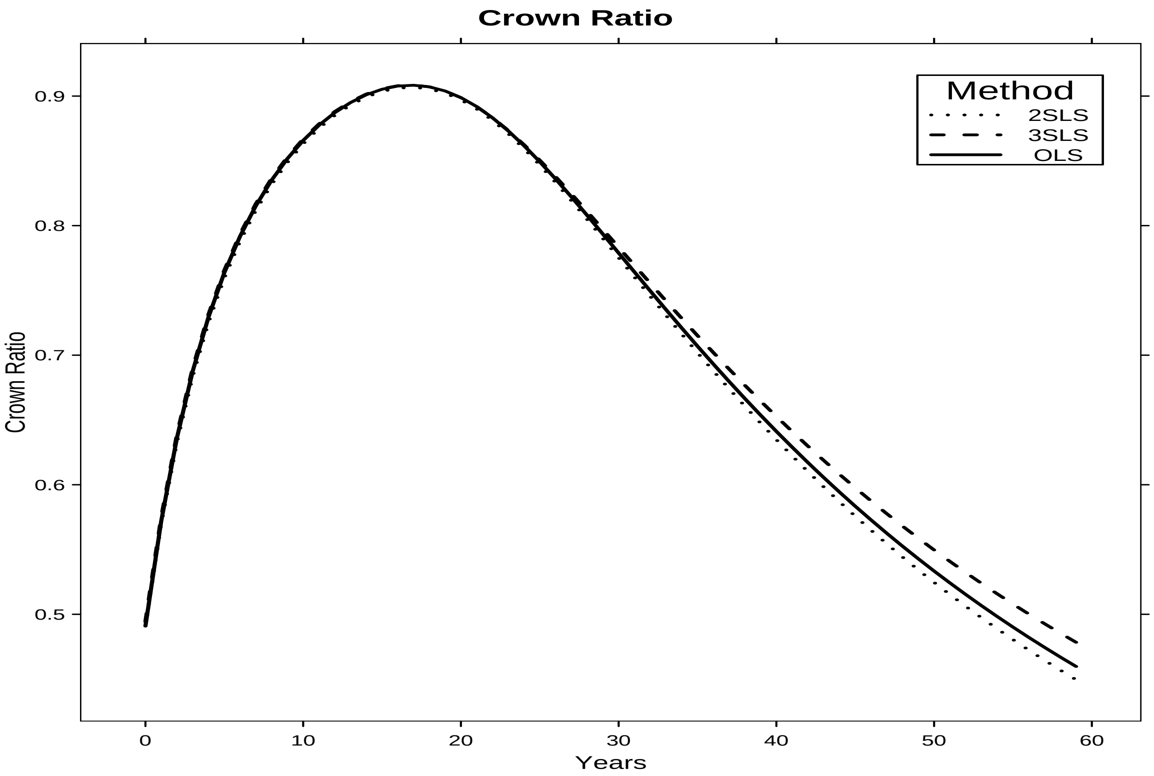
Height Growth



Total Height

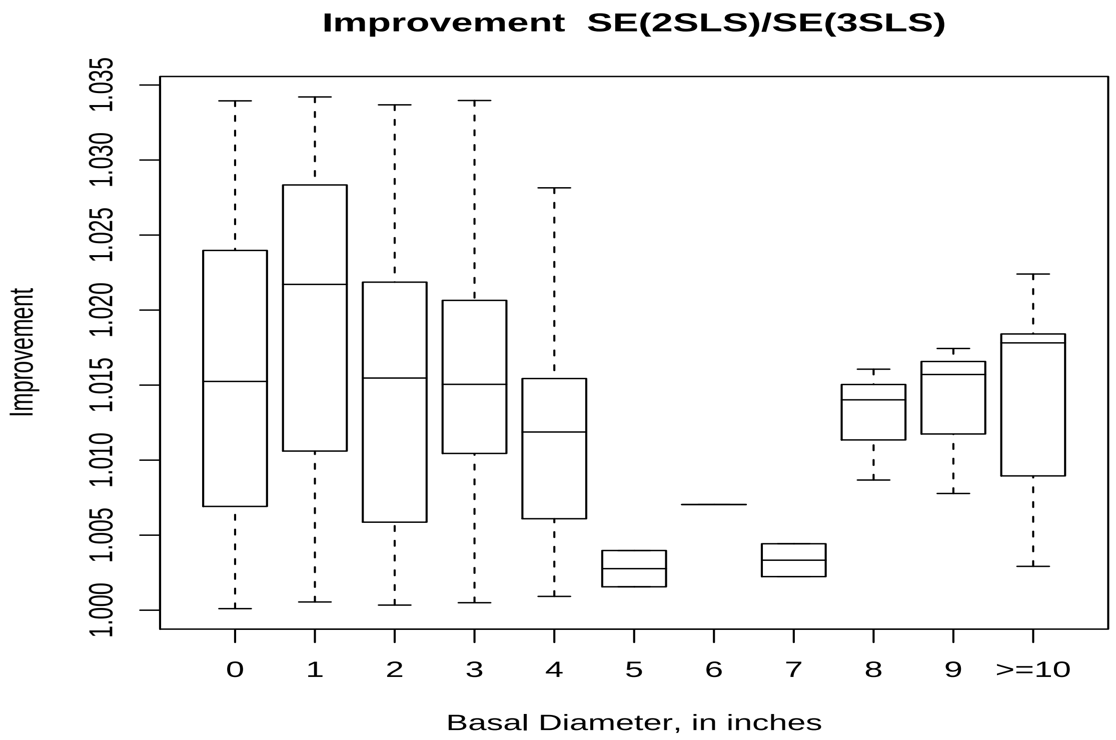


Crown Ratio

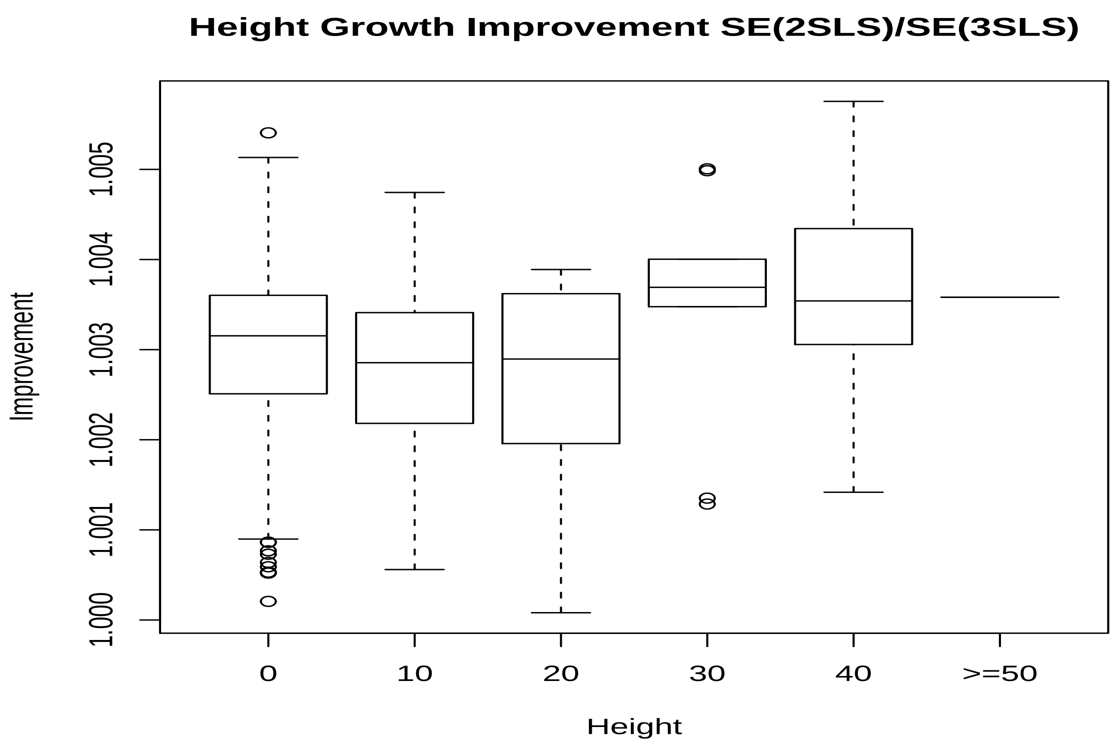


Crown Ratio
Method
2SLS
3SLS
OLS
0.9
0.8
0.7
0.6
0.5
Crown Ratio
0
10
20
30
40
50
60
Years

Basal Diameter Improvement



Height Growth Improvement



R Project

- Command line
- OpenSource version of S/S-Plus
- Runs on multiple platforms
- Freebie!
- Extensive user contributions

systemfit example - model formulation

```
library( systemfit )
data( kmenta )
demand <- q ~ p + d
supply <- q ~ p + f + a
labels <- list( "demand", "supply" )
system <- list( demand, supply )

## OLS estimation
fitols <- systemfit("OLS", system, labels, data=kmenta )
print( fitols )

## 2SLS estimation with different instruments in each equation
inst1 <- ~ d + f
inst2 <- ~ d + f + a
instlist <- list( inst1, inst2 )
fit2sls2 <- systemfit( "2SLS", system, labels, instlist, kmenta )
print( fit2sls2 )

## 3SLS estimation
inst <- ~ d + f + a
fit3sls <- systemfit( "3SLS", system, labels, inst, kmenta )
hg.se.ratio2 <- se.ratio.systemfit( fit2sls, fit3sls, 2 )
print( fit3sls )
```

systemfit example - systemwide results

```
systemfit results
method: 3SLS

        N DF      SSR     MSE    RMSE       R2   Adj R2
demand 20 17  65.7291 3.86642 1.96632 0.754847 0.726005
supply 20 16 107.9138 6.74461 2.59704 0.597508 0.522041

The covariance matrix of the residuals used for estimation
        demand  supply
demand 3.86642 4.35744
supply 4.35744 6.03958

The covariance matrix of the residuals
        demand  supply
demand 3.86642 5.00443
supply 5.00443 6.74461

The correlations of the residuals
        demand  supply
demand 1.00000 0.97999
supply 0.97999 1.00000

The determinant of the residual covariance matrix: 1.03320
OLS R-squared value of the system: 0.676177
McElroy's R-squared value for the system: 0.786468
```

systemfit example - single equation

```
3SLS estimates for demand  (equation 1 )
Model Formula: q ~ p + d
Instruments: ~d + f + a

             Estimate Std. Error   t value Pr(>|t|)
(Intercept) 94.633304   7.920838 11.947385        0 ***
p           -0.243557   0.096484 -2.524313 0.021832   *
d            0.313992   0.046944  6.688695    4e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.966321 on 17 degrees of freedom
Number of observations: 20 Degrees of Freedom: 17
SSR: 65.729088 MSE: 3.866417 Root MSE: 1.966321
Multiple R-Squared: 0.754847 Adjusted R-Squared: 0.726005
```

Was it worth it?

- Yes
  - 3SLS produces correct standard errors
  - Helpful in system model development
  - Developed systemfit package
- No
  - Little Difference in range of data
  - Little improvement in precision

Questions/Discussion

- Thank You
- R Project
- systemfit package