Statistical Software and Freeware for Data Analysis in Management

Matthew Liao-Troth
Olin School of Business, Washington University
liao-troth@olin.wustl.edu

Joel West
Graduate School of Management, U.C. Irvine
joelwest@uci.edu
http://pobox.com/~joelwest

Academy of Management
MED/RM Divisions
Toronto, Ontario

August 6, 2000

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Format of Presentation

- **Overview of packages and features**
- **Examples for undergrad, MBA and PhD classes**
  - **Research statistics are superset of these analysis techniques**
- **Licensing/administrative issues**
- **Discussion**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

1. Comparing the Packages

- **Evaluation criteria**
- **List of packages**
- **Statistical tools provided**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Evaluation Criteria

- **Statistical analysis capabilities**
- **Supported computer configurations**
- **Graphics capabilities and user friendliness**
- **Extensibility/programmability**
- **Prior experience (switching cost)**
- **Cost**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Basic Analysis Tools

- **Descriptives**
- **Correlation**
- **Chi-square hypothesis testing**
- **1-D, 2-D data plotting**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Regression

- **Regression, multiple regression**
- **ANOVA/MANOVA**
- **Time series (ARIMA, ARCH)**
- **Survival analysis (hazard models)**
- **Logit/probit**
- **GLM**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Other Analysis Techniques

- **Cluster**
- **Conjoint**
- **Factor analysis: ML, principal component**
- **Structural equation models**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Related Capabilities

- **Bootstrap/jackknife estimation**
- **Wald, likelihood ratio tests**
- **Monte carlo simulation**
- **Matrix mathematics**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Commercial Packages

- **JMP***
- **Minitab**
- **SAS**
- **SPSS***
- **Stata***
- **Statistica**
- **WINKS***

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Freeware Packages

- **ADE-4 (U. Lyon)**
- **Lisp-Stat (U. Minnesota/UCLA)**
  - **ViSta (UNC)***
  - **WinAnova/MacAnova (U. Minnesota)**
- **R (U. Auckland/CMU)***

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Supported Configurations

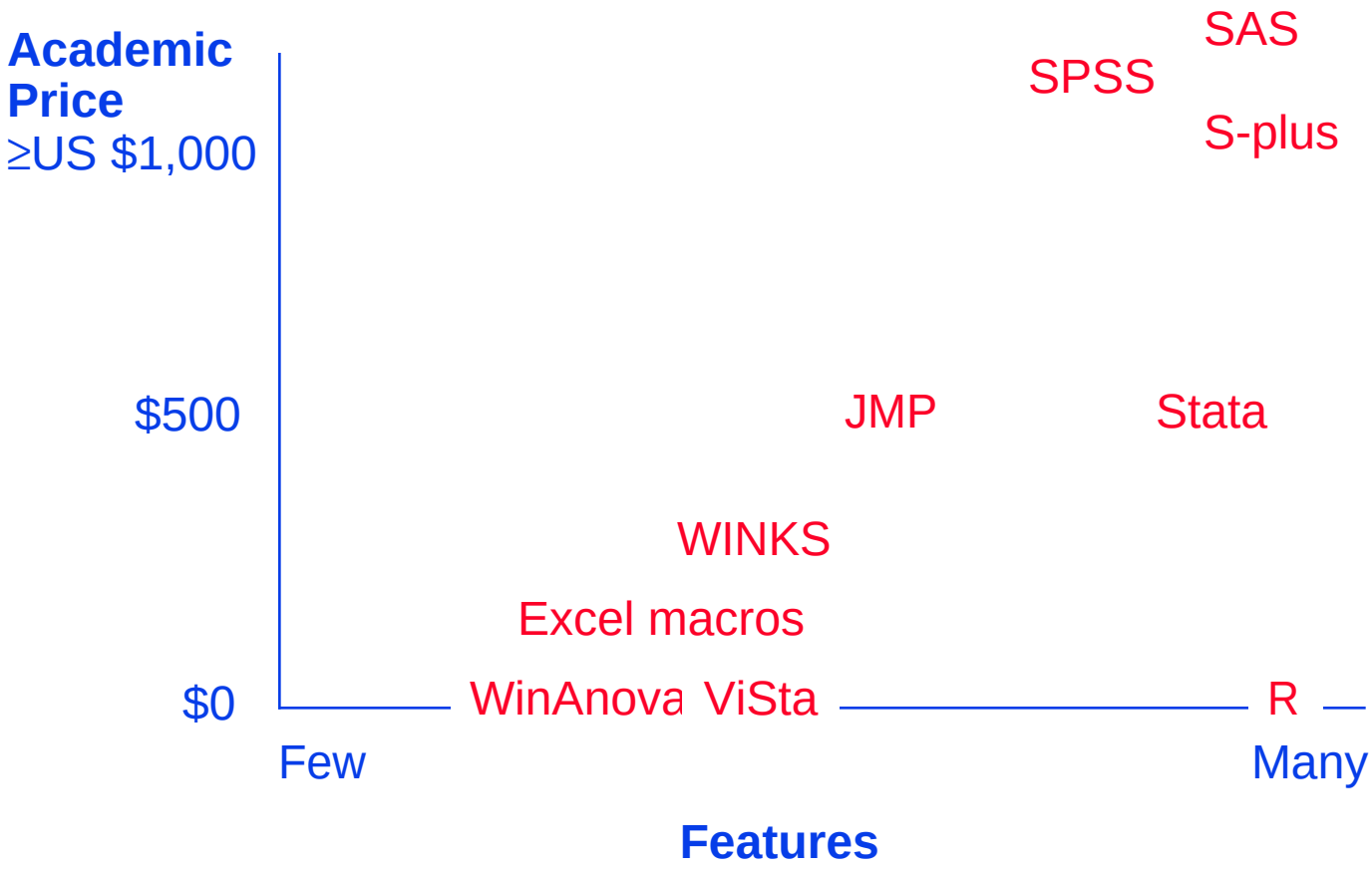
CPU:	Intel			PowerPC		Other	
OS:	95/98	2000	Linux	MacOS	Linux	Unix	M.F.
JMP	•	•		•			
Minitab	•	•		†			
SAS	•	•	*	†		•	•
SPSS	•	•		*			
Stata	•	•	•	•	•	•	
Statistica	•	•		†			
WINKS	•	•					
ADE-4	•	•		•			
Lisp-Stat	•	•		•		•	
R	•	•	•		•	•	
Vista	•	•	•	•		•	
WinAnova	•	•		•		•	

* due later this year † discontinued, but still available

UCI Graduate School of Management

Olin School of Business, Washington University

Price vs. Features



UCI Graduate School of Management

Olin School of Business, Washington University

Statistical Capabilities

	JMP	SPSS	Stata	WINKS	R	ViSta
Basic Analysis Tools						
Descriptives	•	•	•	•	•	•
Correlation	•	•	•	•	•	
1-D, 2-D data plotting	•	•	•	Basic,Pro	•	•
Regression						
Regression, multiple regression	•	•	•	•	•	•
ANOVA/MANOVA	•	•	•	•	•	•
Time series (ARIMA, ARCH)	•	Trends	•	Pro	arima	
Survival analysis (hazard models)	•	Advanced	•	Basic,Pro	survival5	
Logit/probit	•	Regression	•		•	
GLM			•		•	
Other Analysis Techniques						
Cluster	•	•			multiv	
Conjoint		Conjoint				
Multi-dimensional scaling		Categories			cmdscale	•
Principal components	•	Categories	•		multiv	•
ML factor analysis		•	•			
Structural equation models		Amos			pls	
Related Capabilities						
Bootstrap/jackknife		Regression	•		boot	
Wald, LR tests	•	Exact	•			
Monte carlo simulation		Exact	•		mcmc	
Matrix mathematics	•	•	•		•	

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

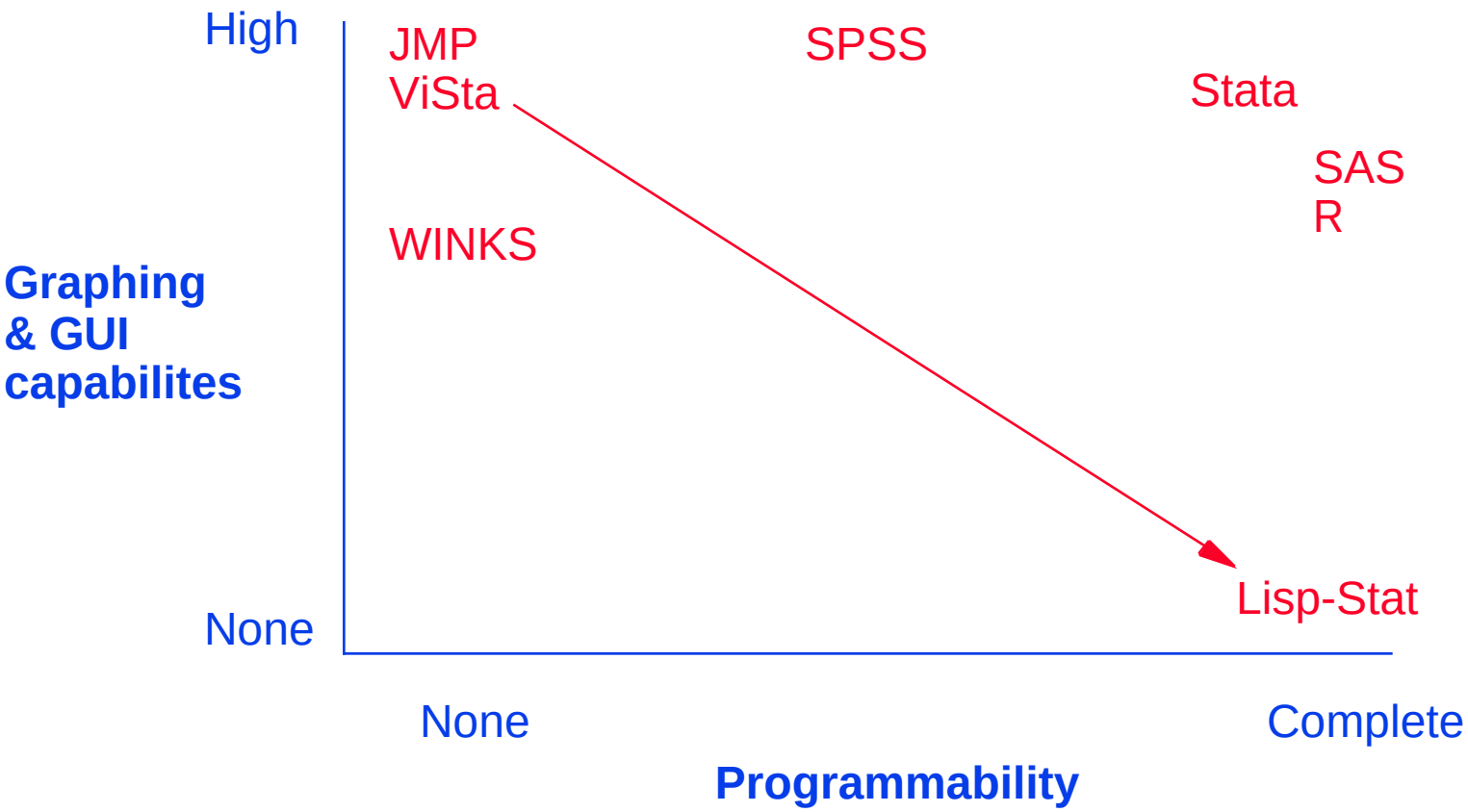
Other Features

- **Graphical design**
  - » **Charting/graphing capabilities**
  - » **Graphical user interface (GUI)**
- **Ease of use**
- **Extensibility**
  - » **Optional modules**
  - » **Programmability**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Graphics/Programming Capabilities



UCI Graduate School of Management

Olin School of Business, Washington University

2. Teaching Undergrad Stats

- **Typical syllabus**
- **Introduce the packages**
- **Demonstrations**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Upper-Division Syllabus

- **Vectors and matrix concepts**
- **Descriptive statistics**
  - **Mean, standard deviation**
  - **Variation, sum of squares**
- **Correlation**
  - **Scatter plots**
- **Least squares regression**
  - **Linear**
  - **Simple non-linear forms (quadratic, cubic, exponential, logarithmic)**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Demonstration

- **Packages:**
  - **JMP**
  - **WINKS**
- **Test problems:**
  - **Cross-tabs**
  - **Correlations**
  - **Visualization**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

3. Teaching MBA Stats

- **Typical syllabus**
- **Introduce the packages**
- **Demonstrations**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

MBA Syllabus

- **Probability concepts**
- **Hypothesis testing**
- **Correlation**
- **Linear regression**
- **ANOVA**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Demonstration

- **Packages:**
  - **SPSS**
  - **ViSta**
- **Test problems:**
  - **Regression**
  - **ANOVA**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

4. PhD Statistics and Research

- **Typical syllabus**
- **Introduce the packages**
- **Demonstrations**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

PhD Syllabus

- **Association:**
  - » **Principal components**
  - » **Factor analysis**
  - » **Cluster analysis**
- **MANOVA**
- **Multiple regression: OLS, GLS, ML**
- **Categorical variables:**
  - » **Discriminant analysis**
  - » **Logit/Probit models**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Demonstration

- **Packages:**
  - **Stata**
  - **R**
- **Test problem:**
  - **Logit**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

5. Licensing/Administrative Issues

- **Configurations**
- **Pricing**
- **Combined texts/software**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Student Versions

- **JMP**
  - » **JMP IN Book/software ($63): almost completely featured**
- **SPSS**
  - » **†Student Version ($60): limited features and datasets**
  - » **Book/software ($70-$90?): usable for 13 months only**
  - » **†Career starter ($85): all "base" module features**
  - » **†Grad pack ($175): full features, with popular modules**
- **Stata**
  - » **StataQuest book/software ($35), limited functionality**
  - » **†Small Stata ($55): full features but limited datasets**
  - » **†Grad Plan ($100): full features**

† Limited printed documentation

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Computer Lab Site Licenses

- **JMP**
  - » **Annual license**
  - » **$160/seat plus $80/seat/year after year 1 (min. 10 seats)**
- **SPSS**
  - » **Annual or perpetual licenses**
  - » **“There are many variables to site license pricing”**
- **Stata**
  - » **Perpetual license**
  - » **$70/seat ($85 for UNIX) plus $125 one-time fee (min. 10 seats)**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Full Academic Licenses

- **JMP**
  - » **Perpetual license: $395**
- **SPSS**
  - » **“Base” license: $450**
  - » **Additional modules: $250-$450 each**
- **Stata**
  - » **Perpetual license: $445**

Prices are less with site licenses

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

For More Information

Robert Stine & John Fox, eds., *Statistical Computing Environments for Social Research*, Sage, 1996.

- **APL2**
- **Gauss**
- **Lisp-Stat**
- **Mathematica**
- **SAS**
- **Stata**
- **S-Plus (and thus R)**
- **ViSta**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

6. Discussion

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Commercial SEM Packages

The most popular SEM packages are:

- **AMOS**
- **EQS**
- **LISREL**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Free SEM Packages

Most of the free packages are implementations of Wold's PLS algorithm, which has very different estimation properties. These packages include:

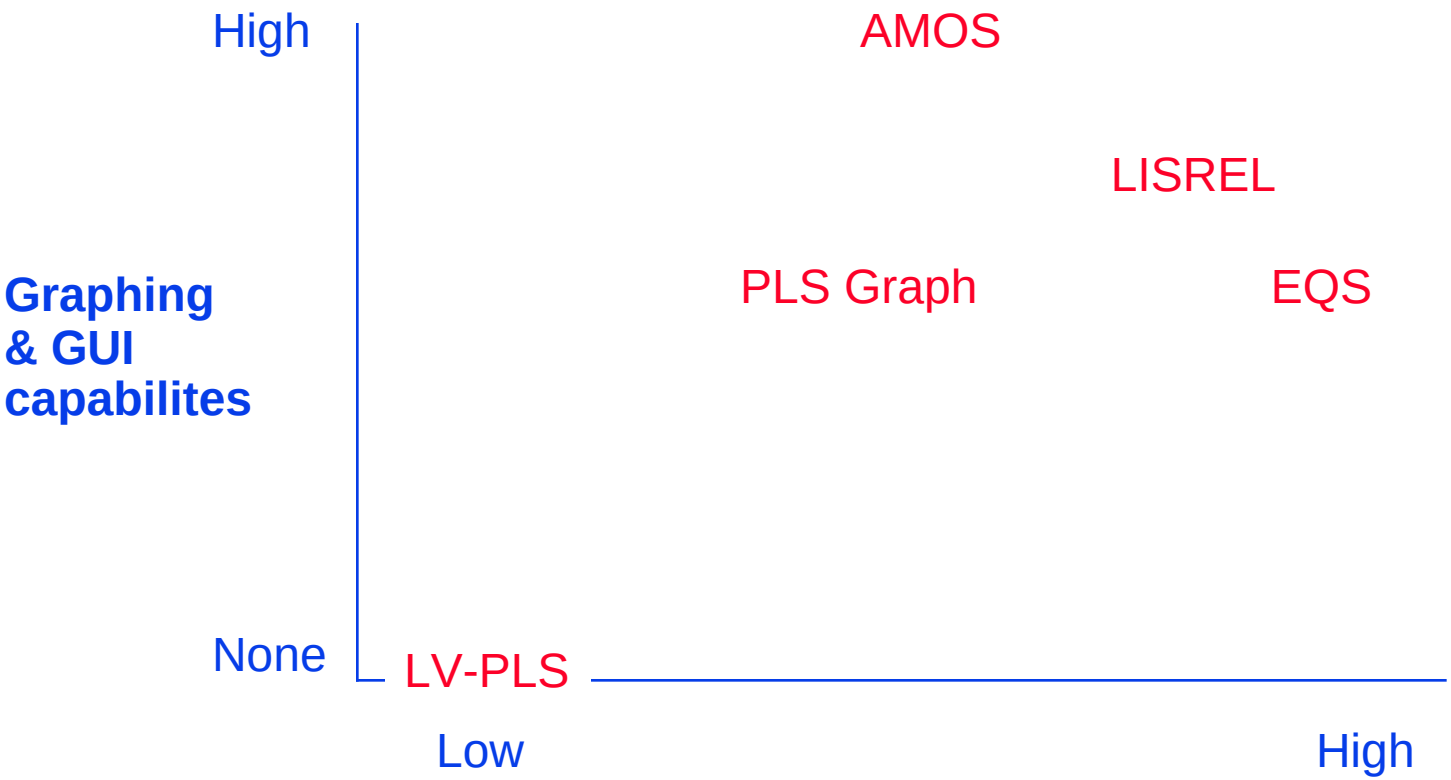
- **LV-PLS (MS-DOS program)**
- **PLS.SAS and PROC PLS (requires SAS)**
- **pls (requires R)**

See also Mx (free) and PLS Graph (beta)

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Graphics vs. Statistics



UCI Graduate School of Management

Olin School of Business, Washington University

Key Issues

- **Statistical analysis capabilities:**
  - » **Dynamically changing methodology**
  - » **Thus, important to keep up to date with developments**
  - » **Major issues (differences between packages):**
    - » **Robust estimation methods**
    - » **Measures of goodness of fit**
- **Graphics & GUI**
- **Price**

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business

Web Page for More Information

Type

http://pobox.com/~joelwest

and follow the links

UCI
GRADUATE
SCHOOL OF
MANAGEMENT

OLIN
Washington University
School of Business