The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

cron, perl and Stata: Automated production and presentation of a business-daily index

North American Stata Users Group, Boston, 11–12 July 2005

Christopher F Baum[1]   Atreya Chakraborty[2]

[1]Department of Economics
Boston College

[2]College of Management
University of Massachusetts–Boston

8th July 2005

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Outline

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Risk Appetite Index for Massachusetts (RAIM)

- RAIM: an index, developed by Profs. Bandopadhyaya and Jones of UMass-Boston College of Management, that measures investors' sentiment toward firms active in the local economy.

- The index deals with the notion that regional economies' market dynamics may differ substantially from national trends, and some investors may have a local preference in forming their portfolios.

- RAIM is based on stockmarket quotes from the Bloomberg "Massachusetts Index" of about 240 traded firms based in or doing considerable business in Massachusetts. The firms' quotes and the Mass Index are available daily from the Bloomberg website.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Risk Appetite Index for Massachusetts (RAIM)

- RAIM: an index, developed by Profs. Bandopadhyaya and Jones of UMass-Boston College of Management, that measures investors' sentiment toward firms active in the local economy.

- The index deals with the notion that regional economies' market dynamics may differ substantially from national trends, and some investors may have a local preference in forming their portfolios.

- RAIM is based on stockmarket quotes from the Bloomberg "Massachusetts Index" of about 240 traded firms based in or doing considerable business in Massachusetts. The firms' quotes and the Mass Index are available daily from the Bloomberg website.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Risk Appetite Index for Massachusetts (RAIM)

- RAIM: an index, developed by Profs. Bandopadhyaya and Jones of UMass-Boston College of Management, that measures investors' sentiment toward firms active in the local economy.
- The index deals with the notion that regional economies' market dynamics may differ substantially from national trends, and some investors may have a local preference in forming their portfolios.
- RAIM is based on stockmarket quotes from the Bloomberg "Massachusetts Index" of about 240 traded firms based in or doing considerable business in Massachusetts. The firms' quotes and the Mass Index are available daily from the Bloomberg website.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Quantifying the Market's Appetite for Risk

- Research in currency markets has shown that the market's changing appetite for risk is the most influential factor affecting currency returns.

- Values of the RAIM index, measuring the market's appetite for risk, allow classification of trading days into risk seeking, risk averse or neutral. Movements in the RAIM are triggered by both good news and bad news relating to the Mass. economy.

- Changes in the RAIM index have been shown to be highly correlated to changes in the Bloomberg Massachusetts Index, a measure of returns from holding a portfolio of Massachusetts firms.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Quantifying the Market's Appetite for Risk

- Research in currency markets has shown that the market's changing appetite for risk is the most influential factor affecting currency returns.

- Values of the RAIM index, measuring the market's appetite for risk, allow classification of trading days into risk seeking, risk averse or neutral. Movements in the RAIM are triggered by both good news and bad news relating to the Mass. economy.

- Changes in the RAIM index have been shown to be highly correlated to changes in the Bloomberg Massachusetts Index, a measure of returns from holding a portfolio of Massachusetts firms.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Quantifying the Market's Appetite for Risk

- Research in currency markets has shown that the market's changing appetite for risk is the most influential factor affecting currency returns.

- Values of the RAIM index, measuring the market's appetite for risk, allow classification of trading days into risk seeking, risk averse or neutral. Movements in the RAIM are triggered by both good news and bad news relating to the Mass. economy.

- Changes in the RAIM index have been shown to be highly correlated to changes in the Bloomberg Massachusetts Index, a measure of returns from holding a portfolio of Massachusetts firms.

The computational challenge

Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

The challenge: implementing a real-time RAIM

- RAIM's developers want a mechanism by which the RAIM will be calculated automatically from each trading day's quotations on firms in the Mass. Index, and web pages automatically updated.
- Elements of the solution, as implemented on Mac OS X 10.3: cron to schedule a daily Unix process; perl to acquire the data and put it in readable format for Stata; console version 8 of Stata to generate the RAIM and produce web pages / graphics

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

The challenge: implementing a real-time RAIM

- RAIM's developers want a mechanism by which the RAIM will be calculated automatically from each trading day's quotations on firms in the Mass. Index, and web pages automatically updated.

- Elements of the solution, as implemented on Mac OS X 10.3: cron to schedule a daily Unix process; perl to acquire the data and put it in readable format for Stata; console version 8 of Stata to generate the RAIM and produce web pages / graphics

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Data acquisition and manipulation

- A Unix cron script schedules a process to run at a specific interval. In our case, the cron script on our Mac OS X system runs every morning at 2:30 AM, and launches a perl program, bloomberg.pl.

- The perl program uses the LWP::Simple package function get() to copy the contents of the web page listing the firms in the Bloomberg Massachusetts Index, and parse that HTML into a tab-delimited .raw file readable by Stata.

- The perl program then writes a Stata program, bloomberg.do, which reads that .raw file and saves it as a .dta file, time-stamped with the quote date. Stata is launched by the cron script to run bloomberg.do.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Data acquisition and manipulation

- A Unix `cron` script schedules a process to run at a specific interval. In our case, the `cron` script on our Mac OS X system runs every morning at 2:30 AM, and launches a `perl` program, `bloomberg.pl`.
- The `perl` program uses the `LWP::Simple` package function `get()` to copy the contents of the web page listing the firms in the Bloomberg Massachusetts Index, and parse that HTML into a tab-delimited `.raw` file readable by Stata.
- The `perl` program then writes a Stata program, `bloomberg.do`, which reads that `.raw` file and saves it as a `.dta` file, time-stamped with the quote date. Stata is launched by the `cron` script to run `bloomberg.do`.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Data acquisition and manipulation

- A Unix `cron` script schedules a process to run at a specific interval. In our case, the `cron` script on our Mac OS X system runs every morning at 2:30 AM, and launches a `perl` program, `bloomberg.pl`.
- The `perl` program uses the `LWP::Simple` package function `get()` to copy the contents of the web page listing the firms in the Bloomberg Massachusetts Index, and parse that HTML into a tab-delimited `.raw` file readable by Stata.
- The `perl` program then writes a Stata program, `bloomberg.do`, which reads that `.raw` file and saves it as a `.dta` file, time-stamped with the quote date. Stata is launched by the `cron` script to run `bloomberg.do`.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Generating a panel of daily quotes

- The `cron` script then launches Stata to run `bloomberg_combine.do`, which uses N.J. Cox's `fs` command to locate all time-stamped daily quote files.

- The program uses Stata's `cf` command to determine whether files are redundant: e.g., those generated on Sunday or Monday morning, which like Saturday morning's file will be based on Friday trading. This also removes holidays from the set under consideration.

- When all unique files are identified, they are appended into `BloombergAll.dta`, set up as a panel (`tsset`) with firm name (encoded to numeric format) and trading day number (not date). The panel is unbalanced since firms may be added or dropped each day.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Generating a panel of daily quotes

- The `cron` script then launches Stata to run `bloomberg_combine.do`, which uses N.J. Cox's `fs` command to locate all time-stamped daily quote files.
- The program uses Stata's `cf` command to determine whether files are redundant: e.g., those generated on Sunday or Monday morning, which like Saturday morning's file will be based on Friday trading. This also removes holidays from the set under consideration.
- When all unique files are identified, they are appended into `BloombergAll.dta`, set up as a panel (`tsset`) with firm name (encoded to numeric format) and trading day number (not date). The panel is unbalanced since firms may be added or dropped each day.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Generating a panel of daily quotes

- The cron script then launches Stata to run bloomberg_combine.do, which uses N.J. Cox's fs command to locate all time-stamped daily quote files.
- The program uses Stata's cf command to determine whether files are redundant: e.g., those generated on Sunday or Monday morning, which like Saturday morning's file will be based on Friday trading. This also removes holidays from the set under consideration.
- When all unique files are identified, they are appended into BloombergAll.dta, set up as a panel (tsset) with firm name (encoded to numeric format) and trading day number (not date). The panel is unbalanced since firms may be added or dropped each day.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Computation of the RAIM index

- The RAIM is based on a specific "window width" of recent quotes, presently set to five trading days

- Stata program `calcRAIM.do` is invoked by the `cron` script to generate a five-day moving standard deviation of daily returns, $\hat{\sigma}_t$, utilizing Cox and Baum's `mvsumm` routine

- The resulting series is used in a Spearman rank correlation (`spearman`) to calculate, for each trading day, the relation between the $\hat{\sigma}_t$ of returns and the percentage changes in firms' stock prices. The resulting index is scaled to lie in the [-100,+100] interval.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Computation of the RAIM index

- The RAIM is based on a specific "window width" of recent quotes, presently set to five trading days
- Stata program `calcRAIM.do` is invoked by the `cron` script to generate a five-day moving standard deviation of daily returns, $\hat{\sigma}_t$, utilizing Cox and Baum's `mvsumm` routine
- The resulting series is used in a Spearman rank correlation (`spearman`) to calculate, for each trading day, the relation between the $\hat{\sigma}_t$ of returns and the percentage changes in firms' stock prices. The resulting index is scaled to lie in the [-100,+100] interval.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Computation of the RAIM index

- The RAIM is based on a specific "window width" of recent quotes, presently set to five trading days
- Stata program calcRAIM.do is invoked by the cron script to generate a five-day moving standard deviation of daily returns, $\hat{\sigma}_t$, utilizing Cox and Baum's mvsumm routine
- The resulting series is used in a Spearman rank correlation (spearman) to calculate, for each trading day, the relation between the $\hat{\sigma}_t$ of returns and the percentage changes in firms' stock prices. The resulting index is scaled to lie in the [-100,+100] interval.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Generating a tabular display of the RAIM

- Program `calcRAIM.do` then uses Stata's `file` command to write a text file, `raim.php`, in HTML format with embedded PHP directives. The RAIM index for each trading day, sorted (`gsort`) in reverse chronological order, is displayed in a HTML table.

- Since the RAIM index is a product of Boston's public business school, UMass-Boston, headers and footers from the College of Management web site are added to give the RAIM display the proper brand identity

- Links are added for additional views of the available history of RAIM values

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Generating a tabular display of the RAIM

- Program `calcRAIM.do` then uses Stata's `file` command to write a text file, `raim.php`, in HTML format with embedded PHP directives. The RAIM index for each trading day, sorted (`gsort`) in reverse chronological order, is displayed in a HTML table.

- Since the RAIM index is a product of Boston's public business school, UMass-Boston, headers and footers from the College of Management web site are added to give the RAIM display the proper brand identity

- Links are added for additional views of the available history of RAIM values

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Generating a tabular display of the RAIM

- Program `calcRAIM.do` then uses Stata's `file` command to write a text file, `raim.php`, in HTML format with embedded PHP directives. The RAIM index for each trading day, sorted (`gsort`) in reverse chronological order, is displayed in a HTML table.

- Since the RAIM index is a product of Boston's public business school, UMass-Boston, headers and footers from the College of Management web site are added to give the RAIM display the proper brand identity

- Links are added for additional views of the available history of RAIM values

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Generating a graphical display of the RAIM

- A simple `tsline` graph of the RAIM for each trading day is generated and saved in PostScript (`.ps`) format via `graph export`.
- A call to the Unix shell invokes the GhostScript utility `ps2png`, which translates the graph into `.png` format, suitable for inclusion on a web page

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Generating a graphical display of the RAIM

- A simple `tsline` graph of the RAIM for each trading day is generated and saved in PostScript (`.ps`) format via `graph export`.
- A call to the Unix shell invokes the GhostScript utility `ps2png`, which translates the graph into `.png` format, suitable for inclusion on a web page

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Generating a tab-delimited listing of the RAIM

- Another web page is generated by Stata's `file` command: `raim.txt`, listing the trading day and RAIM value in tab-delimited format
- This page, linked to the main RAIM page, may be saved to disk if analysis of the RAIM values in Stata, Excel or other packages is desired

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Generating a tab-delimited listing of the RAIM

- Another web page is generated by Stata's `file` command: `raim.txt`, listing the trading day and RAIM value in tab-delimited format
- This page, linked to the main RAIM page, may be saved to disk if analysis of the RAIM values in Stata, Excel or other packages is desired

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Some issues...

- The routine must be run in Stata 8.2, since a console version for Mac OS X is not available for Stata 9, and the "hands-off" mode of operation cannot be accomplished with the GUI version of Stata.

- Officlal spearman produces a *t*-test for the hypothesis that $\rho = 0$, but no standard error of ρ. I would like to produce 95% confidence bands, and have written spearman2 to return the standard error of ρ as well.

- The console version of Stata interfaces very well with other Unix components: cron, perl, PHP and the Apache web server to provide a fully automated solution to the challenge of generating a business-daily index.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Some issues...

- The routine must be run in Stata 8.2, since a console version for Mac OS X is not available for Stata 9, and the "hands-off" mode of operation cannot be accomplished with the GUI version of Stata.

- Officlal spearman produces a *t*-test for the hypothesis that $\rho = 0$, but no standard error of ρ. I would like to produce 95% confidence bands, and have written spearman2 to return the standard error of ρ as well.

- The console version of Stata interfaces very well with other Unix components: cron, perl, PHP and the Apache web server to provide a fully automated solution to the challenge of generating a business-daily index.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Some issues...

- The routine must be run in Stata 8.2, since a console version for Mac OS X is not available for Stata 9, and the "hands-off" mode of operation cannot be accomplished with the GUI version of Stata.

- Officlal spearman produces a *t*-test for the hypothesis that $\rho = 0$, but no standard error of ρ. I would like to produce 95% confidence bands, and have written spearman2 to return the standard error of ρ as well.

- The console version of Stata interfaces very well with other Unix components: cron, perl, PHP and the Apache web server to provide a fully automated solution to the challenge of generating a business-daily index.

The computational challenge
Assembling the data automatically
Computing the index with Stata
Generating the web pages
Some issues

Screen shots

Risk Appetite Index in Massachusetts [RAIM]

Quote date	RAIM	P-value	# Firms
06jul2005	-12.7782	0.047	242
05jul2005	2.7844	0.666	242
01jul2005	7.9994	0.215	242
30jun2005	-22.7382	0.000	242
29jun2005	-4.6362	0.473	242
28jun2005	31.2105	0.000	242
27jun2005	-18.2089	0.004	242
24jun2005	1.4824	0.819	242
23jun2005	-25.3281	0.000	241
22jun2005	0.6189	0.924	241
21jun2005	-11.2244	0.082	241
20jun2005	11.8934	0.066	240
19jun2005	21.7152	0.001	240
17jun2005	22.9104	0.000	241
16jun2005	13.0527	0.043	241
15jun2005	-1.6338	0.801	241
14jun2005	2.2860	0.723	242
13jun2005	2.1315	0.741	242
10jun2005	2.3420	0.717	242
09jun2005	10.7858	0.095	241
08jun2005	3.2513	0.615	241
07jun2005	-6.6923	0.300	242
06jun2005	0.3261	0.960	242
05jun2005	-12.1187	0.060	242
03jun2005	-11.3955	0.076	243
02jun2005	0.4771	0.941	243
01jun2005	1.5849	0.806	243
31may2005	-4.1381	0.520	244
29may2005	9.6685	0.132	244
28may2005	7.4121	0.249	244
26may2005	21.6962	0.001	244
25may2005	-5.3862	0.402	244
24may2005	-6.1347	0.340	244
20may2005	7.5806	0.238	244
19may2005	-3.7461	0.561	243
18may2005	23.4830	0.000	243
17may2005	0.7538	0.907	243

16may2005	-5.2037	0.419	243
13may2005	-4.2227	0.511	244
12may2005	-7.8106	0.224	244
11may2005	-0.8179	0.899	244
10may2005	-11.5883	0.070	245
09may2005	17.7676	0.005	245
06may2005	-1.0903	0.865	245
05may2005	0.6442	0.920	245
04may2005	14.5780	0.022	245
03may2005	-1.9823	0.758	245
02may2005	3.2063	0.617	245
01may2005	19.5772	0.002	246
29apr2005	12.3132	0.054	246
28apr2005	-25.6828	0.000	245
27apr2005	1.7532	0.785	245
26apr2005	-18.8534	0.003	245
25apr2005	5.7389	0.371	245
22apr2005	-25.8827	0.000	245
21apr2005	35.2729	0.000	245

RAIM graph | downloadable text file

UMass Boston College of Management
Boston's Public Business School

COLLEGE OF
Management
UMASS BOSTON

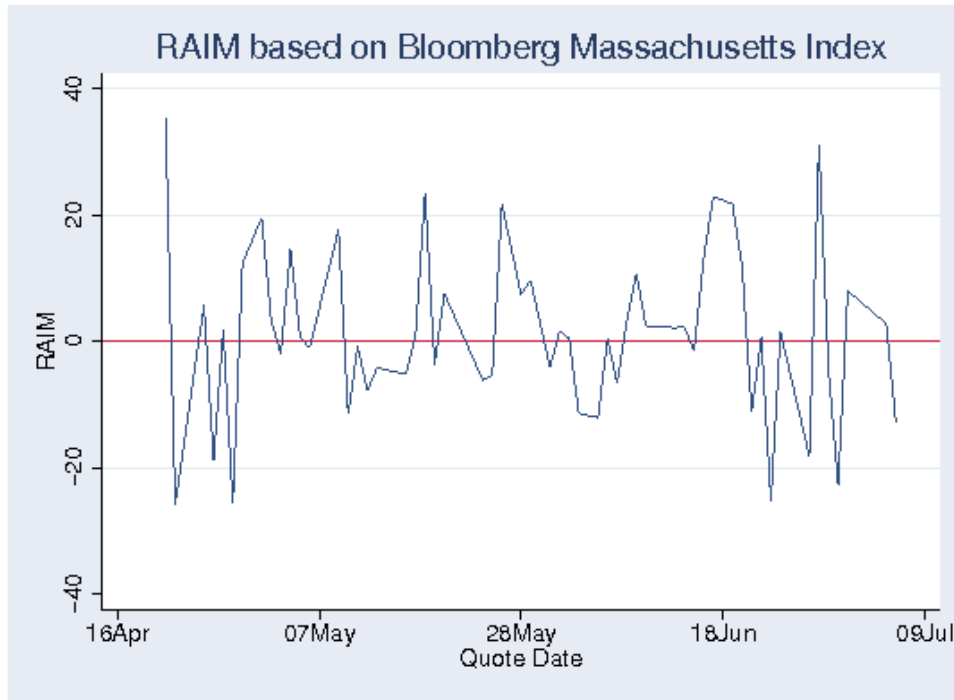## Risk Appetite Index in Massachusetts [RAIM]

RAIM based on Bloomberg Massachusetts Index

Home | Contact | Sitemap | About Us | Undergraduate Program | Graduate Programs | Faculty & Research | Business Centers | Calendar

UMass Boston Homepage | Privacy Policy

University of Massachusetts Boston College of Management © 2005

```
quotedate      RAIM
21apr2005      35.2729
22apr2005      -25.8827
25apr2005      5.7389
26apr2005      -18.8534
27apr2005      1.7532
28apr2005      -25.6828
29apr2005      12.3132
01may2005      19.5772
02may2005      3.2063
03may2005      -1.9823
04may2005      14.5780
05may2005      0.6442
06may2005      -1.0903
09may2005      17.7676
10may2005      -11.5883
11may2005      -0.8179
12may2005      -7.8106
13may2005      -4.2227
16may2005      -5.2037
17may2005      0.7538
18may2005      23.4830
19may2005      -3.7461
20may2005      7.5806
24may2005      -6.1347
25may2005      -5.3862
26may2005      21.6962
28may2005      7.4121
29may2005      9.6685
31may2005      -4.1381
01jun2005      1.5849
02jun2005      0.4771
03jun2005      -11.3955
05jun2005      -12.1187
06jun2005      0.3261
07jun2005      -6.6923
08jun2005      3.2513
09jun2005      10.7858
10jun2005      2.3420
13jun2005      2.1315
14jun2005      2.2860
15jun2005      -1.6338
16jun2005      13.0527
17jun2005      22.9104
19jun2005      21.7152
20jun2005      11.8934
21jun2005      -11.2244
22jun2005      0.6189
23jun2005      -25.3281
24jun2005      1.4824
27jun2005      -18.2089
28jun2005      31.2105
29jun2005      -4.6362
30jun2005      -22.7382
01jul2005      7.9994
05jul2005      2.7844
06jul2005      -12.7782
```