

Syllabus for Summer Mini-Course on STATA/MATLAB/C/Perl/ArcGIS

Welcome to the IAP “Programming for Economists” mini-course! Based on the positive feedback we received, we have decided to offer the mini-course again this IAP. This intensive five-day course is designed to help anyone interested in doing empirical work become familiar with the basic structure of STATA/MATLAB/C/Perl/ArcGIS. It will also introduce useful programming tricks and tools that are relevant for economic research, such as methods of dealing with large data sets, dynamic programming and optimization. The course will be taught by graduate students, and it will meet in 14-0637 between January 12th and January 16th. If you are planning to attend the course, please send email to Matt Notowidigdo <noto@mit.edu> to reserve a space.

Unlike past versions of this course, we will have more in-depth practice exercises; each day after class we will have informal “office hours” to help you work through the more difficult exercises. We believe the best way to build programming expertise is to work through examples and solve problems on your own.

The best way to prepare for this mini course is to browse through the lecture notes that we used last year <http://web.mit.edu/econ-gea/14.170>, and to implement a few simple examples. Please bring your questions to the class.

As this course is (still) an experiment, feedback would be very helpful. After the course, please let one of us know if the course is worthwhile. Suggestions for improving this course would also be very welcome.

Thanks,

Panle Jia

Matt Notowidigdo

Paul Schrimpf

Melissa Dell

<u>MONDAY, JANUARY 12th</u> MORNING: Introduction to STATA AFTERNOON: Intermediate STATA (Lecturer: Matt Notowidigdo)	
8am-10am STATA (Matt)	Lecture S1: STATA must-know info <ul style="list-style-type: none"> - language review: data types, basic commands - syntax review: control structures, local variables, procedural programming - parsing, reshaping, merging effectively
10am-11am	Practice session S1
11am-12pm, 1pm-2pm STATA (Matt)	Lecture S2: Intermediate Stata <ul style="list-style-type: none"> - quick tour of built-in commands - ADO files - bootstrapping, Monte Carlo simulations - Stata matrix language
2pm-3pm	Practice session S2
3pm-5pm STATA (Matt)	Lecture S3: NLLS and MLE <ul style="list-style-type: none"> - Non-linear least squares (NLLS) estimators - Maximum likelihood estimation (MLE)
5pm-6pm	Practice session S3 and “Office Hours”

<u>TUESDAY, JANUARY 13th</u> MORNING: Advanced Stata - NLLS, MLE AFTERNOON: Advanced Stata - Mata in Stata, GMM in Stata (Lecturer: Matt Notowidigdo)	
9-11 am STATA (Matt)	Lecture S4: Mata, GMM <ul style="list-style-type: none"> - NLLS and GMM in Stata ML - Introduction to Mata - GMM in Stata using Mata - Using Mata in Stata effectively
11-12 pm	Practice session S4
12pm-1pm	LUNCH
1-3 pm STATA (Matt)	Lecture S5: Large data sets and numerical precision <ul style="list-style-type: none"> - Tips/tricks - Breaking abstraction barrier - Using Mata when Stata won't do - When to leave Stata - Matrix equilibration - Binary representation
3:00-6:00 pm	Practice Session S5 and “Office Hours”

WEDNESDAY, JANUARY 14th
MORNING: Basic MATLAB
AFTERNOON: More MATLAB
(Lecturer: Paul Schrimpf)

9-11 am MATLAB (Paul)	Lecture M1: Basic MATLAB <ul style="list-style-type: none">- how to write a basic m-file: global variables, cleaning memory, saving data files- basic language concepts: creating matrices, operations (matrix and by element)- manipulating matrices: vector, selecting rows or columns- other data types: cells and structures- writing functions
11am-12pm	Practice Session M1
12pm-1pm	LUNCH
1pm-2:30pm MATLAB (Paul)	Lecture M2: Optimization Routines <ul style="list-style-type: none">- how to use them, what they do, why they might fail
2:30pm-3:30pm	Practice session M2
3:30-5pm MATLAB (Paul)	Lecture M3: Further MATLAB tools <ul style="list-style-type: none">- creating nice output: tables and graphs- debugging and profiling- object-oriented MATLAB
5:00-6:00 pm	Practice session M3

<p align="center"><u>THURSDAY, JANUARY 15th</u> MORNING: Perl, Stata Recap, and ArcGIS AFTERNOON: ArcGIS (Lecturers: Matt Notowidigdo and Melissa Dell)</p>	
9am-10:30am Perl (Matt)	Lecture P1: Perl for Economists <ul style="list-style-type: none"> - auto-downloading data from the internet (web crawlers) - parsing data effectively using regular expressions
10:30am-12:30pm ArcGIS (Melissa)	Lecture G1: Intro to ArcGIS <ul style="list-style-type: none"> - What is GIS? - Accessing ArcGIS on the MIT network - Brief intro to ArcGIS tools likely to be of use to economists <ul style="list-style-type: none"> - <i>Analysis tools</i> - <i>Data management tools</i> - <i>Spatial analyst tools</i> - <i>3D analysis tools</i> - <i>Conversion tools</i> - <i>Hawth's Tools</i> - <i>ArcGIS quirks and bugs</i> - Projections <ul style="list-style-type: none"> - <i>Geographic coordinate systems</i> - <i>Projected coordinate systems</i> - Introduction to python scripting
12:30pm-1:30pm	LUNCH
1:30-3:30 pm ArcGIS (Melissa)	Lecture G2: More ArcGIS <ul style="list-style-type: none"> - Spatial correlation in GIS analysis - Making maps in ArcGIS - Useful GIS data resources - Examples of GIS processing in applied economics research <ul style="list-style-type: none"> - <i>The economic impacts of climate change</i> - <i>Slope, elevation, and distance in econometric analysis</i>
3:30pm-4:30pm	Practice session G1

<u>FRIDAY, JANUARY 16th</u> MORNING: Intro to C AFTERNOON: More C (Lecturer: Paul Schrimpf)	
9am-10 am C (Paul)	Lecture C1: When Can C Help? - fast & interoperable - an example or two of C being much faster than MATLAB or STATA - downside: harder to use
10am-11am	Practice session C1
11pm-12 pm C (Paul)	Lecture C2: Basics of C - how to edit, compile, and run a program - parts of a C program: #directives, functions, variable declarations, datatypes and scope - flow control: if, loops, and switches - memory and pointers:
12pm-1pm	LUNCH
1pm-3 pm C (Paul)	Lecture C3: Continue 1) pointers: what they do and why they cause nasty bugs 2) the stack and its relation to scope 3) dynamic memory allocation - input and output - debugging
3-4 pm	Practice session C2
4-5 pm C (Paul)	Lecture C4: Using C with Other Programs - useful C libraries: 1) numerical: numerical recipes, gsl, statlib, acm calgo 2) parallelization: mpi, threads, openmp, etc - calling C from MATLAB - calling C from STATA