

**PREDICT 411: Predictive Modeling II Syllabus Fall 2012**

Vivek Ajmani, Ph.D.

**vivek.ajmani@northwestern.edu**

Course Description

Predict 411 is a continuation of Predict 410 and will introduce students to basic to advanced econometric methods including multivariate regression methods for both continuous and discrete dependent variables. Further discussions will involve remedial methods under violations of least squares assumptions (Heteroscedasticity, Serial Correlation, Endogenous Variables). We will also discuss panel data analysis, seemingly unrelated regression methods, simultaneous equations, introduction to time series analysis, and duration analysis.

Texts

Ajmani, V.B. (2009). *Applied econometrics using the SAS system*, Hoboken, NJ: Wiley-Interscience.

[ISBN-13: 978-0470129494]

Allison, P.D. (2010). *Survival Analysis Using SAS®*: A Practical Guide (2nd Edition), Cary, N.C.: SAS Institute, Inc.

[ISBN 13: 9781599946405]

Box, G.E.P., Jenkins, G.M., Reinsel, G.C. (2008). *Time Series Analysis* (4th Edition), Hoboken, N.J.: Wiley

[ISBN 13: 9780470272848]

Greene, W.H. (2011). *Econometric Analysis (7th Edition)*, Upper Saddle River, NJ.

[ISBN-13: 978-0131395381]

Software

A low-cost, academic version of SAS is available for use in this course. Information regarding installation of this software will be provided by your instructor.

Prerequisites

PREDICT 410

Learning Goals

The goals of this course are to:

* Utilize SAS to perform analytic procedures on different types of data.
* Assess predictive analytic questions to choose the most appropriate analytic models in terms of measures of goodness of fit, coefficient interpretation, and assessment of results.
* Implement solutions to problems that result from failing to meet the assumption requirements for performing regression analysis.
* Implement appropriate solutions to model and summarize potential endogeneity.
* Address correlated error terms in modeling.
* Interpret duration analysis, censored regression, and truncated regression analysis results.
* Discuss the fundamentals of other advanced modeling techniques.

Evaluation

The student’s final grade will be determined as follows:

* Course Assignments: 450 pts.
* Final Exam: 350 pts.
* Discussion Board Participation: 200 pts.

Total course points = 1000 points

Grading Scale

A if total course points >=950

1. if 900<= total course points < 950

B+ if 870<=total course points<900

B if 830<=total course points<870

1. if 800<=total course points<830

C+ if 770<=total course points<800

C if 730<=total course points<770

1. if 700 <= total course points < 730

D if 600<=total course points<700

F if total course points < 600

Discussion Board Etiquette

The purpose of the discussion boards is to allow students to freely exchange ideas. It is imperative to remain respectful of all viewpoints and positions and, when necessary, agree to respectfully disagree. While active and frequent participation is encouraged, cluttering a discussion board with inappropriate, irrelevant, or insignificant material will not earn additional points and may result in receiving less than full credit. Frequency is not unimportant, but content of the message is paramount. Please remember to cite all sources—when relevant—in order to avoid plagiarism.

Proctored Assessment

There is no proctored assessment requirement for this class.

Attendance

This course will not meet at a particular time each week. All course goals, session learning objectives, and assessments are supported through classroom elements that can be accessed at any time. To measure class participation (or attendance), your participation in threaded discussion boards is required, graded, and paramount to your success in this class. Please note that any scheduled synchronous or “live” meetings are considered supplemental and optional. While your attendance is highly encouraged, it is not required and you will not be graded on your attendance or participation.

Late Work

Students must provide written notification of late work 24 hours prior to the deadline. One grace day is allowed for those who provide late work notification. Only one grace day without reduction of points is allowed. A 25% reduction is applied to the grade for every 12 hours late. No negative points are applied.

Learning Groups

Learning groups are utilized in this course. More information about learning groups will be provided by the instructor via the Blackboard course site.

Academic Integrity at Northwestern

Students are required to comply with University regulations regarding academic integrity. If you are in doubt about what constitutes academic dishonesty, speak with your instructor or graduate coordinator before the assignment is due and/or examine the University Web site. Academic dishonesty includes, but is not limited to, cheating on an exam, obtaining an unfair advantage, and plagiarism (e.g., using material from readings without citing or copying another student's paper). Failure to maintain academic integrity will result in a grade sanction, possibly as severe as failing and being required to retake the course, and could lead to a suspension or expulsion from the program. Further penalties may apply. For more information, visit <www.scs.northwestern.edu/student/issues/academic\_integrity.cfm>.

Plagiarism is one form of academic dishonesty. Students can familiarize themselves with the definition and examples of plagiarism, by visiting <www.northwestern.edu/uacc/plagiar.html>. A myriad of other sources can be found online.

Some assignments in this course may be required to be submitted through SafeAssign, a plagiarism detection and education tool. You can find an explanation of the tool at <http://wiki.safeassign.com/display/SAFE/How+Does+SafeAssign+Work>. In brief, SafeAssign compares the submitted assignment to millions of documents in large databases. It then generates a report showing the extent to which text within a paper is similar to pre-existing sources. The user can see how or whether the flagged text is appropriately cited. SafeAssign also returns a percentage score, indicating the percentage of the submitted paper that is similar or identical to pre-existing sources. High scores are not necessarily bad, nor do they necessarily indicate plagiarism, since the score does not take into account how or whether material is cited. If a paper consisted of one long quote that was cited appropriately, it would score 100%. This would not be plagiarism, due to the appropriate citation. However, submitting one long quote would probably be a poor paper. Low scores are not necessarily good, nor do they necessarily indicate a lack of plagiarism. If a 50-page paper contained all original material, except for one short quote that was not cited, it might score around 1%. But, not citing a quotation is still plagiarism.

SafeAssign includes an option in which the student can submit a paper and see the resultant report before submitting a final copy to the instructor. This ideally will help students better understand and avoid plagiarism.

Other Processes and Policies

Please refer to your SCS student handbook at <www.scs.northwestern.edu/grad/information/handbook.cfm> for additional course and program processes and policies.

Course Schedule

***Important Note:*** Changes may occur to the syllabus at the instructor's discretion.  
When changes are made, students will be notified via an announcement in Blackboard.

Session 1

Learning Objectives

After this session, the student will be able to:

* Explain simple regression concepts.
* Perform simple regression analysis on continuous dependent variable data in SAS.
* Explain multiple regression concepts.
* Perform multiple regression analysis on continuous dependent variable data in SAS.
* Generate forecasts in SAS.
* Summarize regression results and limitations.
* Interpret regression results from SAS.
* Explain generalized linear hypothesis (GLH).
* Analyze data sets in SAS using GLH concepts.

Course Content

Textbook Reading

Ajmani, chapters 1, 2, and 3

Greene, chapters 1,2,3,4, and 5

Multimedia

Multiple Linear Regression

Discussion Board

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students' comments is graded. For this session’s discussion topic(s), visit the discussion board in Blackboard.

Assignments

Airlines Data Analysis—Multiple Linear Regression is due Sunday, October 7th, 2012 at 11:55 p.m. (central time). For more information, click Assignments on the left navigation panel in Blackboard, and scroll to this assignment’s item.

**Sync Session: October** 1st, 2012 from 7:00PM to 9:30PM CST

Session 2

Learning Objectives

After this session, the student will be able to:

* Explain differences between endogenous and exogenous variables.
* Explain the impact of measurement errors in data.
* Analyze data with known measurement errors.
* Explain heteroscedasticity in data sets.
* Analyze data with known heteroscedasticity.
* Explain autocorrelation.
* Analyze data using ARCH and GARCH concepts.

Course Content

Textbook Reading

Ajmani, chapters, chapter 4.1–4.3, 5.1–5.7, and 6

Greene, chapters 8 and 9

Discussion Board

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students’ comments is graded. For this session’s discussion topic(s), visit the discussion board in Blackboard.

Assignments

Credit Card data Analysis—Heteroscedasticity is due Sunday, October 14th, 2012 at 11:55 p.m. (central time). For more information, click Assignments on the left navigation panel in Blackboard, and scroll to this assignment’s item.

**Sync Session:** None

Session 3

Learning Objectives

After this session, the student will be able to:

* Perform multivariate regression analysis on discrete dependent variable data in SAS.
* Explain the basic elements of logistic regression.
* Explain the basic elements of Poisson regression.
* Explain the basic elements of negative binomial regression.
* Conduct logistic, Poisson, and negative binomial regressions in SAS.

Course Content

Textbook Reading

Ajmani, chapter 10

Greene, chapters 17, and 18

Online Reading

Greene, *Functional Forms for the Negative Binomial Model for Count Data*

SAS, *The LOGISTIC Procedure*

SAS, *PROC GENMOD Statement*

Multimedia

Logistic Regression

Discussion Board

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students’ comments is graded. For this session’s discussion topic(s), visit the discussion board in Blackboard.

Assignments

Unemployment Data Analysis—Logistic Regression is due Sunday, October 21st, 2012 at 11:55 p.m. (central time). For more information, click Assignments on the left navigation panel in Blackboard, and scroll to this assignment’s item.

**Sync Session:** None

Session 4

Learning Objectives

After this session, the student will be able to:

* Explain the concept of panel data.
* Explain the basic elements of pooled panel data analysis.
* Explain the basic elements of fixed effects panel data analysis.
* Explain the basic elements of random effects panel data analysis.
* Conduct panel data analysis in SAS.
* Explain dynamic panel data analysis.
* Conduct dynamic panel data analysis in SAS.
* Explain heterogeneity and autocorrelation in panel data.
* Use alternative analysis techniques in SAS to analyze panel data sets with known heterogeneity and autocorrelation issues.

Course Content

Textbook Reading

Ajmani, chapters 7 and 12.7–12.9

Greene, chapter 11

Multimedia

Panel Data Analysis

Discussion Board

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students’ comments is graded. For this session’s discussion topic(s), visit the discussion board in Blackboard.

Assignments

Airlines Data Analysis—Panel Data Analysis is due Sunday, October 28th, 2012 at 11:55 p.m. (central time). For more information, click Assignments on the left navigation panel in Blackboard, and scroll to this assignment’s item.

**Sync Session:** None

Session 5

Learning Objectives

After this session, the student will be able to:

* Differentiate between exogenous and endogenous variables.
* Explain how a system of equations works.
* Estimate a system of equations in SAS.
* Explain the seemingly unrelated regression (SUR) analysis.
* Perform SUR using two-stage-least-squares (2SLS) in SAS.

Course Content

Textbook Reading

Ajmani, chapters 4.1, 8, and 9

Greene, chapters 10, and 15

Discussion Board

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students’ comments is graded. For this session’s discussion topic(s), visit the discussion board in Blackboard.

Assignments

Grunfeld Data Analysis is due Sunday, November 4th, 2012 at 11:55 p.m. (central time). For more information, click Assignments on the left navigation panel in Blackboard, and scroll to this assignment’s item.

**Sync Session:** November 5th, 2012 from 7:00PM to 9:30PM CST

Session 6

Learning Objectives

After this session, the student will be able to:

* Explain how time series analysis is different than pooling time variant data.
* Discuss basic principles behind autocorrelation.
* Perform an autocorrelation analysis in SAS.

Course Content

**Textbook Reading**

Ajmani, Chapter 6

Greene, Chapter 20

Online Reading

SAS, *The ARIMA Procedure*

Discussion Board

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students’ comments is graded. For this session’s discussion topic(s), visit the discussion board in Blackboard.

Assignments

Gasoline Consumption Data Analysis is due Sunday, November 11th, 2012 at 11:55 p.m. (central time). For more information, click Assignments on the left navigation panel in Blackboard, and scroll to this assignment’s item.

**Sync Session:** None

Session 7

Learning Objectives

After this session, the student will be able to:

* Describe the important attributes of time series data.
* Describe how to identify an appropriate model for time series data using ACF, and PACF functions.
* Run time series analysis using SAS.
* Discuss basic MA, AR, and ARMA models.

Course Content

**Textbook Reading**

Box et. al., Chapters 1 and 2

Greene, Chapter 20

Online Reading

SAS, *The ARIMA Procedure*

Discussion Board

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students’ comments is graded. For this session’s discussion topic(s), visit the discussion board in Blackboard.

Assignments

Steel Export Data Analysis is due Sunday, November 18th, 2012 at 11:55 p.m. (central time). For more information, click Assignments on the left navigation panel in Blackboard, and scroll to this assignment’s item.

**Sync Session:** None

Session 8

Learning Objectives

After this session, the student will be able to:

* Explain the basic concepts of time series analysis.
* Explain the basic concepts of model identification.
* Explain the basic concepts of AR, MA, ARMA, ARIMA models.
* Explain model validation in time series.
* Explain forecasting in time series.
* Estimate ARIMA models in SAS.
* Validate ARIMA models in SAS.
* Use time series models to generate forecasts.

Course Content

Textbook Reading

Box et. al., Chapters 1 through 4

Greene, Chapter 20

Online Reading

SAS, *The ARIMA Procedure*

Discussion Board

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students’ comments is graded. For this session’s discussion topic(s), visit the discussion board in Blackboard.

Assignments

Wine Data Analysis is due Sunday, November 25th, 2012 at 11:55 p.m. (central time). For more information, click Assignments on the left navigation panel in Blackboard, and scroll to this assignment’s item.

**Sync Session:** None

Session 9

Learning Objectives

After this session, the student will be able to:

* Discuss duration analysis models.
* Discuss the application of duration analysis models.
* Discuss basic duration analysis procedures.
* Conduct duration analysis in SAS.

Course Content

Online Reading

Allison, Chapters 1 through 5

Discussion Board

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students’ comments is graded. For this session’s discussion topic(s), visit the discussion board in Blackboard.

Assignments

Recidivism Data Analysis is due Sunday, December 2nd, 2012 at 11:55 p.m. (central time). For more information, click Assignments on the left navigation panel in Blackboard, and scroll to this assignment’s item.

**Sync Session:** December 3nd, 2012 from 7:00PM to 9:30PM CST

Session 10

Learning Objectives

After this session, the student will be able to:

* No new learning objectives will be introduced.

Course Content

None.

Discussion Board

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students’ comments is graded. For this session’s discussion topic(s), visit the discussion board in Blackboard.

Assignments

Final exam is due Sunday, December 9th, 2012 at 11:55 p.m. (central time). For more information, click Assignments on the left navigation panel in Blackboard, and scroll to this assignment’s item.

**Sync Session**