***Excel* as a Teaching Platform for Managerial Economics:**
The Experience at *Dickinson College*

June 24, 2015

Stephen Erfle
Associate Professor and Chair
International Studies / International Business & Management
Dickinson College
Carlisle, PA 17013
erfle@dickinson.edu

One of the most important jobs of managers is to make decisions, and managers make better decisions if they employ the tools of economics in their analysis. Therefore, the goal of a managerial economics course should be to prepare the student to use the tools of economics to make decisions. Those decisions oftentimes will be based on at least some empirical information and that information is no longer typically stored on pieces of paper. Rather, empirical information (data) is typically stored in some sort of spreadsheet, the most common of which is *Excel*. In my view, managerial economics is most appropriately taught using the same toolkit that business managers rely on. Today, that means use *Excel* as the platform to teach the course.

Most managerial economics texts assume students to have had introductory microeconomics, but they do not require a background in calculus or statistics. Those elements of differential calculus and statistics that are used are explained in the body of the text. The course described in this paper was designed for students who have had introductory microeconomics, but who have not necessarily had an exposure to calculus, statistics, or *Excel*. Classes meet for three hours of lecture and one hour of computer lab per week. (I have found the 75 minute-twice a week format is preferable to the 50 minute-three times a week format due to the complexity of this material.) The lecture room must have some method for projecting computer output onto a video screen. The computer lab should have a faculty machine with projection capabilities, as well as one computer per student, although two students per computer would be adequate in all but exam situations. Students require individual access to computers in exam situations as each exam contains both an in-class and a lab component. As a result, multiple lab sections may be necessary for a single lecture section, depending on the computer lab capacity relative to size of the lecture at your school. Given the nature of the material being covered and the synergy that develops between lab and lecture, the lab is essential to the course.

None of the texts on the market in the mid-1990s were integrated with general-purpose spreadsheet programs commonly used in business[1]. Therefore, I began to develop *Excel* based lab assignments that would supplement a text. Subsequently, I began to develop *Excel* programs for use during lectures as well. I have tied these files to Mansfield's text, but most files would work well with other texts, as most texts devote significant resources to estimation techniques that form the core of these materials. These lab and lecture files are listed in Table 1 in the order in which they are used in the course[2]. I have also begun work on a more "user-friendly" set of files based on the materials discussed in this paper. These will be published as part of the ancillary package accompanying the new edition of that textbook[3].

****

Table 1 about here

****

Because students have not necessarily had an exposure to calculus, statistics, or *Excel,* these concepts must be taught from the ground up. The course is broadly divided into three parts. The first examines the concept of maximization, the second provides an introduction to statistical analysis with the goal of understanding basic regression modeling, and the third essentially combines the ideas presented in the first two parts. By the end of the semester, students are able to answer complicated but realistic questions using the tools they have at their disposal. *Excel* is used throughout, both in lab, for homework, as well as in most classes.

Lab Classes

Students are not required to have a background in *Excel*, therefore the first three labs provide a necessary introduction to using *Excel* by attacking successively more complicated

present value problems[4]. Students seem quite comfortable with the concept of present value, therefore labs can focus on how *Excel* can be used to do comparative static analysis.

I do not use *Excel's* built-in financial functions during these labs because the focus is on learning *Excel*. Students quickly learn about entering formulas, relative versus absolute cell referencing, moving between worksheets, dragging cells and setting up the general architecture of problems so that comparative static analysis is easily performed. The only advanced tool introduced in these introductory labs is *Excel's* Goal Seek function. The first lab midterm tests basic *Excel* techniques by requiring students to design a structure to answer a present value problem that is given as a word problem.

The first lab of the middle third of the semester is a <u>Statistics Lab</u> that provides students the opportunity to examine statistical data using very real information – their own scores and the scores of previous classes on the first in-class and lab midterm[5]. Students use the Paste Function menu to obtain basic statistical functions and apply these concepts to answer questions. I also ask students to predict their lab performance as a function of their in-class performance. This provides a nice introduction to univariate regression, and it allows them to explain what their residual means. Because they are explaining their own behavior, students seem particularly interested in being able to interpret this simple regression.

The next three labs introduce regression modeling. <u>RLab 1</u> estimates a multivariate demand equation. Students must grapple with dependent versus independent variables, and learn to interpret coefficient size and significance. <u>RLab 2</u> provides a cross-sectional analysis of soft drink consumption in the United States. A goal of this lab is to have students focus on residuals in order to see if patterns emerge, and then to model those patterns by creating dummy variables to test for regional differences in consumption. <u>RLab 3</u> examines airline passenger miles flown

3

over time. This lab provides the opportunity to transpose and manipulate data and to create

monthly dummy variables to search for seasonality. A goal of this lab is to examine linear and

nonlinear trends, and to do simple forecasting.

The final three labs use cost information to examine how managers make decisions to

maximize profits. RLab 4 and Rlab 5 are both based on simulated data for variable cost of

production for two different sized plants. In RLab 4, students obtain per-unit cost functions for

both plants from this information to figure out in which plant they would recommend investing,

based on various expectations regarding demand per unit of time. In RLab 5, the reverse

question is asked: suppose you own both plants and wish to produce a given amount of output at

minimum cost – how should you distribute production between plants? This lab also examines

long run versus short run issues. This is the most challenging lab in the semester. Students must

first solve a constrained minimization problem with pen and paper, and then program their

solution into *Excel* to answer the lab assignment. RLab 6 examines cost analysis in

noncompetitive markets.

Excel during Lectures

I use *Excel* sparingly in lecture during the first third of the course as I want students to

focus on the derivative concepts rather than have them focus on how to program derivatives into

Excel. *Excel* becomes a daily part of lectures once we begin to study empirical estimation.

The standard exposition of regression analysis proceeds from univariate to multivariate

regression for good reason: two-dimensional graphs are more readily understood than are higher

dimensional graphs. Students are able to visualize how an independent variable x relates to the

dependent variable y, because they see that some lines fit the data better than others. RLec 1

examines the univariate model by utilizing a small data set that Mansfield uses to make his case

for the method of least squares.[6] I tell the class to read the material on least squares, but to try to just get the big picture, rather than get bogged down in the mathematical detail because *Excel* takes care of all of those details in a matter of seconds. It is useful to initially show a graph of the data and ask what the best-fit line should look like. One can then run the regression and show that our visual best guess is mirrored by the results of the regression (both from the line fit plot, and from the estimated coefficients themselves). I hand out a copy of *Excel's* regression output so that students can take notes directly on the output regarding its various parts. I use this regression to explain the ANOVA part of the regression output that examines the concept of total variation in the dependent variable, and how this is broken down into variation explained by the regression, and unexplained variation. The coefficients are chosen so that unexplained variation is as small as possible. Finally, I relate these concepts to R^2.

In the next lecture, I move to multiple regression and discuss the interpretation of coefficients for each independent variable as simply slopes in different directions. With two independent variables, we are trying to find the best-fit plane, and *Excel* can do this just as easily as it found the best-fit line in the univariate case. From here students seem able to make the leap to how a set of n independent variables x_1, …, x_n relate to the dependent variable y without having to try to visualize the best fitting n-dimensional hyperplane. RLec 2 provides the results of such a multivariate analysis and allows a preview of RLab 1. RLec 2 also allows further discussion of other parts of the regression output such as the standard error of coefficients, and standard error of estimate (unfortunately, both are titled "Standard Error" in *Excel's* output, so the difference must be explicitly pointed out to avoid confusion).

Simulations using *Excel* allow me to examine an array of issues that must be understood in order to correctly use regression analysis. The simulated data in RLec 3a examines the

difference between magnitude and significance. This allows for an extended discussion of what we mean when we say a coefficient is significantly different from zero, as well as how we might interpret the 95% confidence interval for the coefficient. I also begin an extended discussion of residual analysis in this lecture with two topics: heteroscedasticity using RLec3a; and nonlinearities using RLec 3b.

In the fourth regression lecture, I describe a variety of situations in which dummy variables can be used to help explain the dependent variable before going to *Excel*. The example I use in RLec 4a is based on bond sales in the time period surrounding WWII[7]. I also examine multicollinearity during this lecture, a topic I found difficult to explain until I created a simulation model, RLec 4b, which shows the effect of multicollinearity using peanut butter and jelly sandwiches. I posit that students like their PB&J's in the following ratios: 1-ounce peanut butter, 1-ounce jelly, and two slices of bread. People aren't perfect at keeping to this ratio, some people might have a bit too much peanut butter, and others might have a bit too much jelly, but overall, this is what we would expect. Suppose we wish to predict bread consumption on the basis of peanut butter and jelly consumption in the dining hall. If we run a regression of bread consumption as a function of both peanut butter and jelly consumption, we have a multicollinearity problem. The model allows you to introduce varying amounts of noise between peanut butter, jelly, and bread. By varying the amount of error between peanut butter and jelly, one sees the effect of multicollinearity: peanut butter's and jelly's coefficients and significance levels vary wildly as you would expect given the multicollinearity problem. Nonetheless, the sum of peanut butter's and jelly's coefficients remains constant at 2 (you need two slices of bread for each sandwich).

The next class, on serial correlation, tends to be the most difficult of the semester. It is imperative that students understand serial correlation, even though many of the techniques used to correct for it are beyond the scope of this course. The discussion is made easier to understand using RLec 5a, a model that simulates different degrees of positive or negative serial correlation. I show serial correlation visually using both a standard residual plot, and by showing a scatter plot of residuals vs. lagged residuals. I calculate a Durbin-Watson statistic from these residuals, and work through how to apply the Durbin-Watson test. One of *Excel's* shortcomings is that a Durbin-Watson statistic is not automatically calculated among the regression statistics, so one of the tasks that students must learn is how to calculate it from the residuals that are provided within *Excel*. I show students how to do this calculation using both this simulated dataset, and with RLec 5b that reexamines the bond data introduced in RLec 4a. RLec 5b is instructive because there is strong positive serial correlation when the dummy variable is not included, and it vanishes upon introduction of the dummy variable. This provides a nice lesson for why serial correlation sometimes exists – specification error that results from omitted variables.

The final two lectures in this section of the course examine linear and nonlinear trend analysis, and seasonality. The growth rate simulation, RLec 6a, compares linear, quadratic, and logarithmic estimates based on a constant growth rate dataset. The clear patterns that emerge from the residuals in the first two specifications suggest that these models are misspecified. By contrast, the logarithmic model is the correct specification if the data has a constant rate of growth; the lack of pattern from this model's residuals therefore comes as no surprise. The seasonality simulation, RLec 6b, examines three quarterly datasets, the first with linear trend and no seasonality, the second with linear trend and seasonality, and the third with nonlinear trend and seasonality. Inevitably, students seem hesitant to use logarithms, so it is important to

carefully explain the differences that arise in interpreting coefficients of time as well as the quarterly dummy variables when using ln(Q) as opposed to Q as the dependent variable. The second lecture on this topic begins with RLec 7a, a file that reexamines the airline lab, RLab 3, because it allows an extended discussion of compounding and seasonality as well how to create seasonal indices from regression output. In RLec 7b, I begin by using seasonal indices for soft drink consumption presented in Mansfield and show how these could have been obtained from a regression model similar to the airline model, and conclude by using the airline results from RLab 3 to create seasonal indices for the airline industry.

The final part of the course combines the tools learned in the first two parts of the course. The central focus is cost analysis, but production functions and market structures round out the semester's discussion. To understand costs, it is useful to begin by discussing production theory. The Production Functions file examines returns to scale, relative factor intensity, and cost minimization.

Using simulated data based on a variable cost function used early in the semester, I work through the rationale for the standard cubic form for variable cost in Cost Estimation I. Residual plots for lower order forms exhibit patterns that are eliminated with the cubic form. Similarly, multicollinearity issues may be examined if the form is actually quadratic, but is estimated as cubic. The estimated coefficients of the variable cost function form the basis for the per-unit cost curves used for decision-making. I recommend a format within *Excel* for creating per-unit cost curves that minimizes algebraic errors and allows a visual check of the curves in Cost Estimation II.

Estimated cost functions form the basis for many of the topics examined at the end of the semester. The long run/short run distinction is easily put into perspective, as are issues such as

when the firm should shut down, and profit contribution analysis. Profit maximization is examined using cost and demand information for competitive, monopolistically competitive, and monopolistic market structures but time constraints preclude discussion of oligopoly markets.

Concluding Remarks

The course described omits a number of topics covered in many managerial economics texts. Instead, it focuses on two of the most important tools in the managerial economist's arsenal: constrained optimization and econometric estimation. Since our students are not required to take calculus or statistics these concepts must be taught from the ground up. Similarly, our students are not required to know *Excel* so it must also be taught from scratch. One might question whether the additional burden imposed by learning *Excel* is worth the effort, given the above focus. My sense is that, although there is an initial cost imposed in learning *Excel*, the value added far outweighs the cost. The regression material would be much more difficult to teach were it not for the ease with which *Excel* allows the student to focus on the various issues involved in running a regression as well as interpreting the results of that regression. *Excel* acts in an enabling rather than a limiting capacity in this course.

Given the constraint imposed by that lack of mathematical and statistical preparation of students in this class, it is more important to cover these basic topics carefully, rather than cover a wider array of topics at a more hurried pace. My sense from the five times I have taught this course is that this tradeoff is very reasonable. I covered more pages of the text in my first attempt at this course, but the current course teaches much more by focusing on fewer topics while giving students time to reflect on what they have learned.

If I were teaching managerial economics at a school that had formal calculus and statistics prerequisites, I would pursue a more broad-based strategy. In these circumstances, *Excel* would remain my choice as an analytical platform due to its prominence and versatility.

Bibliography

Mansfield, Edwin. (1999). <u>Managerial economics: theory, applications, and cases</u> (4th ed.). New York: W. W. Norton & Company, Inc.

Wonnacott, Ronald J. and Wonnacott, Thomas H. (1970). <u>Econometrics</u>. New York: John Wiley & Sons, Inc.

Table 1: files available at *http://www.dickinson.edu/~erfle/managerialexcelfilefolder.html*.	
Excel Lab Exercises*	*Excel* Programs for Lectures
• Introduction to *Excel* 1 & 2 (Showboat) • Intro. Lab 3 (should you go to college?) • Statistics Lab (how did you do on midterm 1?) • Regression Lab (RLab) 1(estimating demand) • RLab 2 (cross-sectional modeling) • RLab 3 (time series modeling) • RLab 4 (cost analysis I) • RLab 5 (cost analysis II) • RLab 6 (profit maximization)	• Regression Lecture (RLec) 1 (univariate analysis) • RLec 2 (multivariate analysis) • RLec 3a (significance vs. magnitude) • RLec 3b (introduction to residual analysis) • RLec 4a (dummy variables) • RLec 4b (multicollinearity) • RLec 5a (serial correlation I) • RLec 5b (serial correlation II) • RLec 6a (trend analysis: growth rates) • RLec 6b (seasonality I) • RLec 7a (seasonality II) • RLec 7b (calculating seasonal indicies) • Production Functions • Cost Estimation I • Cost Estimation II
* Lines show placement of labs relative to lectures (eg. Regression Lab 1 falls between Regression Lecture 2 and 3)	

[1] Managerial textbooks have subsequently begun to show up which provide some integration, for example, Edwin Mansfield's 4th edition text <u>Managerial economics: theory, applications, and cases</u> has a "spreadsheet exercises" supplement using *Excel*.

[2] Copies of these *Excel* programs and labs together with a class syllabus, student guide, and day-by-day outline, are available from my website: *http://www.dickinson.edu/~erfle/managerialexcelfilefolder.html*.

[3] W. Bruce Allen, Neil Doherty, Keith Weigelt, and Edwin Mansfield, <u>Managerial Economics</u> (Fifth Edition), Norton, New York, 2002. The new Excel supplement to the textbook will be comprised of 40 files with at least one file for each chapter of the text. Most of the files on my website will be included in modified form, but there will be approximately 20 new files as well. These files are being annotated with notes, comments, questions, and answers and they are being write protected so that students can work directly with these files. Alternatively, faculty members could assign a file as a homework assignment and have the student fill in answers. These files will be released with the textbook, early in 2002.

[4] The first lab provides a break-even analysis of an investment project discussed in the text, the revival of the Broadway musical *Showboat*. The second lab extends *Showboat* by examining the more realistic case of facing a probability distribution of when shutdown will occur. The third lab provides students with wage and college cost assumptions and asks them to provide their own architecture to answer the question: is college a good deal on a present value basis?

[5] Underlined text refers to *Excel* files noted in Table 1 and available at the website cited in Endnote 2.

[6] Most other texts have similar sample datasets that are used for the univariate exposition. I find it helpful to use the same dataset in class as in the text as this material is very difficult to grasp given all the summation signs at this point in the text.

[7] This example is borrowed from Wannacott and Wonnacott, *Econometrics* pp. 68-72 because it provides such a crisp example of a situation that requires the use of a dummy variable. When I plot the data and ask if anyone can explain what they see, inevitably someone will say that this is three separate datasets, one pre-WWII, WWII, and post WWII. They also see that WWII differs from the other two, so creating a dummy variable called "War" takes care the problem that exists when you regress bond sales as a function of national income.