More about Final Project
Econ 29000, Principles of Statistics
Kevin R Foster, CCNY
Spring 2011

Default Guidelines for Final Project

You don't have to follow this format – use your own if you have a better idea. If you don't, though, this will serve as a basis. Also please ensure that your project does not go through and give bullet points in response to each question! You should write a narrative that gracefully includes the answers to these questions. The quality of the writing is a large factor determining the grade you get. There are writing tutors available – use them!

Introduction

A concise description of the project: include the dataset used, the key interesting results (don't reproduce everything), and why those results are interesting. Should be about a page so every word must count!

Literature Review

Describe the papers you've read that also look at this topic. Explain the differences among the results found in different previous studies. You can point out challenges that remain (even if your project doesn't solve them all). Do different authors come to different conclusions? Why might this be? Are their regressions valid (e.g. do they take adequate account of endogeneity issues)? These should be academic papers – serious studies not newspaper accounts. You can cite a newspaper to indicate why the result is interesting (e.g. to show that policymakers or the public cares about knowing the real answer, or to give some background on why you're interested in it) but you can't end there.

Means (simple graphs, correlations, differences of means)

First carefully note the dataset you're using, both the original source and any subsequent restrictions (e.g. if you're only looking at children or only those who are working or whatever). Present a table where each important variable in your regression has its mean and standard deviation as well as any other relevant summary statistics (min/max, median, whatever). Verify that the units all make sense.

This is a good place for simple graphs of the sort that we talked about. Does a two-dimensional scatterplot show your regression results? Why or why not? This is also a good place to discuss functional forms: does the graph show that squared or cubic terms could be useful (or logarithms)? What about subgroups? Medians? (Look over past homework assignments for examples.)

Simple Regressions

Present a few different models in easy-to-read tables. Don't just cut-and-paste the SPSS output! That is unacceptable.

Complicated Regressions

Present some more regressions (again, in easy-to-read tables). Show your main conclusion then do some robustness checks (i.e. what if the sample were limited to only males or females or only those of certain ages or whatever is relevant). Go back to the homework assignments from class and do just those sorts of regressions; for example if you have age plus its square and cube, do the results (the coefficients on the variables of interest) change when you put in 5-year age dummies?

Explain Results

Clearly state what you have found and why it is interesting. Do your results confirm what other researchers have found? Or do they contradict earlier research? Why might this be?

Hand in: Paper, dataset, SPSS output

Don't Plagiarize! Remind yourself of the rules for academic honesty (many many previous references are available). The consequences for violations are substantial – up to expulsion.

Example

Using 2010 CPS data, restrict to only fulltime workers with a non-zero wage. Run two sets of regressions to explain earnings: with earnings (annual wage and salary) as the dependent variable; with log of earnings as the dependent.

The first set of basic explanatory variables is hypothesized to be factors such as age, sex, education, race/ethnicity, marital status, veteran status, and if a union member.

Wage/Salary (annual)	$	49,773.79
Age		41.88
Female		44.5%
White		79.7%
African-American		11.8%
Asian-American		5.8%
Native American/ Indian/ Alaskan/ Inuit/ Hawaiian		2.8%
Hispanic		16.1%
Mexican		9.8%
Puerto Rican		1.4%
Cuban		0.6%
Immigrant		17.5%
1 or more Parents were immigrants		23.8%
Education: no high school		8.6%
Education: High School Diploma		28.9%
Education: Some College (incl no degree or Assoc degree)		27.9%
Education: Some College but no degree		17.5%
Education: Associate in vocational		5.0%
Education: Associate in academic		5.4%
Education: 4-yr degree		22.5%
Education: Advanced Degree		12.1%
Married		62.0%
Divorced or Widowed or Separated		14.8%
Unmarried		23.2%
Union member		2.2%
Veteran (any)		7.4%

The regression estimates are made with three basic specifications: Spec 1 has just the listed variables; Spec 2 included dummies for industry, occupation, and state of residence; Spec 3 has dummy interactions for female*age, African-American*age, female*African-American*age, Hispanic*age, female*Hispanic*age, and female*education.

Spec 1		Spec 2	Spec 3
	estimated value	estimated value	estimated value

intercept	-$28,685.56 *	$13,744.52 *	-$10,978.43 *
	1954.106	*3025.180*	*3685.959*
Age	$2,517.92 *	$2,012.04 *	$3,052.09 *
	93.814	*88.514*	*133.158*
Age-squared	-$23.60 *	-$18.55 *	-$29.40 *
	1.055	*.994*	*1.504*
Female	-$17,380.74 *	-$14,587.20 *	$26,912.27 *
	360.019	*393.294*	*4202.955*
African American	-$6,136.77 *	-$5,315.62 *	$17,924.27 *
	552.138	*545.564*	*7559.610*
Asian	-$783.89	-$3,140.09 *	-$3,196.33 *
	861.879	*851.007*	*849.324*
Native American Indian or Alaskan or Hawaiian	-$4,615.72 *	-$3,077.92 *	-$3,030.05 *
	1054.697	*1025.422*	*1022.749*
Hispanic	-$5,176.56 *	-$4,433.05 *	$32,492.36 *
	596.068	*588.188*	*5715.141*
Immigrant	-$7,377.88 *	-$4,669.63 *	-$4,080.20 *
	776.395	*731.493*	*733.482*
1 or more parents were immigrants	$4,513.48 *	$1,231.87	$892.78
	718.087	*677.532*	*677.771*
Education: High School Diploma	$7,658.27 *	$3,819.68 *	$4,208.53 *
	701.918	*667.305*	*826.691*
Education: Some College but no degree	$15,430.94 *	$7,791.73 *	$9,434.14 *
	756.430	*734.022*	*900.898*
Education: Associate in vocational	$15,719.42 *	$8,376.06 *	$9,873.19 *
	1003.190	*966.454*	*1098.448*
Education: Associate in academic	$19,907.99 *	$9,660.31 *	$11,310.63 *
	978.304	*948.764*	*1091.644*
Education: 4-yr degree	$35,565.50 *	$20,756.84 *	$24,651.87 *
	738.325	*761.377*	*949.760*
Education: Advanced Degree	$63,729.94 *	$40,911.95 *	$46,708.57 *
	815.818	*896.308*	*1109.431*
Married	$8,100.77 *	$7,074.38 *	$6,912.90 *
	486.083	*459.856*	*459.565*
Divorced or Widowed or Separated	$1,646.98 *	$1,893.12 *	$1,881.97 *
	633.993	*595.046*	*594.911*
Union member	-$3,992.75 *	$2,282.96 *	$2,372.64 *
	1169.615	*1108.181*	*1105.552*
Veteran (any)	-$1,186.63	-$884.41	-$905.22
	687.786	*648.453*	*659.002*

R-squared	0.213	0.315	0.319

Discussion….