Department of Economics
University of California
Berkeley

Problem Set #1 Suggested solutions

Due: in lecture, Thursday, February 5, no later than 11:10 a.m.

Bring your problem set to lecture and turn it in before class begins. Problem sets are late after 11:10. Late problem sets lose 5 points. No problem sets accepted after 3 p.m. Friday February 6.

Use the regression results summarized in the table below to answer the questions. **Write your answers on this sheet.** (Neat hand writing is fine.) Attach your work for #3.

In "City Size and Ethnic Discrimination: Michigan Agricultural Implements and Iron Working Industries, 1890" (*J. of Economic History* 42 (Dec. 1984): 825-845; http://www.jstor.org/stable/2121110) Joan Hannon asks whether incomes of immigrants and the children of immigrants were lower than those of native-born workers. Her data set covers Michigan in 1890 and distinguishes between large cities (Grand Rapids and Detroit) and small cities. The average annual earnings was $487.02 in large cities and $476.42 in small cities. You are not required nor expected to do so, but if you would like to download the data set and play with it in Stata yourself, it is available at http://eh.net/database/historical-labor-statistics-project-series/, mi08a and mi08b (click on data & codebook zip files at bottom of the webpage). Table 2 summarizes Hannon's findings.

TABLE 2
MICHIGAN AGRICULTURAL IMPLEMENTS AND IRON WORKING INDUSTRIES, 1890: ESTIMATED REDUCED FORM EARNINGS EQUATIONS
(Standard Errors in Parentheses)

Independent Variable	*Dependent Variable = Ln (Y)*		
	Total Sample	*Grand Rapids and Detroit*	*Small Cities*
Constant	3.9795[a]	4.0243[a]	3.8135[a]
	(.1237)	(.1459)	(.2256)
N	−.5184[a]	−1.1731[a]	.4152
	(.1831)	(.2472)	(.3008)
P	1.0529[a]	1.0282[b]	.4331[c]
	(.1824)	(.2884)	(.2486)
MARRIED	.0359	.0059	.0887[b]
	(.0264)	(.0350)	(.0377)
AGEMIG	.1065[a]	.1046[a]	.1140[a]
	(.0075)	(.0090)	(.0134)
YRSUS	.1157[a]	.1146[a]	.1234[a]
	(.0077)	(.0091)	(.1403)
AGE	.1501[a]	.1920[a]	.1020[a]
	(.0096)	(.0150)	(.0133)
PAGE	−.0668[a]	−.0635[a]	−.0315[b]
	(.0115)	(.0189)	(.0150)
AGE^2	−.1300[a]	−.1264[a]	−.1421[a]
	(.0102)	(.0120)	(.0185)
$NAGE^2$	−.0523[a]	−.1152[a]	.0205
	(.0173)	(.0273)	(.0256)
$PAGE^2$	.0892[a]	.0990[a]	.0410[b]
	(.0168)	(.0292)	(.0210)
R^2	.3467	.4386	.2929
$\bar{R}^2$	.3437	.4338	.2850
NOBS	2210	1178	1032

[a] Significant at less than 1 percent level.
[b] Significant at less than 5 percent level.
[c] Significant at less than 10 percent level.

Where each observation is a different person; NOBS is the number of observations; N (nativity) = 1 if the person is native born, 0 if foreign born; P (parents' nativity) = 1 if parents are native born, 0 if foreign born; MARRIED = 1 if married, 0 if single; AGEMIG = age at migration (=AGE-YRSUS) if foreign born, 0 if native born; YRSUS = Years in US if foreign born, 0 if native born; AGE = age if native born, 0 if foreign born; PAGE = P * AGE; $AGE^2 = Age^2/100$; $NAGE^2 = N * AGE^2$; $PAGE^2 = P * AGE^2$.

(2 points for each question)

1. Go through the table on the other side and use asterisks to indicate which variables are statistically significant at the 95% level and at the 99% level. In the space below, list the variables that are statistically significant in the Large Cities sample but not in the Small Cities sample, or vice versa.

For asterisks: The superscript 'a' should be ***, superscript 'b' replaced with **, and superscript 'c' with *.

Statistically significant in Large Cities (Detroit & Grand Rapids) but not Small Cities:
N (native born)
N*AGE^2 (nativity * age^2 / 100)
Statistically significant in Small Cities but not Large Cities (Detroit & Grand Rapids) :
Married

2. All else constant, how different are the wages of married versus single men in large cities? In small cities? Are the results statistically significant? Do the results have practical (or economic) significance, as defined in the section exercise, week of February 2? Explain.

Important: part of the info you need in order to assess practical (or economic) significance is in the prompt: "The average annual earnings was $487.02 in large cities and $476.42 in small cities."

Also important: remember here and below that the dependent variable in the regression is Ln(Y). But the questions are asking just about income (earnings, wages – three synonymous terms in this instance). So be sure to translate from ln(Y) to just Y.

Remember your log rules: $\ln(Y_M) - \ln(Y_S) = \ln(Y_M / Y_S)$. So if some value $X = \ln(Y_M) - \ln(Y_S)$, then $e^X = Y_M / Y_S$, the ratio of married men's wages to single men's wages.

Note that there's a quick math trick that you may or may not already know. $e^{Small\#}$ is approximately equal to 1 + that Small #. This only works for small numbers. For instance, $e^{0.02} \approx 1.02$ but $e^4 = 54.5$, not 5. So if the difference between two logged values is 0.02, you can interpret that as a percent difference: the larger value is about 2 percent bigger than the small value. If $X = \ln(Y_M) - \ln(Y_S)$ and X is small, then the ratio $Y_M / Y_S \approx (1+X)$. Which is to say, Y_M is about X% bigger than Y_S.

Large Cities: Coefficient = 0.0059. So the ratio of married men's wages in large cities to single men's wages in large cities are $e^{0.0059}$ = 1.01. Married men make 1 percent more than single men in large cities. The result is not statistically significant (t = 0.0059 / 0.0350 < 2). The result doesn't seem to have practical significance: a 1% difference is pretty small.

Small Cities: Coefficient = 0.0887. So the ratio of married men's wages in small cities to single men's wages in large cities are $e^{0.0887}$ = 1.09. Married men make 9 percent more per year than do single men in small cities. The result is statistically significant (t = 0.0887 / 0.0377 > 2). The result has some practical significance: a 9% difference is worth talking about.

3. Hannon uses a number of variables to measure age. Compare two individuals: one is native born of native-born parents, the other is foreign born and migrated to the U.S. at age 18. They are alike in all other regards. Use the regression results to calculate the values for the table below. Attach your work or no points.

	Difference in Income between Native Born (of Native Parents) and Foreign Born	
Age	Large Cities	Small Cities
20	28% higher	23% higher
30	36% higher	2% lower

Be careful to read the notes to the table so that you know what values to enter.

Large Cities, age 20, Native Born of Native Parents:
4.0243-1.1731*1 + 1.0282*1 + 0.0059*X + 0.1046*0 + 0.1146*0 + 0.1920*20 - 0.0635*(1*20) - 0.1264*(20^2/100) - 0.1152*(1 * 20^2/100) + 0.0990*(1 * 20^2/100) = 5.8790 + 0.0059*X (from the prompt, we don't know whether the person is married or not but that's not going to be problematic)
Large Cities, age 20, Foreign Born:
4.0243-1.1731*0 + 1.0282*0 + 0.0059*X + 0.1046*18 + 0.1146*(20-18) + 0.1920*0 - 0.0635*(1*0) - 0.1264*(20^2/100) - 0.1152*(0 * 20^2/100) + 0.0990*(0 * 20^2/100) = 5.6307 + 0.0059*X (from the prompt, we don't know whether the person is married or not but that's not going to be problematic)

Difference in logs, Large Cities, age 20, Native Born of Native Parents minus Foreign born

5.8790 + 0.0059*X - (5.6307 + 0.0059*X) = 5.8790 - 5.6307 = 0.2483

Ratio of wages, Large Cities, age 20, Native Born of Native Parents relative to Foreign born

$e^{0.2483}$ = 1.28. 20 year old native born men of native parents have 28% higher wages than 20 year old foreign born men.

Large Cities, age 30, Native Born of Native Parents:

4.0243-1.1731*1 + 1.0282*1 + 0.0059*X + 0.1046*0 + 0.1146*0 + 0.1920*30 - 0.0635*(1*30) - 0.1264*(30^2/100) - 0.1152*(1 * 30^2/100) + 0.0990*(1 * 30^2/100) = 6.4510 + 0.0059*X (from the prompt, we don't know whether the person is married or not but that's not going to be problematic)

Large Cities, age 30, Foreign Born:

4.0243-1.1731*0 + 1.0282*0 + 0.0059*X + 0.1046*18 + 0.1146*(30-18) + 0.1920*0 - 0.0635*(1*0) - 0.1264*(30^2/100) - 0.1152*(0 * 30^2/100) + 0.0990*(0 * 30^2/100) = 6.1447 + 0.0059*X (from the prompt, we don't know whether the person is married or not but that's not going to be problematic)

Difference in logs, Large Cities, age 30, Native Born of Native Parents minus Foreign born

6.4510 + 0.0059*X - (6.1447 + 0.0059*X) = 6.4510 - 6.1447 = 0.3063

Ratio of wages, Large Cities, age 30, Native Born of Native Parents relative to Foreign born

$e^{0.3063}$ =1.36. 30 year old native born men of native parents have 36% higher wages than 30 year old foreign born men.

Small Cities, age 20, Native Born of Native Parents:

3.8135+0.4152*1 + 0.4331*1 + 0.0887*X + 0.1140*0 + 0.1234*0 + 0.1020*20 - 0.0315*(1*20) - 0.1421*(20^2/100) + 0.0205*(1 * 20^2/100) + 0.0410*(1 * 20^2/100) = 5.7494 + 0.0887*X (from the prompt, we don't know whether the person is married or not but that's not going to be problematic)

Small Cities, age 20, Foreign Born:

3.8135+0.4152*0 + 0.4331*0 + 0.0887*X + 0.1140*18 + 0.1234*(20-18) + 0.1020*0 - 0.0315*(0*20) - 0.1421*(20^2/100) + 0.0205*(0 * 20^2/100) + 0.0410*(0 * 20^2/100) = 5.5439 + 0.0887*X (from the prompt, we don't know whether the person is married or not but that's not going to be problematic)

Difference in logs, Small Cities, age 20, Native Born of Native Parents minus Foreign born

5.7494 + 0.0887*X - (5.5439 + 0.0887*X) = 5.7494 - 5.5439 = 0.2055

Ratio of wages, Small Cities, age 20, Native Born of Native Parents relative to Foreign born

$e^{0.2055}$ = 1.23. 20 year old native born men of native parents have 23% higher wages than 20 year old foreign born men.

Small Cities, age 30, Native Born of Native Parents:

3.8135+0.4152*1 + 0.4331*1 + 0.0887*X + 0.1140*0 + 0.1234*0 + 0.1020*30 - 0.0315*(1*30) - 0.1421*(30^2/100) + 0.0205*(1 * 30^2/100) + 0.0410*(1 * 30^2/100) = 6.0514 + 0.0887*X (from the prompt, we don't know whether the person is married or not but that's not going to be problematic)

Small Cities, age 30, Foreign Born:

3.8135+0.4152*0 + 0.4331*0 + 0.0887*X + 0.1140*18 + 0.1234*(30-18) + 0.1020*0 - 0.0315*(0*30) - 0.1421*(30^2/100) + 0.0205*(0 * 30^2/100) + 0.0410*(0 * 30^2/100) = 6.0674 + 0.0887*X (from the prompt, we don't know whether the person is married or not but that's not going to be problematic)

Difference in logs, Small Cities, age 30, Native Born of Native Parents minus Foreign born

6.0514 + 0.0887*X - (6.0674 + 0.0887*X) = 6.0514 - 6.0674 = -0.0160

Ratio of wages, Small Cities, age 30, Native Born of Native Parents relative to Foreign born

$e^{-0.016}$ = 0.98. 20 year old native born men of native parents have 2% lower wages than 20 year old foreign born men.

4. Native born men of native born parents earn more than native born men of foreign born parents in large cities but not in small cities. Based on Hannon's results, what factors could explain the difference in income?

Hannon's results consider very few factors: dummy variables for nativity and parentage, and age. The result is not true at every age: at age 20, NB men earned more than FB men in both large and small cities. The difference between NB and FB disappears once we reach age 30. So the only variables left from the analysis to explain the result are nativity (foreign or native born) and parentage.

But from the results, we don't know why nativity and parentage matter. We would need additional analysis – perhaps empirical, perhaps not – to develop and test some hypotheses that explain the result.

5. If you were to test differences in incomes of native born and foreign born men in 1890 Michigan, what two additional variables would you want to include? What effect would you expect each of those variables to have?

Answers here will vary. You needed to come up with two variables that a reasonable person might think would explain differences in income between native born and foreign born men in 1890 Michigan. Some obvious choices would be previous occupation, previous amount of work experience (generally, or specific to a particular position), skill level.