

PROBLEM SET 2

due April 4, 2003

The following assignment is based on an Australian data set of young workers taken from the 1985 wave of the Australian Longitudinal Survey. The data have been used, among others, by Rummery, S., F. Vella and M. Verbeek (1999), “Estimating the Returns to Education for Australian Youth via Rank-Order Instrumental Variables”, *Labour Economics*, Vol. 6, 491-507. A codebook is provided at the end of the assignment. The dataset can be downloaded from my web page as a Matlab data file, called *ausdata.mat*

1. Compute weekly earnings using the hourly wage rate and hours worked per week. Run an OLS regression of $\log(\text{weekly earnings})$ (y) on the following individual characteristics: age (age); age squared ($age2$); years of schooling (sch); father’s education (fe); mother’s education (me); number of siblings (sib); a trade union dummy which is equal to 1 if the worker is a union member, zero otherwise ($union$); a government sector dummy which is equal to 1 if the worker is a government employee, zero otherwise (gov); a place of origin dummy which is equal to 1 if the worker was born in Australia, zero otherwise (aus); a marital status dummy which is equal to 1 if the worker is married, zero otherwise (mar). Don’t forget to include a constant term. Report the OLS parameter estimates b , the estimates $\sqrt{\widehat{Var}(b_j)}$, $\forall j = 1, \dots, K$, and the adjusted R^2 from the regression.
2. Perform the following tests, under the assumption of normality of the error terms. For each test, report the p -value:
 - (a) Test the hypothesis that $b_{sch} = 0$.
 - (b) Test the hypothesis that $b_{age2} = 0$.
 - (c) Test the hypothesis that $b_{me} = b_{fe}$.

- (d) Test the hypothesis that $(b_{union} = 0) \cap (b_{gov} = 0) \cap (b_{aus} = 0) \cap (b_{mar} = 0)$.
 - (e) Separately, test the hypotheses that $b_{union} = 0$; $b_{gov} = 0$; $b_{aus} = 0$; $b_{mar} = 0$.
 - (f) Interpret the test results.
3. Now suppose that the “true” model of weekly earnings is one in which the variable “both parents present” (*pres*) enters the set of covariates, as a proxy for the presence of role models, attitudes towards work, self-discipline, etc.
- (a) Perform a “long” regression of y on the extended set of covariates that includes *pres*: how does the estimate of the coefficient associated with years of schooling b_{sch} change? What about the coefficients associated to father’s education b_{fe} and mother’s education b_{me} ?
 - (b) In order to explain this change, run an auxiliary regression of $X_2 = [pres]$ on $X_1 = [const\ age\ age2\ sch\ fe\ me\ sib\ union\ gov\ aus\ mar]$ to obtain the auxiliary parameters Γ , and compute $b_1^{short} = b_1^{long} + \Gamma b_2^{long}$. Interpret this relationship.
 - (c) Test the hypothesis that $b_{sch}^{long} = 0$.
4. Question 20.3 in Goldberger, p. 221.
5. Question 21.3 in Goldberger, p. 231.
6. Question 21.4 in Goldberger, p. 231.

CODEBOOK

<i>WAGE</i>	Hourly wage rate
<i>AGE</i>	Age
<i>MAR</i>	Married (dummy)
<i>AUS</i>	Born in Australia (dummy)
<i>PRES</i>	Both parents present in household at age 14 (dummy)
<i>ME</i>	Mother has university degree (dummy)
<i>FE</i>	Father has university degree (dummy)
<i>SIB</i>	Number of siblings
<i>RC</i>	Roman Catholic (dummy)
<i>SCH</i>	Years of schooling
<i>GOV</i>	Government employee (dummy)
<i>UNION</i>	Union member (dummy)
<i>HOURS</i>	Hours worked per week