Economics 312 **Spring 2013**
Daily Problem #32 **April 15**

1. Suppose that we have two sets of observations on wages W_i and a row vector of other variables $\mathbf{x}_i$ that are expected to affect wages. One set is of individuals drawn from the U.S. Census of 1990 and one is drawn from 2000. They are not the same people in the two sub-samples. We regress

$$\ln(W_i) = \beta_0 + \beta_1 D_i + \mathbf{x}_i \boldsymbol{\gamma} + e_i$$

where

$$D_i = \begin{cases} 1 \text{ if the observation is from the 2000 Census} \\ 0 \text{ if the observation is from the 1990 Census.} \end{cases}$$

How would we interpret an estimated β_1 of 0.10? Why would it be important to include this term?

2. Suppose that we now want to estimate the effect of aggregate per-capita GDP in the economy on wages, so we include a term $\beta_2 \ln(GDP_i)$ in the regression. Note that $\ln(GDP_i)$ takes on only two values, one for 1990 and one for 2000. Explain why the model would suffer from perfect collinearity if both $\ln(GDP_i)$ and D_i were included.

3. If we leave out the dummy, can we estimate the effect of $\ln(GDP_i)$? Why might this estimate be unreliable?