

3 questions based on 3 datasets from Wooldridge package

Submit your answer in Word (or other word processor) as well as an R executable script.

1. Wages and productivity.

One way to consider the relationship between wage and productivity is to estimate the elasticity of hourly wage with respect to output per hour. The file *earn9* contain annual data for the non-farm business sector in the U.S. from 1947 to 1987. The variable *hrwage* is average hourly wage, and *outphr* is output per hour, one measure for labor productivity.

(a) What might be the problem if we directly run the model below? (hint: check out the time plots for these two key variables.)

$$\log(hrwage_t) = \beta_0 + \beta_1 \log(outphr_t) + u_t$$

(b) Now add a time trend to the above model and report the results. Interpret the estimated elasticity.

(c) Do you find high or low first order autocorrelation? Provide the plots for the Autocorrelation Function (ACF) of these two variables.

(d) Now re-estimate the equation in first differences (no longer need a time trend). Interpret the estimated elasticity. Is the estimated elasticity statistically different from a unity elasticity?

(e) We define the growth in hourly wage and output per hour as the change in the natural log:

$$\begin{aligned} ghrwage &= \Delta \log(hrwaeg) \\ goutphr &= \Delta \log(outphr) \end{aligned}$$

Allow an increase in productivity growth to have both a current and lagged effect on wage growth by revising the previous model to the following:

$$ghrwage_t = \beta_0 + \beta_1 goutphr_t + \beta_2 goutphr_{t-1} + u_t$$

Report the results. Is the lagged value of *goutphr* statistically significant?

(f) If $\beta_1 + \beta_2 = 1$, a permanent increase in productivity growth is fully passed on in higher wage growth after one year. Test $\beta_1 + \beta_2 = 1$ as a null hypothesis and what can you conclude?

(g) Does $goutphr_{t-2}$ need to be in the model? Explain.

2. Do changes in traffic laws affect traffic fatalities?

The file *traffic2* contains 108 monthly observations on automobile accidents, traffic laws, and some other variables for California from January 1981 through December 1989. Use this data set to answer the following questions.

- (a) During what month and year did California's seat belt law take effect? When did the highway speed limit increase to 65 miles per hour?
- (b) Regress the variable $\log(\text{totacc})$ on a linear time trend and 11 monthly dummy variables, using January as the base month. Interpret the coefficient estimate on the time trend. Would you say there is seasonality in total accidents?
- (c) Add to the regression from part (b) the variables *wkends*, *unem*, *spdlaw*, and *beltlaw*. Interpret the coefficients on the unemployment variable, *spdlaw* and *beltlaw*. Do their signs and magnitude make sense to you? Are the estimated effects what you expected? Explain.
- (d) The variable *prcfat* is the percentage of accidents resulting in at least one fatality. What is the average of *prcfat* over this period? Use *prcfat* as the dependent variable in place of $\log(\text{totacc})$. Discuss the estimated effects and significance of the speed and seat belt law variables.
- (e) Compute the first order autocorrelation coefficient for the variable *prcfat*. Are you concerned that *prcfat* contains a unit root? Do the same for the unemployment rate.
- (f) Estimate a multiple regression model relating the first difference of *prcfat*, $D\text{prcfat}$, to the same variables in part (d), except you should first difference the unemployment rate, too. Then, include a linear time trend, monthly dummy variables, the weekend variable, and the two policy variables; do not difference these. Do you find any interesting results?
- (g – extra-credit) Using the standard Dickey-Fuller ($k=0$) regression, test whether *prcfat* has a unit root. Can you reject a unit root at the 5% level?

3. The effect of minimum wage.

Use data in *minwage* for this exercise. In particular, use the employment and wage series for sector 232 (Men's and Boy's Furnishings). The variable *gwage232* is the monthly growth (change in logs) in the average wage in sector 232, *gemp232* is the growth in employment in sector 232, *gmwage* is the growth of federal minimum wage, and *gcpi* is the growth in the urban Consumer Price Index. Read the data set and set it up as a time series object.

(a) Run the regression *gwage232* on *gmwage* and *gcpi*, and the regression *gemp232* on *gmwage* and *gcpi*. Interpret the estimated coefficients for *gmwage*, $\hat{\alpha}_1$ and $\hat{\beta}_1$. Does minimum wage growth appear to have a contemporaneous effect on the wage growth and on the employment growth in this sector? Is *gmwage* statistically significant in these two models?

$$\begin{aligned}gwage232_t &= \alpha_0 + \alpha_1 gmwage_t + \alpha_2 gcpi_t + u_t \\gemp232_t &= \beta_0 + \beta_1 gwage232_t + \beta_2 gcpi_t + v_t\end{aligned}$$

(b) Add lags 1 through 12 of *gmwage* to the equations in part (a). What are the estimated long-run impacts of the growth in minimum wage on the wage growth and on the employment growth in this sector?

(c) Now estimate the dynamic model

$$gwage232_t = \beta_0 + \beta_1 gwage232_{t-1} + \beta_2 gmwage_t + \beta_3 gcpi_t + u_t$$

by OLS and report both usual and HAC (heteroscedasticity and auto-correlation consistent) standard errors. Holding fixed last months' growth in wage and the growth in CPI, does an increase in the federal minimum wage result in contemporaneous increase in *gwage232*? Explain.

(d) Now add the lagged growth in employment, *gemp232_{t-1}*, to the above model (c), in order to control for the labor market condition. Does the effect of *gmwage* change much? Is it still statistically significant? Does the answer change if you use the usual OLS standard error in comparison to the use of HAC standard errors?

(e) Compared with the model without *gwage232_{t-1}* and *gemp232_{t-1}*, does adding the two lagged variables have much of an effect on the coefficient of *gmwage*? Why? (hint: consider the discussion on omitted variable bias.)

(f) Create time plots for *lwage232* and *lemp232*. Do you see clear trends there?

(g - extra-credit) Confirm that *lwage232* and *lemp232* are non-stationary. best characterized as I (1) processes. Use the augmented Dickey-Fuller test with one lag of *lwage232* and *lemp232*, respectively, and a linear time trend if needed.