Time series (1)

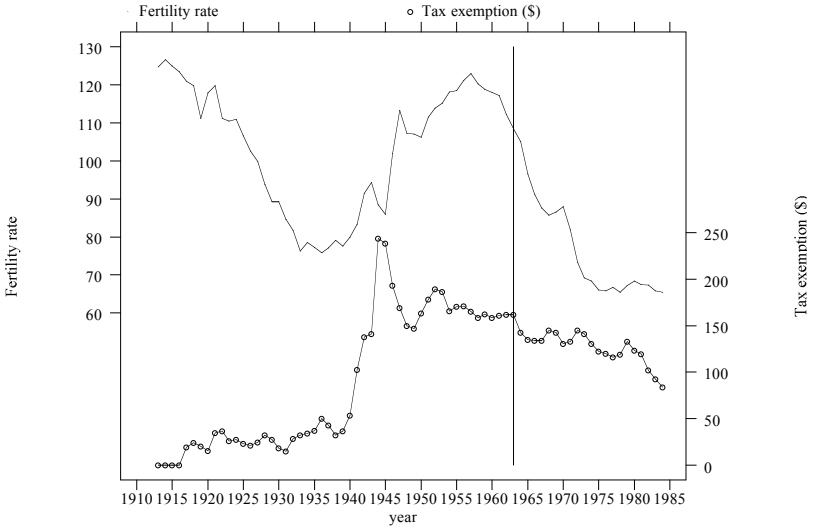
A very important source of time series data is at Economagic: http://www.economagic.com/

Source: FERTIL3, from L.A. Whittington, J. Alm, and H.E. Peters (1990), "Fertility and the Personal Exemption: Implicit Pronatalist Policy in the United States," *American Economic Review* 80, 545-556.

```
gfr     births per 1000 women 15-44
pe      real value pers. exemption, $
year    1913 to 1984
t       time trend, t=1,...,72
pill    =1 if year >= 1963
ww2      =1, 1941 to 1945
```

```
. list year gfr pe pill ww2, sep(0)
```

	year	gfr	pe	pill	ww2
1.	1913	124.7	0	0	0
2.	1914	126.6	0	0	0
3.	1915	125	0	0	0
4.	1916	123.4	0	0	0
5.	1917	121	19.27	0	0
6.	1918	119.8	23.94	0	0
7.	1919	111.2	20.07	0	0
25.	1937	77.1	42.79	0	0
26.	1938	79.1	32.22	0	0
27.	1939	77.6	36.53	0	0
28.	1940	79.9	53.33	0	0
29.	1941	83.4	102.49	0	1
30.	1942	91.5	137.7	0	1
31.	1943	94.3	141.2	0	1
32.	1944	88.4	243.83	0	1
33.	1945	85.9	238.4	0	1
34.	1946	101.9	193.16	0	0
35.	1947	113.3	168.9	0	0
49.	1961	117.2	160.71	0	0
50.	1962	112.2	161.58	0	0
51.	1963	108.5	161.61	1	0
52.	1964	105	142.73	1	0
71.	1983	65.8	92.49	1	0
72.	1984	65.4	83.9	1	0



1. Static time series regression

```
. reg gfr pe ww2 pill
```

```
      Source |       SS       df       MS              Number of obs =      72
-------------+------------------------------           F(  3,    68) =   20.38
       Model |  13183.6215     3  4394.54049           Prob > F      =  0.0000
    Residual |  14664.2739    68  215.651087           R-squared     =  0.4734
-------------+------------------------------           Adj R-squared =  0.4502
       Total |  27847.8954    71  392.223879           Root MSE      =  14.685

------------------------------------------------------------------------------
         gfr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          pe |     .08254   .0296462     2.78   0.007     .0233819    .1416981
         ww2 |   -24.2384   7.458253    -3.25   0.002    -39.12111   -9.355684
        pill |  -31.59403   4.081068    -7.74   0.000    -39.73768   -23.45039
       _cons |   98.68176   3.208129    30.76   0.000     92.28003    105.0835
------------------------------------------------------------------------------
```

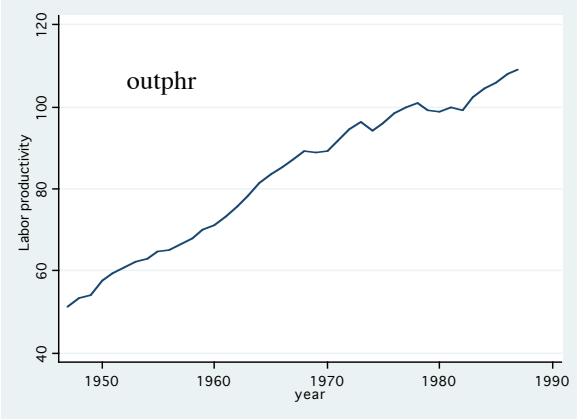
2. Regression between two variables that have a trend: spurious correlation

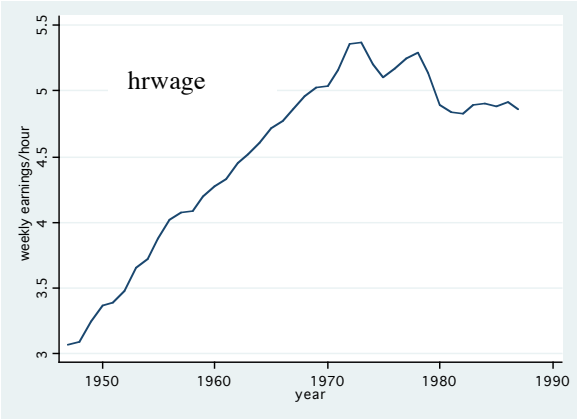
Source: EARNS in Wooldridge. *Economic Report of the President*, 1989, Table B-47. The data are for the nonfarm business sector.

```
  1. year                   1947 to 1987
  4. outphr                 output per labor hour
  5. hrwage                 average earnings / hour
```

```
. use http://fmwww.bc.edu/ec-p/data/wooldridge/EARNS
. sum
```

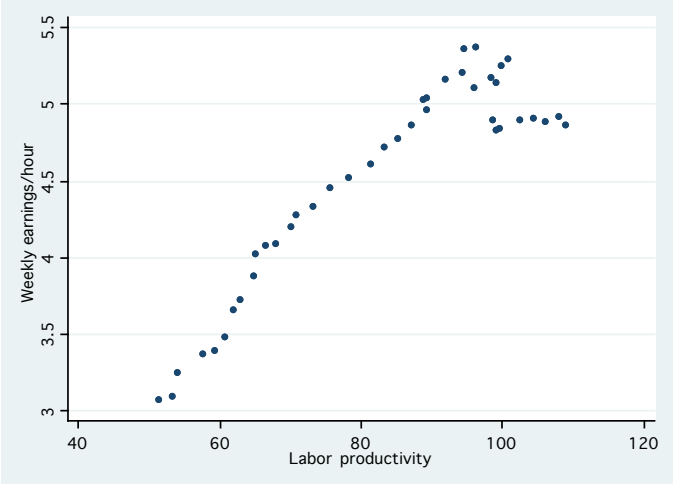
Variable	Obs	Mean	Std. Dev.	Min	Max
year	41	1967	11.97915	1947	1987
outphr	41	82.91463	17.5342	51.4	109
hrwage	41	4.509252	.6854282	3.065012	5.37533941





Looking at observations before 1972, both series have a regular trend.
If you represent a scatter diagram hourly wage against labor productivity, you may have a spurious correlation:

```
. twoway scatter hrwage outphr
```



```
. reg hrwage outph if year<=1972

      Source |       SS       df       MS              Number of obs =      26
-------------+------------------------------           F(  1,    24) = 1173.28
       Model |  11.4674843     1  11.4674843           Prob > F      =  0.0000
    Residual |  .234573166    24  .009773882           R-squared     =  0.9800
-------------+------------------------------           Adj R-squared =  0.9791
       Total |  11.7020574    25  .468082297           Root MSE      =  .09886

------------------------------------------------------------------------------
      hrwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      outphr |   .0511744    .001494    34.25   0.000     .0480909    .0542578
       _cons |    .492528   .1100935     4.47   0.000     .2653061    .7197498
------------------------------------------------------------------------------

. reg hrwage outph

      Source |       SS       df       MS              Number of obs =      41
-------------+------------------------------           F(  1,    39) =  213.87
       Model |  15.8941543     1  15.8941543           Prob > F      =  0.0000
    Residual |  2.89831975    39  .074315891           R-squared     =  0.8458
-------------+------------------------------           Adj R-squared =  0.8418
       Total |  18.7924741    40  .469811852           Root MSE      =  .27261

------------------------------------------------------------------------------
      hrwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      outphr |   .0359503   .0024582    14.62   0.000      .030978    .0409226
       _cons |   1.528445   .2082233     7.34   0.000     1.107274    1.949616
------------------------------------------------------------------------------
```

a. You could add the (omitted) time variable

```
. reg hrwage outphr t if year<=1972

      Source |       SS       df       MS              Number of obs =      26
-------------+------------------------------           F(  2,    23) = 2104.28
       Model |  11.6384528     2  5.81922639           Prob > F      =  0.0000
    Residual |  .063604645    23  .002765419           R-squared     =  0.9946
-------------+------------------------------           Adj R-squared =  0.9941
       Total |  11.7020574    25  .468082297           Root MSE      =  .05259

------------------------------------------------------------------------------
      hrwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      outphr |   -.010874   .0079313    -1.37   0.184    -.0272811    .0055331
           t |   .1079083   .0137239     7.86   0.000     .0795183    .1362983
       _cons |   3.536657   .3915594     9.03   0.000     2.726654    4.346659
------------------------------------------------------------------------------
```

b. General method is to use first difference:

```
. sort year
. g dwage=hrwage-hrwage[_n-1]
(1 missing value generated)
. g doutphr=outphr-outphr[_n-1]
(1 missing value generated)
. reg dwage doutphr if year<=1972

      Source |       SS       df       MS              Number of obs =      25
-------------+------------------------------           F(  1,    23) =    1.39
       Model |  .003598404     1  .003598404           Prob > F      =  0.2497
    Residual |  .059336242    23  .002579837           R-squared     =  0.0572
-------------+------------------------------           Adj R-squared =  0.0162
       Total |  .062934646    24  .002622277           Root MSE      =  .05079

------------------------------------------------------------------------------
       dwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     doutphr |   .0133027   .0112637     1.18   0.250     -.009998    .0366034
       _cons |   .0688566   .0219951     3.13   0.005     .0233563    .1143568
------------------------------------------------------------------------------
```

This also works for time dependences that are not necessarily a regular sustained time trend, as in our example when we consider all the years:

```
. reg dwage doutphr

      Source |       SS       df       MS              Number of obs =      40
-------------+------------------------------           F(  1,    38) =   22.09
       Model |  .124788428     1  .124788428            Prob > F      =  0.0000
    Residual |  .214659545    38  .005648935            R-squared     =  0.3676
-------------+------------------------------           Adj R-squared =  0.3510
       Total |  .339447973    39  .008703794            Root MSE      =  .07516

------------------------------------------------------------------------------
       dwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     doutphr |   .0463143    .009854     4.70   0.000      .026366    .0662626
       _cons |  -.0217087   .0185087    -1.17   0.248    -.0591776    .0157602
------------------------------------------------------------------------------
```

```
. twoway scatter dwage doutphr
```