Serial Correlation [Optional; Very brief overview]

Richard Williams, University of Notre Dame, https://www3.nd.edu/~rwilliam/
Last revised January 22, 2015

[NOTE: These notes draw very, very heavily from Pindyck & Rubinfeld's Econometric Models and Economic Forecasts (1991 edition).]

Introduction. I've never had much reason to worry about serial correlation in my work. But, if you are working with data that are collected repeatedly across time (something which I suspect is much more common in Economics than in Sociology) this may be a concern for you; or, you may read other work in which serial correlation is an issue. I'll therefore briefly discuss the problem, and you can turn to Pindyck and Rubinfeld or other sources if you need to know more.

What Serial Correlation is. When error terms from different (usually adjacent) time periods (or cross-section observations) are correlated, we say that the error term is *serially correlated*. Serial correlation occurs in time-series studies when the errors associated with a given time period carry over into future time periods. For example, if we are predicting the growth of stock dividends, an overestimate in one year is likely to lead to overestimates in succeeding years.

There are different types of serial correlation. With *first-order serial correlation*, errors in one time period are correlated directly with errors in the ensuing time period. (Errors might also be lagged, e.g. if data are collected quarterly, the errors in Fall of one year might be correlated with the errors of Fall in the next year.) With *positive serial correlation*, errors in one time period are positively correlated with errors in the next time period.

Consequences of Serial Correlation. Serial correlation will not affect the unbiasedness or consistency of OLS estimators, but it does affect their efficiency. With positive serial correlation, the OLS estimates of the standard errors will be *smaller* than the true standard errors. This will lead to the conclusion that the parameter estimates are more precise than they really are. There will be a tendency to reject the null hypothesis when it should not be rejected. This is illustrated by the following graphs:
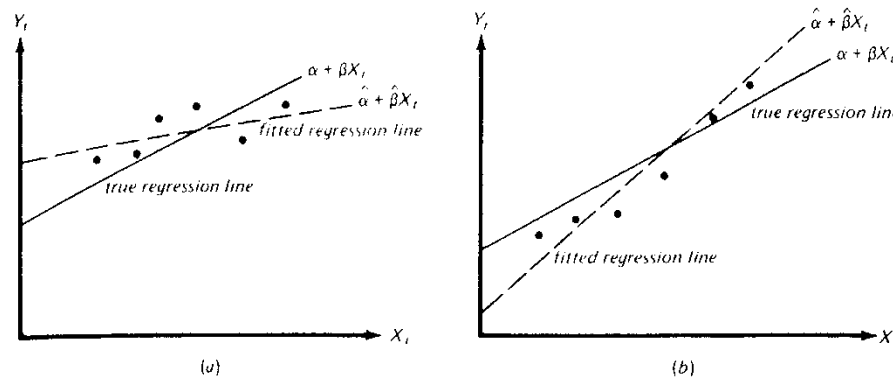

FIGURE 6.1
Positive serial correlation.

In figure 6.1a, the error term associated with the first case happens to be positive. The first 4 error terms are then positive, followed by two negative errors. In figure 6.1b, just the opposite occurs: the first 4 errors are negative and the last two are positive. Each of these are equally likely to occur, so OLS estimates will be unbiased. But, in both cases, the OLS line fits the observed cases more closely than the true regression line does; this leads to an R^2 that is too large and an error variance that is smaller than the true variance.

Tests for Serial Correlation. A popular test for serial correlation is the Durbin-Watson statistic:

$$DW = \frac{\sum_{t=2}^{T}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{T}\hat{\varepsilon}_t^2}$$

T = the number of time periods. Note that the numerator cannot include a difference for the first observation in the sample since no earlier observation is available. When successive values of $\hat{\varepsilon}_t$ are close to each other, the DW statistic will be low, indicating the presence of positive serial correlation.

The DW statistic will lie in the 0-4 range, with a value near two indicating no first-order serial correlation. Positive serial correlation is associated with DW values below 2 and negative serial correlation with DW values above 2. As Pindyck and Rubinfeld explain, exact interpretation of the DW statistic can be difficult.

Corrections for Serial Correlation. Pindyck and Rubinfeld describe various procedures for dealing with serial correlation. These include generalized differencing, the Cochrane-Orcutt Procedure, and the Hildreth-Lu procedure.

Related Topics. Stata has a variety of commands and procedures for time series data. Also, if you are working with cross-sectional time series data (i.e. longitudinal/ panel data) you should familiarize yourself with Stata's XT commands, e.g. `xtreg`. Often these problems involve large N/ small T (e.g. 2,000 subjects that have been interviewed 2 or 3 different times.) In other cases you have small N / large T (e.g. 50 countries that you have 60 years worth of annual data on). I probably won't have time to go over these in class but I can loan you some books and manuals if your research project requires such methods.

Stata Example. This example and discussion is shamelessly stolen pretty much verbatim from the Stata 12 Time Series Manual, pp. 323-324. I make no claims to original thinking or even original words here. It did take me a few minutes to cut and paste everything though.

We have data on quarterly sales, in millions of dollars, for 5 years, and we would like to use this information to model sales for company X. First, we fit a linear model by OLS and obtain the Durbin–Watson statistic by using estat dwatson:

```
. use http://www.stata-press.com/data/r12/qsales, clear
. regress csales isales

      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  1,     18) =14888.15
       Model | 110.256901      1  110.256901           Prob > F      =  0.0000
    Residual | .133302302     18  .007405683           R-squared     =  0.9988
-------------+------------------------------           Adj R-squared =  0.9987
       Total | 110.390204     19  5.81001072           Root MSE      =  .08606


------------------------------------------------------------------------------
      csales |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      isales |   .1762828   .0014447   122.02   0.000     .1732475    .1793181
       _cons |  -1.454753   .2141461    -6.79   0.000    -1.904657   -1.004849
------------------------------------------------------------------------------

. estat dwatson

Durbin-Watson d-statistic(  2,    20) =  .7347276
```

Because the Durbin–Watson statistic is far from 2 (the expected value under the null hypothesis of no serial correlation) and well below the 5% lower limit of 1.2, we conclude that the disturbances are serially correlated. (Upper and lower bounds for the d statistic can be found in most econometrics texts; for example, Harvey [1990]. The bounds have been derived for only a limited combination of regressors and observations.)

Harvey, A. C. 1990. The Econometric Analysis of Time Series. 2nd ed. Cambridge, MA: MIT Press.

See the Stata Manual for discussions on how to deal with the serial correlation in this specific problem.