

Figure 1: A Trend with Flattening

The webpost <u>http://tamino.wordpress.com/2009/12/15/how-long/</u> has a nice step-by-step exposition of how to estimate whether there is a warming trend in temperature data 1975-2008, first using OLS, then using an AR-1 process, then an ARMA. The trend is significant. But the post is responding to the observation that the trend has flattened out since 2000. It doesn't really respond to that.

To see why, note the graph above. It has artificial temperatures that rise from 1975 to 2000 and then flatten out. If you do an OLS regression, though, YEAR comes in significant with a t-statistic of 25.33 and an R2 of .95. I just did it with Excel, because I haven't installed StarOffice or STATA on my new computer here, but I'm sure that doing a serial correlation correction wouldn't alter the result much. Yet eyeballing it, we can see that though it is clear that temperatures have risen since 1975, it is also clear that they've flattened out since 2000. A linear regression just doesn't summarize the data correctly.

Let's do a couple more examples for fun and to drive home the point. In the second figure, the temperature levels out in 1982 but year is still highly significant, with a t-stat of 4.89, though the R2 drops to .42 (what's the R2 with the real data? –very small, I'd expect).



Okay, now look at the third figure, in which the trend actually reverses. The t-stat is actually bigger— 4.98--- and the R2 is .43.



So don't go and use a linear model when eyeballing the data tells you it isn't appropriate. When you have a simple regression in which only one variable explains another, use your eyes first, and software second. Do remember, though, that checking for statistical significance--- and autocorrelation and all those other things-- are useful too, so long as you start off right. Here, the question is not just "Have temperatures been rising with time over the past 30 years?" but, separately, "Have temperatures been rising with time over the past 10 years?"

The way to start addressing that with regression, by the way, is to do a regression of temperature on four variables: Constant, Year, a dummy equaling 1 if the year is after 1999 and 0 otherwise, and an interaction of that dummy with Year.

If a lot of people are interested, I could apply the serial correlation corrections to the artificial data or do this 4-variable regression on the real data, but maybe somebody else can take over now. My Excel spreadsheet is at <u>http://rasmusen.org/t/2009/warming.xlsx</u>, this document at <u>http://rasmusen.org/t/2009/warming.pdf</u>, I'm Eric Rasmusen at <u>erasmuse@indiana.edu</u>, and this is December 29, 2009.