The Stata file var.dta (which you can download from the courses page under my web page http://people.hamilton.edu/christophre-geeorges) contains quarterly macroeconomic data for the US from 1950:1-2010:4.

Specifically, the file includes data on Real GDP (y), the GDP Deflator (p), the unemployment rate (u), the growth rate of utilization adjusted (purified) TFP (dptfp), the real price of oil (rpoil), and a real interest rate (Moody's corporate BAA interest rate index less the inflation rate).[1]

For this exercise, we will use only a subset of this data. But you are welcome to try other variations.

A simple VAR exercise: Let's consider the relationship between Real GDP (y), the aggregate price level (p), and real oil prices (rpoil) for the post-war period in the U.S. We can run a VAR in Stata by selecting Statistics>Multivariate Time Series>Basic VAR. This will run the VAR and give impulse response (IR) plots for each pair of variables. For the purposes of generating the IR plots, Stata here uses the Sims method of asking for a recursive ordering of the variables (in terms of contemporaneous effects on one another).

So let's try the following ordering. Assume that oil prices respond (if at all) to changes in P and Y only with a lag (i.e., not contemporaneously). Assume further that the aggregate price level P responds (if at all) to changes in Y only with a lag (i.e., not contemporaneously). Note that we are not ruling out that changes in oil prices affect P immediately. Finally, assume that Y can respond to changes in both oil prices and P immediately. All variables can respond to changes in any other with a lag (we are putting no restrictions on the coefficients on lagged variables in the VAR).

The assumptions above make sense in terms of a standard new-Keynesian synthesis model (synthesis of RBC and Keynesian theory) that takes the prices of goods and services as being sticky with respect to spending shocks, but not necessarily cost shocks (like oil price shocks). Continuing with this interpretation, we can think of the P equation as an aggregate supply equation, and the Y equation as an aggregate demand equation, and the oil price equation as indicating that global oil prices could be at least partly endogenous to the state of the U.S. macroeconomy. This is clearly an overly simplified model (i.e., the VAR is likely to be misspecified),[2] but it may still give us some stylized evidence for the causal interactions between these variables.

One last thing before we start. Let's run the VAR on logged levels of the three variables. So please first create three new variables: lny = ln(y), lnp = ln(p), and lnrpoil = ln(rpoil).

To implement this VAR in Stata, select Statistics>Multivariate Time Series>Basic VAR, and then for the "dependent variables" select – lnrpoil lnp and lny – *in that order* (i.e., this both selects the variables to include and specifies the recursive order of the variables). Then increase the number of lags to include in the regression from 2 to 4 (this will include lags of 1,2,3 and 4 quarters for each explanatory variable in each equation), keep the selection of OIRF ("orthoganalized impulse

[1] The original data sources are BLS (prices and unemployment), Commerce Department (GDP), John Fernald (utilitation adjusted TFP), and the Federal Reserve Board (interest rates). All but Fernald's data can be easily downloaded from FRED.

[2] Further, the variables we are looking at are nonstationary (there are not fixed long run averages for these variables - e.g., Y grows on average over time), and if we were being more careful, we would try to correct for this.

response function" - this says use the recursive ordering to estimate the true shocks to each variable) and increase the number of periods in the horizon for the impulse response plots from 8 periods to 20 (this will create impulse response plots for 20 quarters (5 years) following hypothetical shocks rather than 8 quarters (2 years)).

Fire away, and Stata will run the regression and then (with a bit of a lag) generate a panel of OIRF plots. Unfortunately, it scales each plot the same, and so you will need to re-plot some of them separately (e.g., if lnrpoil rises by 0.1 in one graph, and lngdp rises by .01 in another graph, the gdp change will look miniscule on the second plot. So, e.g., to re-plot the response of gdp to an aggregate price shock, select Graphics>Time Series Graphs>Multivariate Time Series Graphs>IRF And FEVD After IRF Create, and then under "statistics to graph" select oirf, under "impulse variable" select lnp, and under "response variable" select lny.

Questions:

a. How consistent are the various impulse response plots with the predictions of the new-Keynesian synthesis model?

b. In order to identify the VAR, we assumed that oil prices are contemporaneously exogenous. Is there evidence in these results that oil prices may be endogenous to the state of the U.S. economy?