A Student's Primer to Reading Empirical Macroeconomics

Michael Smitka, Professor of Economics
Washington and Lee University, Lexington, VA 24450
http://econ398.academic.wlu.edu

Abstract

The basics of linear regressions that undergraduates encounter in a rigorous econometrics course do not leave them able to read work in empirical macroeconomics. This guide seeks to make the empirical literature more accessible.

A firm grasp of macroeconometric jargon is not in itself sufficient. It needs to be paired with a knowledge of modern theory, the mix of demand and supply components, of short- and long-run dynamics and of expectations that underly (modern) DSGE models. That effort (OLG models, growth models and rational expectations) comprises a major portion of the senior Econ 398 macro seminar. This handout does not touch upon such issues.

I. Introduction

Empirical macroeconomics faces three main challenges that are qualitatively different from those in microeconomics, both in substance and in the manner in which they are addressed. These are the time series nature of the data, the general equilibrium starting point that assumes key variables are endogenous, and the difficulty of identifying relationships among variables (particularly policy variables). Three additional complications also plague macroeconometricians, the small size of their datasets, the infrequency of events that are of interest (recessions, inflations and especially bubbles) and the challenge of determining what is a good fit when you have multiple equations.

Micro vs macro

Macro data consist of time series such as GDP that have one or another sort of trend. The techniques to handle such issues are at best covered in passing in a good undergraduate course. Hence the basic jargon of macroeconometrics – SVARs, unit roots and cointegration, impulse functions – conveys little to even a diligent economics major. What follows concentrates on parsing that language so as to better prepare students to read empirical papers.

A second issue is that macroeconomics focuses on interactions among key variables. As such, the underlying assumption is that all variables are endogenous. That makes it natural to work with systems of equations rather than individual equations. Furthermore, a century of empirical observation makes it clear that many key variables react only gradually to "shocks" to the economy. So these must not only be systems of equations, but systems that contain multiple lagged values – interest rates last quarter and the quarter before – and not merely contemporaneous values. In principle basic regressions will have dozens of dependent variables.

Third, and more subtly, policy is likely to be at least in part endogenous. Interest rates, for example, surely reflect a systematic response by monetary authorities both to inflation and to unemployment. If policy typically responds to developments in an economy, then that should be reflected in the structure of the model, and so it will be difficult to identify how policy matters. One way to capture that is by adding an explicit Taylor Curve for monetary policy; the error term is then a way to introduce "unanticipated" monetary policy. A similar approach for fiscal policy is used in the IMF working paper by Corsetti and others. Now the "Lucas Critique" argues that unanticipated policy changes the expectations mechanism and hence models will not accurately capture the impact of policy. But so might any large (that is, noticeable) shock to the system. So the Lucas critique is really an empirical statement about the stability of expectations formation, and so is amenable to testing: is there a "break" in models after a "policy innovation"? No study that I have come across trumpets such a finding.

Complications

The first complication is a paucity of good data. In the case of the US, in 1980 interstate banking was not allowed while transactions were settled by clearing checks, that is, physically sending pieces of paper around the country. Similarly, in 1980 exports plus imports were 20% of GDP, but were 30% in 2008; international capital flows, minimal in size in the aftermath of World War II, grew markedly while the US switched from being a (large) net creditor in 1980 to being a large net debtor in global markets. All of these might reasonably affect the nature of monetary policy and would certainly shift the size of parameters. However, if we limit ourselves to the last 30 years of data, then we are left with only 120 quarterly observations. That's not many, given the dozens of dependent variables that we'd ideally like to include. In practice, macroeconometricians must compromise lest they run out of degrees of freedom.

A second complication is that events of interest, such as recessions and inflations, occur infrequently, and in intervening periods neither monetary nor fiscal policy may change. Bubbles are (fortunately) even rarer than recessions, and so are harder to study. Japan's, for example, broke in 1991, though the implosion of the banking system did not occur until 1997. So we might want data from 1981 through 2007 as a minimum. However, the current version of Japan's System of National Accounts includes data only back to 1994 while data under the old methodology were carried forward only through 2001. The two series overlap, but there is no consistent dataset covering the entire period. (See the appendix for a detailed example.)

A third and final complication is that models include multiple equations, so it is unclear what might be meant by a "good" fit. When one variable tracks well but others do not, does that mean the model is of little value? In addition, many macroeconomic models are primarily simulations, which use parameter values "drawn from the literature", resulting models that can be loosely fitted to data but not estimated. What are we to make of such models?

Now I personally continue to read empirical work with great interest; the Great Recession has stimulated innovative research on identifying the impact of policy. However, those who rely on simulations can easily generate contrary findings. Understanding why requires delving into the structure of modern macro models. That is not the purpose of what follows, which focuses on the jargon and concepts of the statistical approaches of more narrowly empirical work.

II. Time series issues: unit roots

The first challenge of empirical macroeconomics is that data exhibit time trends; GDP, consumption and bank assets all grow, and so any simple regression comparing them will find coefficients near 1.0 in value with R^2s of 99%. But that would be true even if you put in a time series such as the annual height of trees as they grow from seedlings. The time trend dominates everything.

Detrending data turns out to be hard, because different types of trends require different statistical techniques but distinguishing the type of trend is not straightforward, particularly in the small-sized datasets typical of macro. Almost always using a first difference (using $\Delta y_t = y_t - y_{t-1}$, that is, using growth rates and not levels) will eliminate the trend, but at the cost of losing at least one observation, and of mixing in variables that aren't differenced (such as interest rates). Logged values of variables lessen remaining heteroskedasticity. Even then, serial correlation is likely to remain, as recessions and inflations, and high and low interest rates persist for several quarters; a high value relative to trend this period is likely to be followed by a high one next period. Adding lagged terms is natural.

An example helps illustrate the challenge. A simple time trend of GDP takes the following form (here abbreviated with the standard symbol Y with y for the log of GDP):

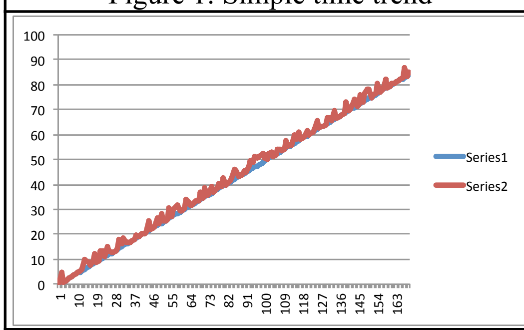
(1) $\mathbf{y_t} = y_0 + b\mathbf{t} + \boldsymbol{\varepsilon_t}$ where ε is a well-behaved normal error term $N(0,\sigma^2)$ of mean zero and constant variance.

When variables are in log form, that implies exponential growth:

(1') $\mathbf{Y_t} = Y_0 B^t \boldsymbol{\mu_t}$ where μ is the exponential form of our error term, and well-behaved.

A time series with a simple time trend such as Equation (1) will look like the graph in Figure 1, deviating around the trend but never moving away from it; on a random basis you find a few positive ε in a row (this example did not build in serial correlation), but then the next ε could be negative and you're below the line. On average, you're zero away.

Figure 1: Simple time trend



In macroeconomics, however, we frequently believe that variables will reflect past values plus a growth increment. So we might instead prefer to write the following equation. That way, if we have a recession – y_{t-1} is lower – then y_t will tend to be lower, too.

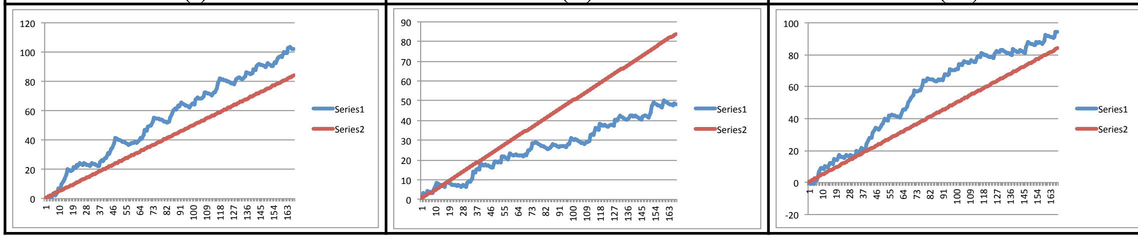
(2) $\mathbf{y_t} = b + \mathbf{y_{t-1}} + \boldsymbol{\varepsilon_t}$

If we add that up – substituting recursively until we are back at y_0 – we get a form that looks identical to (1) except for a cumulating error. Note as well that everything is lower case, indicating log values. Hence what this in fact reflects is exponential growth (as represented by the non-logarithmic version (1') above). The only difference would seem to be an innocuous change in the error term, which now includes past error terms as well.

(2') $\mathbf{y_t} = y_0 + bt + \sum \boldsymbol{\varepsilon_t}$

However, if we graph this as per Figure 2 we see something strikingly different: the data no longer stick close to the time trend, but can wander away from it – in fact, the process is now a random walk around a time trend.

Figure 2: time trend with a lagged dependent variable		
(i)	(ii)	(iii)



In the case of Equation (1) an error isn't propagated; the series returns to its underlying trend. However, in the case of Equation (2) the impact of a large error term persists; a large negative ε leaves GDP permanently lower. In the first graph a "bad" early draw is only offset at the very end by a couple favorable values of ε while in the second and third graphs the result is (after an initial blip or two) permanently higher or lower values. To reiterate, there's no tendency to revert to trend. The final summation sign thus hides a crucial difference: because we are summing our error term ε variance increases over time, tending to infinity.

To use the jargon of the trade, the series is **non-stationary**, tending to wander away; what you want is a series that is **stationary** around the mean. As you may suspect, in practice many macro time series prove *non-stationary.*

This matters a lot. When the error term has increasing variance, regressions produce spurious results. It's not just that the statistical fit is less good, it's that coefficients can be statistically significant but have the wrong sign. If you eliminated the time trend and ran a regression with (i) and (iii) you'd get a positive correlation, with (ii) and either (i) or (iii) you'd get a negative correlation. Yet the underlying process in all three is the same. In other words, failing to spot a *non-stationary* series leads to results that are at best meaningless and are potentially misleading, despite the computer telling you that you have a statistically significant coefficient.

A simple test for the presence of a random walk is to regress a variable on its lagged self:

$$(3)\quad \mathbf{y}_t = \alpha + \beta\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

When $\beta < 1$ we get a convergent geometric series for the error term ($\sum \beta^t\varepsilon_t$) so the variance is finite and a regression will work; if $\beta = 1$ we have equation (2) above. (If $\beta > 1$ the series diverges, which is not meaningful in a macro context. In other words, there's something very wrong with your data!) This leads naturally to testing for a "**unit root**" (that $\beta = 1$). As it happens, the standard t-statistic is commonly used, but the critical values for significance are much stricter, needing a value of over 3-4 in the size samples encountered in macro. In other words, it's hard to "reject a unit root" because the critical value is so high.

Now if there is a unit root – or rather, because of low statistical significance of the coefficient, you can't reject a unit root – then you may have a time series with a random walk, one that is *non-stationary*. Such a time series may also be called **I(1)** (integrated of order 1). In macroeconomic data, taking a first-difference (using $\Delta y_t = y_t - y_{t-1}$ instead of y_t) will almost always eliminate the random walk, making the series *stationary* or (equivalently) turning it into an *I(0)* series, with some loss of information and with a regression that mixes variables expressed in levels and variables expressed in differences.

Macroeconometricians try to avoid that. Doing so leads to a second test and associated jargon. Frequently several *non-stationary* variables appear in the same regression. In macroeconomic work, it is not unusual that such series – such as Y (GDP) and C (consumption) – will move together. In that case, a (linear) combination of the two can sometimes eliminate the random walk behavior, as a movement away from trend in one is matched by a movement away from trend in the other and so can be used to offset it. In such cases these variables **cointegrate.** The bottom line is that an appropriate combination of the variables can be used, because the combination will no longer have a random walk component. In such cases, there's then no need to use first differences.

Table 1: Time series jargon

one variable			
random walk	non-stationary	I(1)	unit root
simple trend	stationary	I(0)	reject unit root
two or more variables			
each is a random walk	non-stationary	I(1)	unit root
combined not a random walk	stationary	I(0)	cointegrate

In sum, when you encounter the above jargon, you need to know that it reflects an effort to discern the type of time trend and to make appropriate statistical corrections. Of course that leads to a whole array of names of tests for unit roots. (You may see "Dickey-Fuller" after the pair to first calculate the critical value for the t-statistic.) There are also specialized statistical procedures to implement corrections other than by taking first differences or utilizing cointegration. For better or for worse, to be published papers need to demonstrate something new, so you're likely to be treated to a lengthy discussion of such niceties; if nothing else, it shows that the authors are appropriately trained members of the confraternity of macro people and know what they're doing. So don't get your hopes up that you'll encounter many papers that simply state "we did the normal tests and corrections for time series issues."

Error-correcting models (**ECM**s) assume lie between a random walk and a time trend, putting in a factor that forces the variable to revert to trend. The "error" (deviation) from trend is $E_t = y_t - bt$. So here you subtract some fraction of that: (4) $\mathbf{y_t} = b + \mathbf{y_{t-1}} - \alpha E_{t-1} + \boldsymbol{\varepsilon_t}$ [where $0 < \alpha < 1$].

III. Endogeneity issues: vector autoregressions

The standard microeconomic model is built upon partial equilibrium analysis. Bringing data to it has its own challenges. Observations are the product of a market process, jointly determined by supply and demand. Control variables may correlate with those of interest so can't be excluded. Observables may be weak proxies for key theoretical variables, and data can be qualitative. While causation can be ambiguous, the starting point is the assumption that variables are independent, that a market can be isolated from its affect on other markets, that feedback effects are minimal.

Macroeconomics begins from the assumption of general equilibrium, that no variable is independent. Furthermore, the underlying world is dynamic, and what we have are measures that are not yet at an equilibrium. The dynamics are clearly non-linear, involving cycles and overshooting. In that context history matters, the state of the economy constrains where it will be tomorrow, and we are typically more interested in what happens next quarter than in the final equilibrium some years hence. Furthermore, all models have as key variables expectations of the future.

Vector autoregressions are one econometric tool that addresses these issues. As noted above, lagged variables incorporate history into models, but they also produce rich dynamics. Autoregressive (integrated) moving average processes – ARIMA models – were known to do that. The next step was to jointly estimate autoregressions in several variables (hence the "vector" terminology). In the basic form each variable appears as a (lagged) variable in every equation in the model. That makes explicit the interdependence that lies at the core of macroeconomics. It also provided a setup where a series of past variables determine the present value. That captures the role of expectations, however imperfectly, which is a key desideratum of modern macroeconomics.

The result is that what is estimated would look something like this for a model that includes government purchases, money, prices and GDP. Here the setup is for a 4-variable VAR run using logged variables; you may also see VARs run in differences ΔY_t, ΔM_t and so on rather than logged values.

$$\mathbf{y_t} = \alpha_0 + \alpha_{11}\mathbf{y_{t-1}} + \ldots + \alpha_{1n}\mathbf{y_{t-n}} + \alpha_{21}\mathbf{m_{t-1}} + \ldots + \alpha_{2n}\mathbf{m_{t-n}} + \alpha_{31}\mathbf{p_{t-1}} + \ldots + \alpha_{3n}\mathbf{p_{t-n}} + \alpha_{41}\mathbf{g_{t-1}} + \ldots + \alpha_{4n}\mathbf{g_{t-n}} + \boldsymbol{\varepsilon_t} \tag{5}$$

$$\mathbf{m_t} = \beta_0 + \beta_{11}\mathbf{y_{t-1}} + \ldots + \beta_{1n}\mathbf{y_{t-n}} + \beta_{21}\mathbf{m_{t-1}} + \ldots + \beta_{2n}\mathbf{m_{t-n}} + \beta_{31}\mathbf{p_{t-1}} + \ldots + \beta_{3n}\mathbf{p_{t-n}} + \ldots + \beta_{41}\mathbf{g_{t-1}} + \ldots + \beta_{4n}\mathbf{g_{t-n}} + \boldsymbol{\eta_t}$$

$$\mathbf{p_t} = \gamma_0 + \gamma_{11}\mathbf{y_{t-1}} + \ldots + \gamma_{1n}\mathbf{y_{t-n}} + \gamma_{21}\mathbf{m_{t-1}} + \ldots + \gamma_{2n}\mathbf{m_{t-n}} + \gamma_{31}\mathbf{p_{t-1}} + \ldots + \gamma_{3n}\mathbf{p_{t-n}} + \gamma_{41}\mathbf{g_{t-1}} + \ldots + \gamma_{4n}g_{t-n} + \boldsymbol{\zeta_t}$$

$$\mathbf{g_t} = \theta_0 + \theta_{11}\mathbf{y_{t-1}} + \ldots + \theta_{1n}\mathbf{y_{t-n}} + \theta_{21}\mathbf{m_{t-1}} + \ldots + \theta_{2n}\mathbf{m_{t-n}} + \theta_{31}\mathbf{p_{t-1}} + \ldots + \theta_{3n}\mathbf{p_{t-n}} + \theta_{41}\mathbf{g_{t-1}} + \ldots + \theta_{4n}\mathbf{g_{t-n}} + \boldsymbol{\mu_t}$$

The details of actually running this set of equations is messy. There's no objective method for deciding how many lags to include, but degrees of freedom force a severe limit. The parameters generated may be jointly significant in a statistical sense but often will not be individually significant. Furthermore, what you get is a large number of lagged variable coefficients. (See Figure 3.) No economic interpretation can be placed on theses, e.g., there is not obvious meaning to the second price lag coefficient in the money equation.

The basic VAR model is atheoretic: it includes nothing that we believe we know about how a macroeconomy works, over and above the list of variables. Early proponents viewed that as a strength, a way to avoid battling models. Varying the lags, shifting particular measures of variables, even changing the order of the equations in the estimation procedure all lead to shifts in the results. That undermined early hopes that VARs in themselves would resolve disputes. Subsequently, techniques such as structural VARs (SVARs), error-correction models (ECMs) and other variations have been added to our palette. In SVARs, some of the coefficients (or error terms) are forced to take specific values. (If nothing else, you don't want models that can generate negative interest rates!) Error-correction models add terms to push specific variables, for example pushing the GPD variable towards potential GDP.

Figure 3: The Coefficients of a VAR

TABLE 2.—ESTIMATED VECTOR AUTOREGRESSION

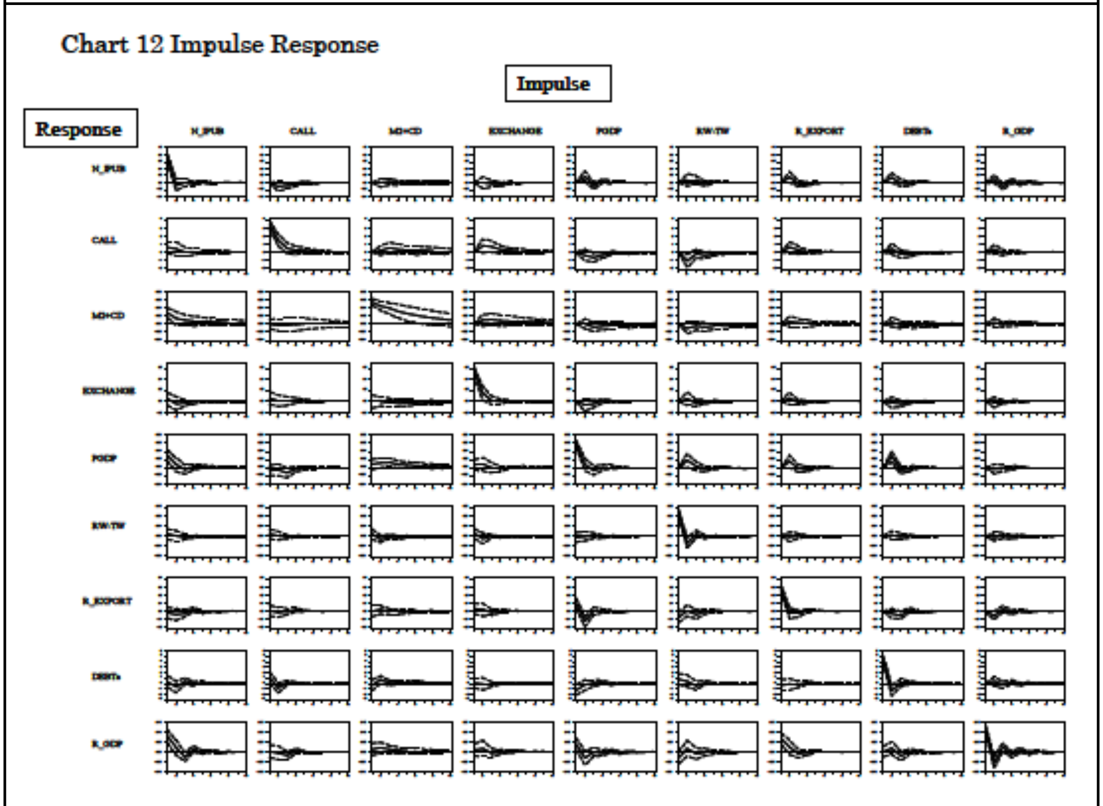
Variable	Equation: e	$(y-y^*)$	$(m-m^*)$	$(p-p^*)$	tb	tb^*	$(g-g^*)$
e_{-1}	0.906[b]	−0.153[a]	0.069	−0.163[b]	−0.038	0.014	0.112
e_{-2}	−0.388[b]	0.062	−0.069	0.148[a]	−0.034	−0.019	−0.319[b]
$(y-y^*)_{-1}$	−0.053	0.496[b]	−0.174	−0.081	−0.227[a]	−0.017	0.043
$(y-y^*)_{-2}$	−0.010	−0.180[a]	−0.094	0.106	0.161[a]	−0.094[a]	−0.271[a]
$(m-m^*)_{-1}$	−0.598[b]	−0.087	0.284[a]	−0.039	0.085	−0.017	0.140
$(m-m^*)_{-2}$	0.627[b]	−0.022	0.329[b]	−0.026	−0.038	0.028	−0.032
$(p-p^*)_{-1}$	−0.619[a]	−0.523[b]	−0.113	1.079[b]	0.068	0.019	−0.494[a]
$(p-p^*)_{-2}$	0.589[a]	0.538[b]	0.148	−0.282[a]	−0.107	−0.083[a]	0.318[a]
tb_{-1}	0.093	−0.542[b]	−0.071	0.171[a]	0.622[b]	−0.047	−0.164
tb_{-2}	−0.036	0.340[a]	−0.433[a]	−0.107	−0.196[a]	−0.040	0.231
tb^*_{-1}	1.727[a]	−0.774[a]	0.398	0.090	0.140	0.919[b]	−0.466
tb^*_{-2}	−1.383[a]	2.011[b]	0.156	0.179	−0.353	0.026	0.129
$(g-g^*)_{-1}$	0.012	−0.020	0.221[a]	0.126[a]	0.173[a]	−0.008	0.392[b]
$(g-g^*)_{-2}$	0.411[a]	0.060	0.191[a]	0.024	−0.130[a]	0.007	−0.171
SE(×10)	0.1603	0.0692	0.0856	0.0657	0.0671	0.0267	0.1173
AR(1)	3.08	2.64	1.73	0.00	0.71	1.55	0.42
	(0.09)	(0.11)	(0.20)	(0.96)	(0.41)	(0.22)	(0.52)
AR(4)	2.03	1.92	2.36	0.53	1.33	3.97	2.05
	(0.12)	(0.14)	(0.08)	(0.71)	(0.28)	(0.01)	(0.12)
R^2	0.9810	0.9633	0.9975	0.9960	0.9215	0.9573	0.9035

Finally, how do you interpret a VAR, when the coefficients aren't in and of themselves meaningful? The tool here is to take the estimated model, and impose an impulse (by convention, of 1.0 standard deviation or, for interest rates, of 1.0 percentage points) on the error term associated with a variable, which requires setting up the technical side of the model so that it is meaningful to tweak a single error term (it needs to be "orthogonal," independent of the other error terms). The impact of the impulse on the variables in the model going forward can then be compared with what happens without the impulse, shown as a graph of the deviation over time. If we look at the VAR above, a 1% increase to the "G" government purchase variable will generate an impulse graph of how "Y" GDP responds. The area under the graph – the cumulative change in Y – then provides an estimate of the simple government purchase multiplier. The shape of what happens when money or interest rates are changed by 1% provides an estimate of how quickly a policy change takes effect, when it peaks, and how long it takes to peter out. In contrast to a traditional (microeconomic) regression, a VAR makes no pretense to capturing what goes on in a dynamic model with a single coefficient.

The output of a VAR is thus not a table, such as that above, but a series of "impulse" graphs, for which additional metrics may (or may not) be provided. Figure 4 provides an example for a 9-variable VAR, which thus has 9 equations, each with lags for all 9 variables. Rather than reporting details of each of the 9 equations (such as how many lags were included for which variable) the output consists of a set of impulse graphs. In this case, with 9 variables, the complete output (the authors weren't being very helpful!) is a set of 9 graphs for each of the 9 variables, comprising 81 graphs in total.

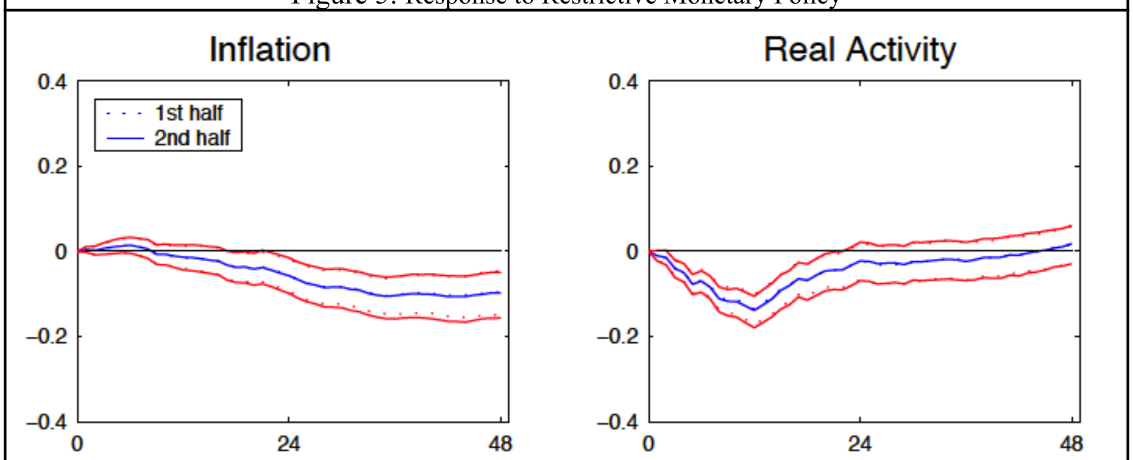
The second, more representative case is one such as Figure 5, which adds 95% confidence bands. Here we can see that the maximum response to a boost in interest rates on real GDP comes after 12 months, and remains negative until the 24 month point, after which it is statistically indistinguishable from 0. In contrast, inflation only falls significantly starting at the 24 month point; monetary policy operates more slowly, but inflation shows no tendency to rebound to its former level.

Figure 4: What you really get with a VAR



Harada and Onishi (2003) Japan's Great Recession, ESRI Discussion Paper Series No. 77.

Figure 5: Response to Restrictive Monetary Policy



Belviso & Milani, "Structural Factor-Augmented VARs…" Topics in Macroeconomics 6:3 (2006), p 29.

IV. Conclusion

Macroeconometrics differs from the content of Economics 203 in the limited availability of data, its time series nature and endogeneity. As a result, regression in macroeconomics vary in the technical issues that must be overcome. That results in a whole set of new terminology. In addition, the way in which you interpret regression results is different. In general, regression coefficients themselves convey little information, and may be statistically insignificant. Hence output takes the form of modified projections versus trend projections. Finally, the multi-equation nature of macro models means that there is no unambiguous way to define a "good" model. Instead, given the limited degrees of freedom, models will be tailored to address specific issues. The focus then becomes more on how well those components fit historic data and (future) out-of-sample data, and less on how well other elements track the economy. Such compromises are not always highlighted, but the deviations from how variables evolve in real data can be disconcertingly large.

Empirical macroeconomics is necessary: policymakers face imperatives to address inflation and unemployment. It is however unrealistic to think that empirics can be used to cleanly "test" theoretical propositions. Furthermore, results that fail to meet expectations lead practitioners to modify their models, not reject them. Macroeconometric work exerts a loose influence on the direction of research.

Egos come into play; those outside economics are predisposed to cherry-pick results that match their priors. Economists, or at least few dozen based in premier research positions, are willing to play into that game. Be suspicious of a macroeconomist who avoids nuance in their blogs, and who cannot point out to strengths (or who use none of the methodologies) of their rivals. But don't expect to find that on the op/ed page; editors are uninterested in even-handed economists, they want someone with a story that has edge to it. For better or for worse we will continue to be treated to a battle of alternate multipliers.

V. Appendix

The structure of the developed world, the OECD economies, evolved markedly since GDP data began to be compiled on a standardized basis in the 1950s. Except for the US and the UK, the first countries to implement national accounts, OECD members follow the System of National Accounts standardized under UN auspices. (The US makes the distinction clear, calling its system NIPA, the National Income and Product Accounts.) In the 1950s, however, agriculture and mining were important in Europe and Japan; services were more modest. Retailing included many mom-and-pop stores that kept no proper books, so it was hard to collect sales data on a timely basis. The SNA has been modified since to take advantage of improved methodology, the need to capture new sorts of economic activity, and greater data availability. However, because of the inclusion of better and more detailed data, the new series cannot be extended backwards, or can be calculated retroactively only with strong assumptions.

Japanese data illustrate this, because their statistical agency did not replace the older series with retroactive ones, choosing to value internal consistency at the cost of a longer but less comparable time series. The top half of Figure 5 graphs GDP, business fixed investment and private consumption for the two main SNA series. In the period where they overlap, the magnitudes of these key variables are visibly different. Worse, as illustrated by the bottom chart on the share of consumption in GDP, the new and old variables do not always move in the same direction. Furthermore, some sub-measures (such as data on business and household net savings) are only available in 3 separate, partially overlapping series. To make matters worse, the 1994-2001 period with inconsistent overlapping data comes immediately after the bursting of Japan's real estate bubble. So a single series is unavailable exactly where it might be most valuable, in contributing to our understanding of how a large, high-income economy responds to that sort of macroeconomic disaster.

GDP accounts are ***not*** consistent across the decades, and they are ***not*** consistent across countries. Any long series will include variables recalculated on a retroactive basis. Despite the common SNA framework, the data used to build particular series will vary by country, and in its methodology the US remains an outlier. While that does not stop economists from running regressions using cross-country data and long time series, it should give us pause in interpreting their results.

Figure 6: Japan: SNA 19968 vs SNA 1993 GDP Series