SGPE Econometrics 1 Projects (* indicates advanced)
(version of 13 October 2016)

1. Conditional convergence

Mankiw-Romer-Weil, in an influential 1992 paper, ask whether an augmented Solow growth model survives an encounter with the data. They conclude it does, but we now have 20 more years of data, estimation methods have also advanced, and other plausible model specifications are possible. Are MRW's results robust to various specifications? To different time periods and country sets? To updated data?

Data:
MRW data (available on network).
World Bank/ICP data on GDP per capita: http://data.worldbank.org/indicator
Penn World Tables GDP data: http://pwt.econ.upenn.edu/

References:
Fumio Hayashi (2000), *Econometrics*, pp. 342-46, 358-68.
Mankiw-Romer-Weil (1992), "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics*, Vol. 107, No. 2.
Jonathan R.W. Temple (1998), "Robustness Tests of the Augmented Solow Model", *Journal of Applied Econometrics*, Vol. 13, pp. 361-375.
Brad DeLong (2008), http://delong.typepad.com/sdj/2008/05/exercise-the-ca.html

2. Forward exchange rates as optimal predictors

The "forward premium" is the difference between the forward and spot exchange rates. A hypothesis that has been the subject of various studies is that the forward premium should equal the expected rate of currency appreciation over the life of the forward contract. An alternative way of stating the hypothesis (see Hayashi) is that the forward rate is the optimal forecast of future spot rates.

Data requirements: spot market price of dollar, 30-day forward price of the dollar, etc. See Hayashi/Verbeek.
Variants: pound sterling, deutchmark/euro, yen, others.

References:
Fumio Hayashi (2000), *Econometrics*, pp. 418-28, 438-40.
Marno Verbeek (2012), *A Guide to Modern Econometrics*, 4th ed., Wiley, pp. 127-34.
G. Bekaert and R. Hodrik (1993), "On Biases in the Measurement of Foreign Exchange Risk Premiums", *Journa of International Money and Finance*, Vol. 12.
M.P. Taylor (1995), "The Economics of Exchange Rates", *Journal of Economic Literature*, Vol. 33.

3. Gender and wages

The gender wage gap is another classic empirical question in applied econometrics and labour economics: how much of the observed pay gap between men and women is attributable to different characteristics (e.g., schooling), to different impacts of different characteristics (e.g., the return to schooling), how big is the unexplained residual, and how do we interpret it? The now-standard decomposition approach was introduced here by Oaxaca (1973) and Blinder (1973). Straightforward extensions are to use other methods and refinements of the decomposition method (readily available as Stata add-ins) and to explore the changes in the gender wage gap over time.

Data requirements:
The standard UK dataset for this kind of analysis is the Quarterly LFS (Labour Force Survey): http://www.esds.ac.uk/government/lfs/
You will need to register to use the data here: http://www.esds.ac.uk/aandp/access/access.asp
The LFS dataset are available in Stata format, but the datasets are large and require some manipulation to create a suitable subset of observations and variables. You will be provided with some Stata do files that will do this for a basic dataset and that you can modify appropriately.

References:
Marno Verbeek (2012), *A Guide to Modern Econometrics*, 4th ed., Wiley, pp. 127-34.
James H. Stock and Mark M. Watson (2012), *Introduction to Econometrics*, 3rd ed., pp. 325-6.
R. Oaxaca (1973), "Male-female wage differentials in urban labour markets", *International Economic Review*, 14(3): 693-709.
A. Blinder (1973), "Wage discrimination: Reduced form and structural variables", *Journal of Human Resources*, 8: 436-465.
An excellent and extensive, though now slightly aging, review of the UK gender pay gap by the NIESR (National Institute for Economic and Social Research):
Tracy Anderson, John Forth, Hilary Metcalf and Simon Kirby (2001), "The Gender Pay Gap: Final Report to the DfEE", http://www.niesr.ac.uk/pubs/searchdetail.php?PublicationID=380

4. Intertemporal asset pricing*

Probably the most common use of nonlinear GMM methods is to estimate intertemporal asset pricing models. The seminal paper here is Hansen and Singleton (1982), who demonstrated the usefulness of the GMM approach. They use a nonlinear rational expectations model with a representative agent who maximises (expected) utility from consumption, and show that the first-order conditions ("Euler equations") are sufficient for estimating the model. In a nutshell, since agents are forward-looking optimisers, the information set of the agent at time t provides exogenous instruments that should be uncorrelated with the error. Stata's `gmm` command will estimation nonlinear models using 2-step efficient GMM; you will be provided with separate estimation routines and Mata code that will enable you to specify and estimate models using the CUE (continuously-updated) GMM estimator.

Data requirements: depends on the approach adopted. You'll need data on consumption and interest rates; the variety arises in the choice of instruments and how they're assembled.

References:
Marno Verbeek (2012), *A Guide to Modern Econometrics*, 4th ed., Wiley, pp. 166-7, 171-75.
Fumio Hayashi (2000), *Econometrics*, pp. 454-55.
William H. Greene (2012), *Econometric Analysis*, 7th ed, pp. 495-96.
Hansen-Singleton (1982), "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models", *Econometrica*, Vol. 50, No. 5, September.
Alastair R. Hall (2005), *Generalized Method of Moments: Advanced Texts in Econometrics*, Oxford University Press. (He uses Hansen-Singleton as an example at various points.)
Patrick Gagliardini (n.d.), presentation notes on GMM estimation of asset pricing models, available at www.people.usi.ch/gagliarp/additional/GMMASSETPRICING.pdf.

5. Crime rates in US cities/counties

Kelly uses a dataset of US cities to examine the relationship between inequality and crime. He distinguishes between violent and property crime, and in some estimations treats policing levels as endogenous using IV techniques. He estimates a Poisson regression, but for this project OLS/IV/GMM can be used.

Data:
US city and county data are available from various websites, e.g., http://www.nhgis.org.
Crime data and other US city/county data are available from ICPSR, http://www.icpsr.umich.edu.

Basic: cross-county or cross-city determinants of crime, with a particular focus on inequality.
Advanced: estimation techniques employed by Kelly, IV estimation, etc.

References:
Jeffrey M. Wooldridge (2009), *Introductory Econometrics: A Modern Approach*, 4th ed, pp. 311, 549, 565-66.
Morgan Kelly (2000), "Inequality and Crime", *Review of Economics and Statistics*, 2000, Vol. 82, No. 4.

6. Colonial origins of comparative development

Acemoglu-Johnson-Robinson use cross-country data to look at the impact of institutions on economic development (income per capita), using IV techniques to address endogeneity; European mortality rates are used as instruments for current institutions. Sachs criticises the results, arguing that geography has direct effects. Albouy criticises them on data grounds. Easterly and Levine extend the analysis to focus specifically on European colonial settlements. Do AJR's results hold up to Sachs' critique? To Albouy's critique? To variations in estimation methods, variables are measured, the use of updated data, etc? Is the Easterly-Levine focus on European colonial settlements justified?

Data:
AJR: http://econ-www.mit.edu/faculty/acemoglu/data/ajr2001
Albouy: http://www-personal.umich.edu/~albouy/
Easterly-Levine: http://faculty.haas.berkeley.edu/ross_levine/papers.htm
Gallup-Mellinger-Sachs geography data: http://www.cid.harvard.edu/ciddata/geographydata.htm
World Bank/ICP data on GDP per capita: http://data.worldbank.org/indicator
Penn World Tables GDP data: http://pwt.econ.upenn.edu/

References:
Acemoglu-Johnson-Robinson (2001), "The Colonial Origins of Comparative Development ", *American Economic Review*, Vol. 91, No. 5, December.
David Albouy (2012), "The Colonial Origins of Comparative Development: An Empirical Investigation: Comment", *American Economic Review*, Vol. 102, No. 6, October, pp. 3059-3076.
Easterly-Levine (2015), "The European Origins of Economic Development". Latest version available on Levine's website (see above); NBER working paper version (2012) available at http://www.nber.org/papers/w18162.pdf.
Hall-Jones (1999), "Why Do Some Countries Produce So Much More Output Per Worker Than Others?", *Quarterly Journal of Economics*, Vol. 114, No. 1, February.
Sachs (2003), "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income", NBER Working Paper 9490, http://www.nber.org/papers/w9490.pdf.
Edward L. Glaeser et al. (2004), "Do Institutions Cause Growth?", *Journal of Economic Growth*, Vol. 9, pp. 271-303.
A 4-part series by Dietz Vollrath at the Growth Economics Blog:
https://growthecon.wordpress.com/2014/11/18/the-skeptics-guide-to-institutions-part-1/
https://growthecon.wordpress.com/2014/11/20/the-skeptics-guide-to-institutions-part-2/
https://growthecon.wordpress.com/2014/11/23/the-skeptics-guide-to-institutions-part-3/
https://growthecon.wordpress.com/2014/11/26/the-skeptics-guide-to-institutions-part-4/

7. Trade and growth

Frankel and Romer try to identify the causal impact of trade on growth using IV methods and instruments constructed from the geographic component of countries' trade. Do their results hold up to alternative estimation methods, variance-covariance estimators (cluster-robust and 2-way-cluster robust), updated data, etc?

Data:
F-R: Data available at Romer's website, http://elsa.berkeley.edu/~dromer/
World Bank/ICP data on GDP per capita: http://data.worldbank.org/indicator
Penn World Tables GDP data: http://pwt.econ.upenn.edu/
See other projects for cross-country datasets on institutions and geography.

References:
Frankel-Romer (1999), "Does Trade Cause Growth?", *American Economic Review*, Vol. 89, No. 3 June.

8. Exploring inflation dynamics

Inflation exhibits very high persistence in industrial economies, approaching that of a random walk. Levin and Piger, in a study of 12 OECD countries for the period 1984-2003, first estimate a univariate AR process for inflation to construct measures of inflation persistence. They then consider the possibility that the appearance of persistence is exaggerated because of unmodelled structural breaks. They find evidence of such structural breaks, which they interpret as corresponding to shifts in the inflation objectives of central banks. Once they allow for structural breaks, inflation persistence declines markedly. Do their findings stand up to extension to new data, different countries, and/or different techniques?

Data:
Quarterly data on inflation for various countries; various measures of inflation can be used.

References:
Andrew T. Levin and Jeremy M. Piger, "Is Inflation Persistence Intrinsic in Industrial Economies?"
ECB version: Working Paper Series No. 334, April 2004.
https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp334.pdf
St. Louis Fed version: Working Paper 2002-023E, October 2002, revised November 2003.
http://research.stlouisfed.org/wp/2002/2002-023.pdf

9. Growth and inequality

Do more countries with more equal income distributions grow faster than those that are less egalitarian? This is a topic that has been analysed in various studies, with conflicting results. The variation in results is attributable at least in part to the range of different specifications employed (cross-section, panel, etc.), estimation methods, and choice of control variables. Banerjee and Duflo (2003) and Benabou (2000) are useful reviews of the literature. A number of these studies used a country-level dataset on inequality assembled by Deninger and Squire (1996). More recently, the University of Texas Inequality Project (UTIP) has extended and improved the Deninger-Squire data. This project will explore the topic using the new UTIP data and various specifications, choices of controls, and/or estimation methods as employed by these earlier researchers.

Data:
University of Texas Inequality Project data: http://utip.gov.utexas.edu/data.html

Deninger-Squire (1996) inequality data:
http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/0,,contentMDK:20699070~pagePK:64214825~piPK:64214943~theSitePK:469382,00.html
Barro-Lee education data: http://barrolee.com/
World Bank/ICP data on GDP per capita: http://data.worldbank.org/indicator
Penn World Tables GDP data: http://pwt.econ.upenn.edu/

References:
Abhijit V. Banerjee and Esther Duflo (2003), "Inequality and Growth: What Can the Data Say?", *Journal of Economic Growth*, Vol. 8, pp. 267-299. http://economics.mit.edu/files/753.
Roland Benabou (2000), "Unequal societies: Income distribution and the social contract", *American Economic Review*, Vol. 90, No. 1, pp. 96-129.
Federico Cingano (2014), "Trends in Income Inequality and its Impact on Economic Growth", *OECD Social, Employment and Migration Working Papers*, No. 163, OECD Publishing. http://dx.doi.org/10.1787/5jxrjncwxv6j-en
Roberto Perotti (1996), "Growth, income distribution and democracy", *Journal of Economic Growth*, Vol. 1.
Robert J. Barro (2000), "Inequality and growth in a panel of countries", *Journal of Economic Growth*, Vol. 5, No. 1.
Kristin J. Forbes (2000), "A reassessment of the relationship between inequality and growth", *American Economic Review*, Vol. 90, No. 4.
Klaus W. Deininger and Lyn Squire 1996, "A New Data Set Measuring Income Inequality", *The World Bank Economic Review*, Vol. 10, No. 3, pp. 565-91.

10. Returns to schooling*

Estimating the return to investment in human capital – schooling – is an important problem in applied econometrics and labour economics, and many studies have been devoted to this. The challenge is addressing a classic omitted variables/ endogeneity problem: because innate talent is unobserved, schooling is endogenous, and OLS estimates of the return to schooling are biased.

The usual approach is instrumental variables. Angrist and Krueger (1991), in an influential study, used quarter-of-birth as an instrument. Bound-Jaeger-Baker (1995) pointed out that A-K's instruments were "weak". How to deal with the weak instruments problem is currently a very active area for econometric research. Techniques for this problem are now becoming available in standard packages. This project will explore various techniques available using the A-K dataset.

Data:
Angrist's website: http://economics.mit.edu/faculty/angrist/data1/data/angkru1991
The Montiel-Olea and Pflueger test is available in Stata as the command **`weakivtest`**, available from SSC (see below).
The weak-instrument-robust tests are available in Stata as the command **`weakiv`**, available from SSC (see below).

References:
Fumio Hayashi (2000), *Econometrics*, pp. 236-44, 250-56.
James H. Stock and Mark M. Watson (2012), *Introduction to Econometrics*, 3rd ed., pp. 325-6 (return to schooling), 506-09 (weak instruments).
Marno Verbeek (2012), *A Guide to Modern Econometrics*, 4th ed., Wiley, pp. 156-60.

Angrist-Krueger, "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*. Vol. 106, No. 4, November.
Joshua D. Angrist and Alan B. Krueger (2001), "Instrumental Variables and the Search

for Identification: From Supply and Demand to Natural Experiments", *Journal of Economic Perspectives*, Vol. 15, No. 4.
John Bound, David Jaeger and Regina Baker (1995), "Problems with Instrumental Variables Estimation when the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak", *Journal of the American Statistical Association*. 90:430, pp. 443–50.
Jose Luis Montiel Olea, Carolin Pflueger and Su Wang, 2013. "WEAKIVTEST: Stata module to perform weak instrument test for a single endogenous regressor in TSLS and LIML".
https://ideas.repec.org/c/boc/bocode/s457732.html. To install in Stata:
`ssc install weakivtest`.
Keith Finlay, Leandro Magnusson and Mark E. Schaffer, 2013. "WEAKIV: Stata module to perform weak-instrument-robust tests and confidence intervals for instrumental-variable (IV) estimation of linear, probit and tobit models".
https://ideas.repec.org/c/boc/bocode/s457684.html. To install in Stata: `ssc install weakiv`.

11. The effect of health on wages

This project will use the British Household Panel Survey (BHPS) to explore the effects of health on wages. Basic human capital theory tells us that poor health should have a negative impact on wages via its impact on productivity. The longitudinal nature of the BHPS means that panel data various panel data estimators can be used and compared. An interesting problem is the possible two-way causal relationship between health and wages: *ceteris paribus*, people in poor health will have lower productivity and hence lower wages; but also *ceteris paribus*, high wage earners can afford to maintain better levels of physical health. Another possibility is that observed health status is correlated with unobserved characteristics that also have an impact on productivity and wages. One possible route to explore is the use of panel data instrumental variables estimators (either standard IV/GMM or more sophisticated estimators such as the Hausman-Taylor estimator), where health status is treated as an endogenous regressor.

Data requirements:
The BHPS is available at
http://www.esds.ac.uk/longitudinal/access/bhps/L33196.asp
You will need to register to use the data.
The BHPS data are available in Stata format, but the dataset is very large and requires some manipulation to create a suitable subset of observations and variables. You will be provided with some Stata do files that will do this for a basic dataset and that you can modify appropriately.

References:
P. Contoyannis and N. Rice (2001), "The impact of health on wages: Evidence from the British Household Panel Survey", *Empirical Economics*, Vol. 26(4): 599-622.
S. Ettner (1996), "New evidence on the relationship between income and health", *Journal of Health Economics*, Vol. 15: 67-85.
H.S. Luft (1975), "The Impact of Poor Health on Earnings", *Review of Economics and Statistics*, Vol. 57, pp. 43-57.
J.S. Feinstein, "The relationship between socioeconomic status and health: A Review of the Literature", *The Milbank Quarterly*, Vol. 71(2): 279-322. Available at:
http://www.jonathanfeinstein.com/PDFs/relationship.pdf

12. Testing purchasing power parity*

Purchasing power parity implies that a unit of currency should be able to buy the same basket of goods (a) at home; (b) in another country, after conversion at the going exchange rate. Short-run deviations from PPP are persistent (and obvious!), but the long run is different. There is a vast literature on whether PPP holds in the long run. That is, do market exchange rates exhibit a tendency

to return to PPP in the long run? Rogoff (1996) notes that "While few empirically literate economists take PPP seriously as a short-term proposition, most instinctively believe in some variant of purchasing power parity as an anchor for long-run real exchange rates." But the then immediately adds, "Warm, fuzzy feelings about PPP are not, of course, a substitute for hard evidence." This project will involve time-series techniques such as unit root and cointegration analysis.

Data:
Data on real and nominal exchange rates, inflation, etc. , for multiple countries. May be supplemented.

References:
Charles Engel (1996), "Long-run PPP may not hold after all", *Journal of International Economics*, 57:243-273.
Kenneth Rogoff (1996), "The Purchasing Power Parity Puzzle", *Journal of Economic Literature*, 34(2):647-668.
Alan M. Taylor and Mark P. Taylor (2004), "The Purchasing Power Parity Debate", *Journal of Economic Perspectives*, 18(4):135-158.
Derick Boyd and Ron Smith (2002), "Testing for Purchasing Power Parity: Econometric Issues and an Application to Developing Countries", *Manchester School*, 63(3):287-303.

13. The impact of the Fukushima meltdown: An event study

The tsunami that hit Japan in March 2011 had a major effect on the nuclear power plant at Fukushima, owned by the Toyko Electric Power Company (TEPCO). One purpose of this project is to examine the impact this meltdown had for returns offered by TEPCO. The exercise can then be extended to take in returns offered by other major companies in Japan and elsewhere. The extension of the analysis would explore how the impact of Fukushima varies with the characteristics of particular companies: major electric power companies that own or operate nuclear power plants, power companies that do not have nuclear plants, engineering companies build nuclear power plants, companies that specialise in competing technologies such as renewables, etc.

Data requirements:
Historical prices for both individual companies and stock markets are readily available from various internet sources. A good starting point for TEPCO is the following:
http://quotes.wsj.com/JP/9501/historical-prices

References:
E.R. Berndt (1991), *The Practice of Econometrics: Classic and Contemporary*, Addison-Wesley
J.J.Binder (1998), "The Event Study Methodology Since 1969", *Review of Quantitative Finance and Accounting*, Vol. 11, pp. 111-137.
S.P. Khotari and J.B. Warner (2005), "Econometrics of Event Studies", chapter 1 in B.E. Eckbo, *Handbook of Corporate Finance: Empirical Corporate Finance*, Elsevier/North-Holland.
R.M.Bowen, R.P.Castanias and L.A.Daley (1983), "Intra-Industry Effects of the Accident at Three Mile Island", *Journal of Financial and Quantitative Analysis*, Vol. 18, pp. 87-111.
J. Campbell, A. Lo and A.C. MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton University Press.
R.Ferstl, S.Utz and M.Wimmer (2012), "The Effect of the Japan 2011 Disaster on Nuclear and Alternative Energy Stocks Worldwide: An Event Study", *Business Research Official Open Access Journal*, Vol. 5, pp. 25–41.
A.C.MacKinlay (1997), "Event Studies in Economics and Finance", *Journal of Economic Literature*, Vol. 35, pp. 13-39.
S. Brown and J. Warner (1980), "Measuring Security Price Performance", *Journal of Financial Economics*, 8: 205-258.

S. Brown and J. Warner (1985), "Using Daily Stock Returns: the Case of Event Studies", *Journal of Financial Economics*, 14: 3-31.

14. Weak instruments in country growth regressions*

As Bazzi and Clemens (2013) point out, "One of the great projects of economic research is to establish the causes of economic growth." But as they point out, separating correlation from causality – the endogeneity problem – is a challenging task for empirical research. They survey a range of cross-section and panel-data studies using country-level data that employ IV and GMM methods, and show that the problem of weak and/or invalid instruments is widespread. Their work includes replicating a range of studies and employing various econometric tests. This is an area in which econometric research is advancing rapidly, however, and several tests and estimators are available that Bazzi and Clemens did not use. These include the pre-test for weak instruments of Montiel-Olea and Pflueger (2013) that does not require the assumption of i.i.d. data, and the methods of weak-instrument-robust inference (due to Anderson and Rubin, Moreira, and Kleibergen) that are now available for non-i.i.d. data and can be used with the same sets of instruments and orthogonality conditions as employed by the GMM dynamic panel data estimators that Bazzi and Clemens examine. Your main task in this project is to re-examine the cases that Bazzi and Clemens consider, using these new tests and methods.

Data requirements:
The Bazzi and Clemens data and Stata code are freely available in Stata format at https://www.aeaweb.org/articles.php?doi=10.1257/mac.5.2.152.
The Montiel-Olea and Pflueger test is available in Stata as the command **`weakivtest`**, available from SSC (see below).
The weak-instrument-robust tests are available in Stata as the command **`weakiv`**, available from SSC (see below).

References:
Bazzi, Samuel, and Michael A. Clemens. 2013. "Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth." *American Economic Journal: Macroeconomics*, 5(2): 152-86. DOI: 10.1257/mac.5.2.152
Jose Luis Montiel Olea and Carolin Pflueger, 2013. "A Robust Test for Weak Instruments." *Journal of Business & Economic Statistics*, Vol. 31, No. 3, pp. 358-369.
Jose Luis Montiel Olea, Carolin Pflueger and Su Wang, 2013. "WEAKIVTEST: Stata module to perform weak instrument test for a single endogenous regressor in TSLS and LIML". https://ideas.repec.org/c/boc/bocode/s457732.html. To install in Stata: **`ssc install weakivtest`**.
Keith Finlay, Leandro Magnusson and Mark E. Schaffer, 2013. "WEAKIV: Stata module to perform weak-instrument-robust tests and confidence intervals for instrumental-variable (IV) estimation of linear, probit and tobit models". https://ideas.repec.org/c/boc/bocode/s457684.html. To install in Stata: **`ssc install weakiv`**.

15. The long-run demand for money*

There is a very large literature on the demand for money. Starting around the late 1980s, modern techniques began to be deployed on this topic (accounting for stochastic trends, cointegration, etc.). An early and influential paper on this topic is Stock and Watson (1993), who examined the stability of the demand for money in the US using annual data for the period 1900-89. This paper is discussed in Hayashi (2000) on pp. 659-67, including a section on replicating the Stock-Watson results (pp. 665-67). Your task in this project is to carry out a similar exercise for a country of your choice. The techniques used in this project involve topics and techniques that are not covered in Economics 1 and Econometrics 2 (cointegration, VECMs, DOLS, Johansen's procedure, structural breaks, etc.), and

therefore this topic should be taken up only by groups who either are willing to read up and master their techniques on their own or who already have some expertise in them.

Data requirements:
Stock and Watson use annual data on inflation, money and interest rates for the US covering 90 years (1900-89). You will want to construct a similar dataset for the country of your choice, probably also extended to more recent years.
For replication purposes (getting started) the original Stock-Watson dataset in Stata format is available here: http://ideas.repec.org/p/boc/bocins/swatson93.html

References:
Hayashi (2000), pp. 659-67
J.H. Stock and M.W. Watson (1993), "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems", *Econometrica* 61, pp. 783-820.

16. Business cycles and productivity shocks*

An important question from the "real business cycle" (RBC) literature (now superseded by/incorporated into the DSGE literature): are business cycles mainly the result of permanent shocks to productivity? King, Plosser, Stock and Watson (1991) work with quarterly data from the US covering the period 1947-1988. The paper starts with a review of the implications of balanced growth in the basic neoclassical growth model when combined with productivity shocks. In this model, per capita consumption, investment and output (C, I and Y) all grow at the same rate. Add in a productivity shock and the expected long-run growth path changes. The upshot of this is that the logs of C, I and Y should share a common stochastic trend, i.e., that they should be cointegrated. The question KPSW ask is: are business cycles mostly the result of productivity shocks, or do shocks in nominal variables matter too? Your task in this project is to carry out a similar exercise for a country of your choice. The techniques used in this project involve topics and techniques that are not covered in Economics 1 and Econometrics 2 (cointegration, VECMs, DOLS, Johansen's procedure, structural breaks, etc.), and therefore this topic should be taken up only by groups who either are willing to read up and master their techniques on their own or who already have some expertise in them.

Data requirements:
KPSW work with the following quarterly data for the US for 1947-1988:
log real quarterly consumption, investment, output, money balances (M2) (all per capita), quarterly price inflation, nominal interest rate (both annualised)
You will want to construct a similar dataset for the country of your choice, probably also extended to more recent years.
For replication purposes (getting started), you will be provided with the original KPSW dataset and a Stata do file that replicates some of the results in the paper.

References:
King, Plosser, Stock & Watson (1991), "Stochastic Trends and Economic Fluctuations", *American Economic Review* (Vol. 81, No. 4), September, pp. 819-40

17. Stock markets and oil shocks*

Jones and Kaul (1996) note that oil is not only a key resource for the postwar world economy, it is one for which – they argue – most price shocks have been "exogenous" to the rest of the world economy. Jones and Kaul use this observation to motivate their examination of the relationship between the reaction of international stock markets to oil shocks and whether these can be justified by current and future changes in real cash flows and/or changes in expected returns. Jones and Kaul use postwar data through 1991 for four countries (US, Canada, Japan, UK) and various time series methods

(Granger causality, Granger precedence, etc.). Your task in this project is to re-examine Jones and Kaul's findings in light of some combination of new data, more/different countries, alternative estimation methods, and developments in the literature.

Data requirements:
See Jones and Kaul, pp. 468-9 and 483-486 for the description of the data they used for real cash flows, inflation, oil prices, and various financial series.

References:
Charles M. Jones and Gautam Kau (1996), "Oil and the Stock Markets", *Journal of Finance*, 51(2): 463-491.
Lutz Killian (2008), "The Economic Effects of Energy Price Shocks", *Journal of Economic Literature*, 46(4): 871-909.
And two recent short pieces from the Dallas Fed on the oil price shock of 2002-08:
Michael D. Plante and Mine K. Yücel, "Did Speculation Drive Oil Prices? Futures Market Points to Fundamentals", *Economic Letter*, Vol. 6, No. 10, October 2011.
http://www.dallasfed.org/pages/research/eclett/2011/el1110.cfm
Michael D. Plante and Mine K. Yücel, "Did Speculation Drive Oil Prices? Market Fundamentals Suggest Otherwise", *Economic Letter*, Vol. 6, No. 11, October 2011.
http://www.dallasfed.org/pages/research/eclett/2011/el1111.cfm

18. Interrelated factor demands

This project implements and extends the detailed empirical exercise in Hayashi Chapter 4, pp. 296-307 and 317-320. The translog cost function can be thought of a generalisations of the simple Cobb-Douglas cost function; another way to think about the translog functional form is as giving us a 2nd order Taylor approximation to an arbitrary production function. The exercise is based on the Nerlove (1963) example covered in Hayashi Chapter 1 and in the Stata labs and extended by Christensen and Greene (1976): electricity generation in the US. Estimation in this case requires system estimation: estimation of a system of 3 factor share equations. A full description can be found in Hayashi. The investigation as outlined in Hayashi can be extended further, e.g., by relaxing the assumption of conditional homoskedasticity.

Data requirements:
The dataset of the 99 US electricity utilities is available online from various sources and can be loaded directly into Stata from the SSC data archive:

```
use http://fmwww.bc.edu/ec-p/data/hayashi/cgreene76.dta
```

References:
L. Christensen, D. Jorgenson and L. Lau (1973), "Transcendental Logarithmic Production Frontiers", *Review of Economics and Statistics*, 55, 28-45.
L. Christensen and W. Greene (1976), "Economies of scale in U.S. electric power generation", *Journal of Political Economy*, 84, 655-676.
Further readings provided in Hayashi Chapter 4.

19. Estimating marginal effects in misspecified LDV (limited dependent variable) models: a Monte Carlo study*

Probably the most popular model in econometrics for estimating an LDV model is the probit model. In a probit model, the dependent variable is binary, and the probability that we observe a zero or a one is modelled as $\boldsymbol{P}(\boldsymbol{y_i} = 1 | \boldsymbol{x_i}) = \Phi(\boldsymbol{x_i'}\beta)$ where $\boldsymbol{x_i}$ is a set of independent variables. This is typically estimated using Maximum Likelihood. The coefficients β cannot, however, be interpreted as the

marginal effects of $\boldsymbol{x_i}$ on the probability of observing $\boldsymbol{y_i}$=1; a separate procedure is needed, and in general the magnitude of the marginal effect depends on the values of the different $\boldsymbol{x_i}$. Recently, however, Agrist and Pischke (2009) have argued strongly in favour of the use of the simple Linear Probability Model (LPM). The simple LPM is estimated using OLS with the binary $\boldsymbol{y_i}$ as the dependent variable, and using heteroskedastic-robust standard errors to account for the inherent heteroskedasticity. The advantages of the LPM are simplicity and robustness. In particular, the coefficients β can be interpreted as the marginal effects of $\boldsymbol{x_i}$ on the probability of observing $\boldsymbol{y_i}$=1. But the LPM has a number of disadvantages as well (e.g., the predicted values can be >1 or <0). While the advantage of the probit over the LPM is obvious for the case where the correct distribution is known and corresponds to the assumptions of the probit model, surprisingly little is known about the relative advantages and disadvantages of the probit, LPM and other estimators (e.g., the logit) when the distribution generating the data is not the one assumed by the researcher. This debate spilled over into the blogosphere (your humble lecturer was a participant), with Dave Giles (an econometrician at the University of Victoria, Canada) arguing the case for probit, and Angrist and Pischke arguing the case for the LPM. This project will investigate the performance of the probit, LPM, etc. in obtaining estimates of the average marginal effect (ATE) when the model is correctly specified and under different misspecifications, i.e., when the DGP is not the one assumed by the researcher.

Data:
None required.

References:
Dave Giles' blog, "Econometrics Beat":
http://davegiles.blogspot.co.uk/2012/06/another-gripe-about-linear-probability.html
http://davegiles.blogspot.co.uk/2012/06/yet-another-reason-for-avoiding-linear.html
http://davegiles.blogspot.ca/2012/07/more-comments-on-use-of-lpm.html
Angrist and Pischke's blog, "Mostly Harmless Econometrics":
http://www.mostlyharmlesseconometrics.com/2012/07/probit-better-than-lpm/
Adonis Yatchew and Zvi Griliches (1985), "Specification Error in Probit Models", *Review of Economics and Statistics*, 67(1), 134-39.
Jeffrey M. Wooldridge, *Econometric Analysis of Cross Section and Panel Data*, MIT Press. (a) First edition (2002): Section 15.7, pp. 470-482. (b) Second Edition (2012): Section 15.7, pp. 582-608.
Angrist and Pischke (2009), *Mostly Harmless Econometrics*, Princeton University Press.
Takeshi Amemiya (1977), "Some theorems in the linear probability model", *International Economic Review*, 18, 645-650.
W.C. Horrace and R. L. Oaxaca (2006), "Results on the bias and inconsistency of ordinary least squares for the linear probability model", *Economics Letters*, 90, 321-327. Working paper version available at http://ideas.repec.org/p/wpa/wuwpem/0206002.html

20. MSE as a criterion for a linear regression estimator: a Monte Carlo study

In a series of blog posts, Dave Giles has discussed MSE as a criterion for linear regression. OLS is well-known to be BLUE, i.e., "best" in the class of linear unbiased estimators. But what if we are willing to trade off bias in exchange for a smaller variance? Perhaps we would like to use a linear estimator that is biased but has a smaller MSE (mean-squared error) than OLS. If so, which one should we use? In his first blog post, Giles shows that it is possible to derive a minimum MSE linear estimator (where MSE = sum of the variance and the squared bias of the estimator). But this estimator is infeasible (non-operational) because it depends on the true and ex ante unknown parameters $\boldsymbol{\beta}$ and σ^2. These can be replaced by their OLS estimates, but in that case the estimator is no longer guaranteed to be the minimum MSE linear estimator. In his second blog post, Giles considers the case where the researcher choses an ex ante weight $0 \leq \alpha \leq 1$ and then minimises the weighted sum of the *relative* variance and the *relative* squared bias of the estimator. It turns out that this is a kind of "shrinkage" estimator (such estimators "shrink" OLS towards the origin). It also

turns out that these two "optimal MSE estimators" are closely related. Using the OLS estimates of $\boldsymbol{\beta}$ and σ^2 to implement a "feasible" version of the first estimator is equivalent to choosing a particular α for the second estimator. Giles states that the finite sample performance of these estimators is, to his knowledge, unexplored, and moreover that this would make for a good MSc econometrics project. Your task in this project is to explore the finite sample performance of these estimators, including the design and implementation of the appropriate Monte Carlo experiments.

Data:
None required.

References:
Dave Giles (2013), "A Regression 'Estimator' that Minimizes MSE",
http://davegiles.blogspot.ca/2013/10/a-regression-estimator-that-minimizes.html
Dave Giles (2013), "Beyond MSE – 'Optimal' Linear Regression Estimation",
http://davegiles.blogspot.co.uk/2013/10/beyond-mse-optimal-linear-regression.html
Comments to 2nd blog entry above (where Giles notes this would make a good project!)