DIW Berlin

Deutsches Institut
für Wirtschaftsforschung

www.diw.de

Consultancy to Develop and Implement a
Macroeconomic Model for Lesotho (DIMMoL)

Macro-Econom(etr)ic Modelling

Part 7

Dr. Stefan Kooths
DIW Berlin – Macro Analysis and Forecasting

Course program

- Introduction
- Outline of macroeconom(etr)ic models
- Macroeconomic framework
- **Econometric methodology (cont.)**
- Applied econometrics with EViews
- Lesotho case studies
- Follow-up work

Econometric methodology: Overview

- Fundamentals of probability
- Fundamentals of mathematical statistics
- Principles of regression analysis (cross sections)
- Time series regression models

Conceptional differences to cross sections

- Sequence of random variables indexed by time
  - time series process
  - stochastic process
- Sample = one possible outcome (realization) of the stochastic process
- Sample size = number of time periods observed
- Temporal ordering
- The past can affect the present (and the future)
- Randomness = different historic conditions would have generated a different realization of the observed process
- Population = set of all possible realization of the stochastic process

General strategy

~~random sampling~~

conditions that restrict
temporal correlation in time series

Using OLS in time series analysis

- Case 1: Gauss-Markov-Assumptions
  - strictly exogenous regressors
  - ⇨ OLS estimators are BLUE
- Case 2: Asymptotic Gauss-Markov-Assumptions
  - contemporaneously exogenous regressors
  - weakly dependent time series (asymptotically uncorrelated)
  - ⇨ OLS is consistent, inference methods are asymptotically valid

(trend-) stationary processes

- Case 3: Cointegration analysis
  - strictly exogenous regressors (via leads and lags)
  - highly persistent, cointegrated time series
  - ⇨ OLS is super-consistent, inference methods apply
  - ⇨ error-correction model representation

non-stationary processes

(Trend-) Stationarity

- A process y is stationary if it is identically distributed over time
  - constant mean μ_y
  - constant variance Var(y)
  - constant autocovariance $Cov(y_t, y_{t-h})$
- Trend stationarity
  - stationarity after removing the trend
  - deviations from the trend are stationary

Gauss-Markov assumptions

- **Linearity in parameters**
  population model is characterized by a linear regression function and additive errors
- **No perfect collinearity**
  none of the independent variables is constant nor a perfect linear combination of the others
- **Zero conditional mean (strict exogenity)**
  for each t, the expected value of the error, given the regressors for all time periods, is zero
- **Homoskedasticity**
  error has the same variance given any value of the explanatory variables for all time periods
- **No serial correlation**
  The errors in two different time periods are uncorrelated

OLS estimators are unbiased

OLS estimators are BLUE (Gauss-Markov Theorem)

Why strict exogenity might fail

- Omitted variables
- Measurement errors in some of the regressors
- feedback from the dependent variable on future values of a regressor (policy response)
- Lagged dependent variable as regressor

Assymptotic Gauss-Markov assumptions

- **Linearity and weak dependence**
  population model is characterized by a linear regression function, additive errors, and weakly dependent processes
- **No perfect collinearity**
  none of the independent variables is constant nor a perfect linear combination of the others
- **Zero conditional mean (contemporaneous exogenity)**
  for each t, the expected value of the error, given the regressors in the same period, is zero
- **Homoskedasticity**
  error has the same variance given any contemporaneous value of the explanatory variables
- **No serial correlation**
  The errors in two different time periods are uncorrelated

OLS estimators are consistent

Asymptotic normality of OLS

Weak dependence

- A time series is weakly dependent, if
  - x_t and x_{t+h} are „almost independent“ as h increases without bound (autocorrelation dies out over time)
  - $Cov(x_t, x_{t+h}) \to 0 \quad \text{as} \quad h \to \infty$
- Replaces the assumption of random sampling, making use of the
  - Law of Large Numbers and the
  - Central Limit Theorem

Static and distributed lag Models

- Static models (contemporaneous relationship)

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

- Distributed lag models

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \beta_3 x_{t-2} + u_t$$

- finite distributed lag models
- infinite distributed lag models

⇨ impact propensity (or: impact multiplier)

⇨ long-run propensity (or: long-run multiplier)

Deterministic trends and seasonality

- Trends
  - linear
  - quadratic, cubic (BUT: parsimony condition!)
  - exponential
- Seasonality
  - quarterly: 3 Dummies
  - monthly: 11 Dummies
- Using trending/seasonal variables in regressions
  - including trend and/or seasonal component or
  - removing trends (detrending) and seasonality (seasonal adjustment)

  ⇨ usual inference procedures are asymptotically valid

  ⇨ otherwise: spurious regression problem, artificially high R^2

AR(1) processes

- AR(1) = autoregressive process of order 1

$$y_t = \rho_1 y_{t-1} + e_t$$

- Crucial assumption
  - $\rho < 1$ ⇨ weakly dependent process
    ⇨ integrated of order zero: I(0)
  - $\rho = 1$ ⇨ highly persistent (unit root) process (Random walk)
    ⇨ integrated of order one: I(1)
- Policy implication
  - weakly dependence: policy interventions have temporary effects only
  - high persistence: policy interventions have permanent effects

Estimating the first order autocorrelation

- Case 1: $|\rho| < 1$ (weakly dependent process)
  - regressing y_t on y_{t-1}
  - consistent (but biased) estimator (LLN needed)
- Case 2: $|\rho| = 1$ (unit root process)
  - t-distribution no longer valid
  - Dickey-Fuller tests (based on Monte Carlo Experiments)
- Problem
  - Distribution of the test statistic depends on H_0
  - Both cases might be not rejectable

  ⇨ Power of unit root tests is rather poor

Transforming the AR(1) equation

$$Y_t = \rho \cdot Y_{t-1} + e_t$$

$$\underbrace{Y_t - Y_{t-1}}_{} = \rho \cdot Y_{t-1} + e_t - Y_{t-1}$$

$$\Delta Y_t = (\rho - 1) \cdot Y_{t-1} + e_t$$

$$\Delta Y_t = \gamma \cdot Y_{t-1} + e_t$$

$\uparrow$

$?\quad \gamma = 0$

(Augmented) Dickey-Fuller tests

- Three scenarios
  - Random walk: $\Delta y_t = \gamma y_{t-1} + e_t$
  - Random walk with drift: $\Delta y_t = a_0 + \gamma y_{t-1} + e_t$
  - Random walk with drift and trend: $\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + e_t$
- Scenarios may include lags of Δy (Augmented DF)
  - e.g. $\Delta y_t = \gamma y_{t-1} + \theta_1 \Delta y_{t-1} + \theta_1 \Delta y_{t-2} + e_t$
- Critical values t_c (tabulated) depend on
  - scenario type
  - sample size
- Testing for $\gamma = 0$ (H_0: existence of a unit root)
- Rejection rule: Reject H_0 if $t < t_c$

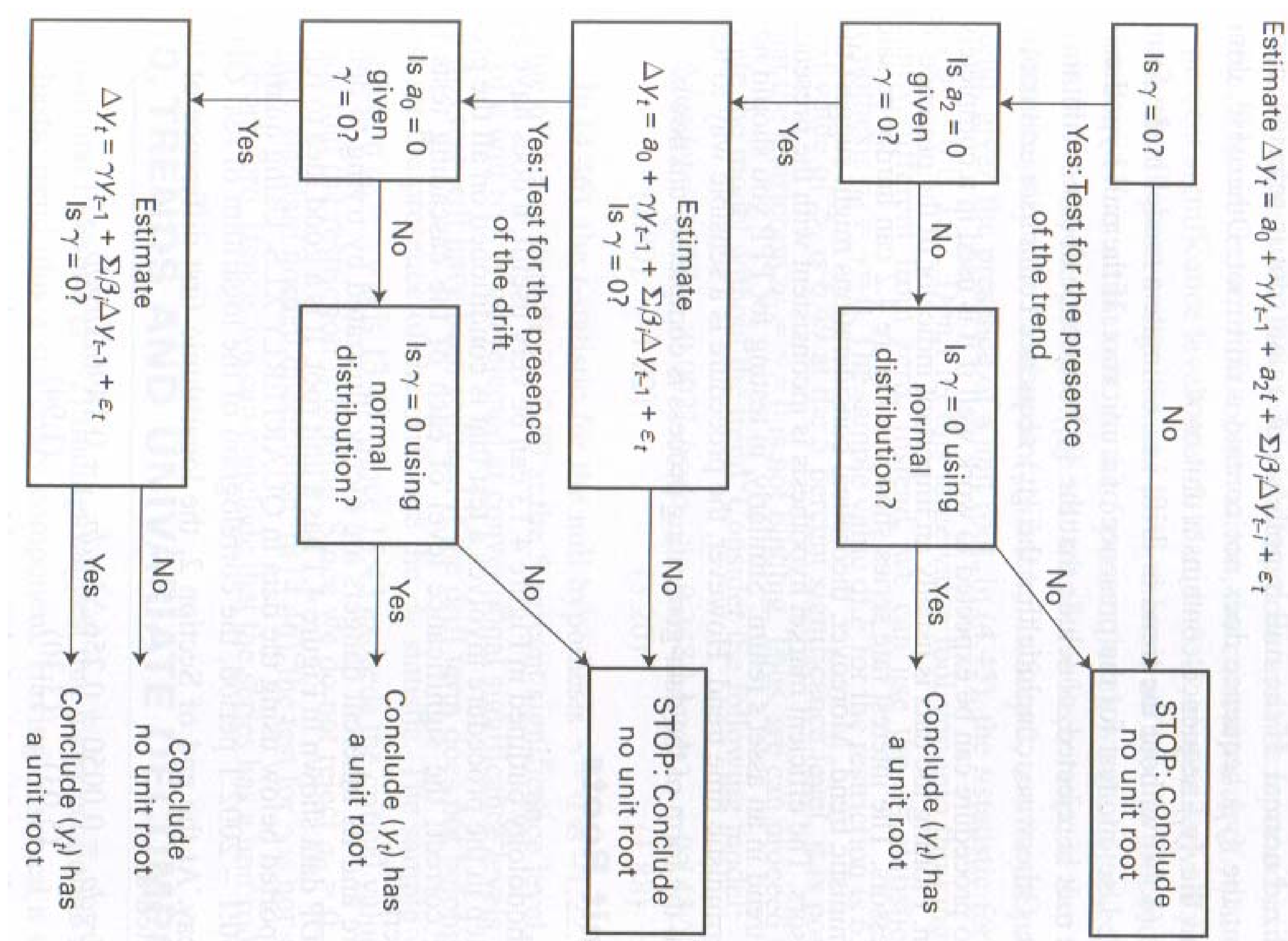
Critical values for Dickey-Fuller test

$\gamma = 0$

Sample Size T	Significance level 0.01	0.025	0.05	0.10
The τ statistic: No Constant or Time Trend ($a_0 = a_2 = 0$)				
25	–2.66	–2.26	–1.95	–1.60
50	–2.62	–2.25	–1.95	–1.61
100	–2.60	–2.24	–1.95	–1.61
250	–2.58	–2.23	–1.95	–1.62
300	–2.58	–2.23	–1.95	–1.62
∞	–2.58	–2.23	–1.95	–1.62
The τ_μ statistic: Constant but No Time Trend ($a_2 = 0$)				
25	–3.75	–3.33	–3.00	–2.62
50	–3.58	–3.22	–2.93	–2.60
100	–3.51	–3.17	–2.89	–2.58
250	–3.46	–3.14	–2.88	–2.57
500	–3.44	–3.13	–2.87	–2.57
∞	–3.43	–3.12	–2.86	–2.57
The τ_τ statistic: Constant + Time Trend				
25	–4.38	–3.95	–3.60	–3.24
50	–4.15	–3.80	–3.50	–3.18
100	–4.04	–3.73	–3.45	–3.15
250	–3.99	–3.69	–3.43	–3.13
500	–3.98	–3.68	–3.42	–3.13
∞	–3.96	–3.66	–3.41	–3.12

Sample size T	Significance level 0.10	0.05	0.025	0.01
$\gamma = a_0 = 0$ — ϕ_1				
25	4.12	5.18	6.30	7.88
50	3.94	4.86	5.80	7.06
100	3.86	4.71	5.57	6.70
250	3.81	4.63	5.45	6.52
500	3.79	4.61	5.41	6.47
∞	3.78	4.59	5.38	6.43
$\gamma = a_0 = a_2 = 0$ — ϕ_2				
25	4.67	5.68	6.75	8.21
50	4.31	5.13	5.94	7.02
100	4.16	4.88	5.59	6.50
250	4.07	4.75	5.40	6.22
500	4.05	4.71	5.35	6.15
∞	4.03	4.68	5.31	6.09
$\gamma = a_2 = 0$ — ϕ_3				
25	5.91	7.24	8.65	10.61
50	5.61	6.73	7.81	9.31
100	5.47	6.49	7.44	8.73
250	5.39	6.34	7.25	8.43
500	5.36	6.30	7.20	8.34
∞	5.34	6.25	7.16	8.27

General-to-specific procedure for testing unit roots



Unit root processes in regression analysis

- Time series x_t, y_t are I(1) processes
  - also applies to higher identical orders of integration and more than two variables
- Case 1: No cointegration
  - any linear combination of x_t and y_t is I(1)
  - ⇨ problem of spurious regression
  - ⇨ first differences as transformation method
- Case 2: Cointegration
  - an linear combination of x_t and y_t (cointegration vector) exists such that $s_t = y_t - \beta x_t$ is I(0)
  - ⇨ OLS estimators show long-run equilibrium relationship
  - ⇨ error-correction model for short-run adjustment dynamics (Granger representation theorem)
- Test for cointegration: Engle-Granger cointegration test

Testing for cointegration: The Engle-Granger Methodology

- Step 1: Test x_t and y_t for integration
  - use Dickey-Fuller test
  - EXIT if both series are stationary or integrated of different orders (= no cointegration)
- Step 2: Estimate long-run equilibrium relationship

$$y_t = \beta_0 + \beta_1 x_t + s_t$$

- Step 3: Check residuals for stationarity

$$\Delta \hat{s}_t = a_1 \hat{s}_{t-1} + e_t$$

  - special critical values apply
  - EXIT if H_0: $a_1 = 0$ cannot be rejected

Testing for cointegration: The Engle-Granger Methodology (cont.)

- Step 4: Estimate the error-correction model

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta y_{t-1} + \eta_0 \Delta x_t + \eta_1 \Delta x_{t-1} + \delta \boxed{(y_{t-1} - \beta_1 x_{t-1})} + u_t$$

- all variables are I(0), therefore OLS is valid
- further lags of Δy and Δx may apply (check u_t for white noise)
- use residuals from step 3 for $(y_{t-1} - \beta x_{t-1})$:

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta y_{t-1} + \eta_0 \Delta x_t + \eta_1 \Delta x_{t-1} + \delta \hat{s}_{t-1} + u_t$$

Critical values for Engle-Granger Cointegratoin test

T	1%	5%	10%	1%	5%	10%
	Two Variables			Three Variables		
50	–4.123	–3.461	–3.130	–4.592	–3.915	–3.578
100	–4.008	–3.398	–3.087	–4.441	–3.828	–3.514
200	–3.954	–3.368	–3.067	–4.368	–3.785	–3.483
500	–3.921	–3.350	–3.054	–4.326	–3.760	–3.464
	Four Variables			Five Variables		
50	–5.017	–4.324	–3.979	–5.416	–4.700	–4.348
100	–4.827	–4.210	–3.895	–5.184	–4.557	–4.240
200	–4.737	–4.154	–3.853	–5.070	–4.487	–4.186
500	–4.684	–4.122	–3.828	–5.003	–4.446	–4.154

The critical values are for cointegrating relations (with a constant in the cointegrating vector) estimated using the Engle-Granger methodology.

Other topics in time series analysis

- Serial correlation, Autokorrelationsfunktion
- ARMA (Box-Jenkins) and ARIMA models
- ARCH processes
- Vector autoregressive models (VAR), interventions and impulse-response analysis
- Structural change
- Non-linear time series models

Course program

- Introduction
- Outline of macroeconom(etr)ic models
- Macroeconomic framework
- Econometric methodology (cont.)
- Applied econometrics with EViews
- Lesotho case studies
- Follow-up work

Course program

- Introduction
- Outline of macroeconom(etr)ic models
- Macroeconomic framework
- Econometric methodology (cont.)
- Applied econometrics with EViews
- Lesotho case studies
- Follow-up work

Work groups

- Private domestic demand
  - private consumption
  - private investment
  - income
- Fiscal affairs
  - public consumption
  - public investment
  - taxation
  - subsidies
  - budgets and MTEF
- External relations and monetary issues
  - trade flows
  - capital and transfer flows
  - real effective exch. rate
  - interest rate forecasts and money demand
- Production and Pricing
  - production function
  - labor demand and wage setting
  - capital accumulation

General tasks (all groups)

- Economic theory and literature review
- Model formulation in African economies
- Functional form specification
- Data base checks
- Preliminary estimation of equations

- Regular work group meetings
- Collective macro level discussions
- Remote assistence from DIW Berlin