Cheat Sheet for the Exam:
Econometric Methods

1 A-, B- and C-Assumptions

Assumption A1: (The true model) The econometric model does not lack any relevant exogenous variables and the exogenous variables used are not irrelevant.

Assumption A2 (Linear in parameters) The true relationship between $\mathbf{X}$ and $\mathbf{y}$ is linear.

Assumption A3 (Constant parameters) The parameters β $(\alpha, \beta_1, \beta_2, \ldots, \beta_K)$ are constant for all T observations $(\mathbf{x}_t, y_t)$.

Assumption B1 $E(\mathbf{u}) = \mathbf{o}$

Assumption B2 $var(u_t) = \sigma^2$ $, fort = 1, 2, ..., T.$ or $\mathbf{V}(\mathbf{u}) = \sigma^2 \mathbf{I}_T$

Assumption B3 $cov(u_t, u_s) = 0$, for all $t \neq s$ where $t = 1, 2, ..., T$ und $s = 1, 2, .., T.$ or $\mathbf{V}(\mathbf{u}) = \sigma^2 \mathbf{I}_T$

Assumption B4 The disturbance u_t is normally distributed.

$$\mathbf{u} \sim N(E(\mathbf{u})\ ,\ \mathbf{V}(\mathbf{u}))$$

Assumption C1 (Fixed exogenous variables) None of the elements of the matrix $\mathbf{X}$ (the exogenous variables $x_{1t}, x_{2t}, \ldots, x_{Kt}$) is a random variable, but can be controlled as in an experiment.

Assumption C2 (Free from perfect multicollinearity) The matrix $\mathbf{X}$ has full rank, i.e. all column vectors are linearly independent. So there is no linear relationship between the columns or in other words, there is no linear relationship between the exogenous variables where at least one parameter $\gamma_k \neq 0$.

2 Simple Regression Model

Estimation

$$\begin{aligned}\widehat{\alpha} &= \overline{y} - \widehat{\beta}\overline{x} \\ \widehat{\beta} &= S_{xy}/S_{xx}\end{aligned}$$

$$\begin{aligned}S_{yy} &\equiv \sum (y_t - \overline{y})(y_t - \overline{y}) \\ S_{xx} &\equiv \sum (x_t - \overline{x})(x_t - \overline{x}) \\ S_{xy} &\equiv \sum (x_t - \overline{x})(y_t - \overline{y})\end{aligned}$$

$$S_{yy} = S_{\widehat{u}\widehat{u}} + S_{\widehat{y}\widehat{y}}$$

$$R^2 = S_{\widehat{y}\widehat{y}}/S_{yy}$$

$$\begin{aligned}E(\widehat{\alpha}) &= \alpha \\ E\left(\widehat{\beta}\right) &= \beta\end{aligned}$$

$$\widehat{\sigma}^2 = S_{\widehat{u}\widehat{u}}/(T-2)$$

$$\left[\widehat{\beta} - t_{a/2} \cdot \widehat{se}(\widehat{\beta}) \ ; \ \widehat{\beta} + t_{a/2} \cdot \widehat{se}(\widehat{\beta})\right]$$
$$\left[\widehat{\alpha} - t_{a/2} \cdot \widehat{se}(\widehat{\alpha}) \ ; \ \widehat{\alpha} + t_{a/2} \cdot \widehat{se}(\widehat{\alpha})\right]$$

Hypothesis Testing (two sided t-test)

$$\begin{aligned}\Pr\left\{-t_{a/2} \leq t \leq t_{a/2}\right\} &= 1 - a \\ \textit{where} \qquad t &= (\widehat{\beta} - q)\Big/\widehat{se}(\widehat{\beta})\end{aligned}$$

3 Multiple Regression Analysis

Estimation

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u} \ ,$$

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + ... + \beta_K x_{Kt} + u_t \ .$$

$$\widehat{\beta} = \left(\mathbf{X}'\mathbf{X}\right)^{-1} \mathbf{X}'\mathbf{y} \ .$$

$$\widehat{\sigma}^2 = S_{\widehat{u}\widehat{u}}/(T - K - 1)$$

Variance-covariance matrix of the error terms

$$\mathbf{V}(\mathbf{u}) = \mathbf{E}[\mathbf{u}\mathbf{u}'] = \begin{bmatrix} var(u_1) & cov(u_1, u_2) & \cdots & cov(u_1, u_T) \\ cov(u_2, u_1) & var(u_2) & \cdots & cov(u_2, u_T) \\ \vdots & \vdots & \ddots & \vdots \\ cov(u_T, u_1) & cov(u_T, u_2) & \cdots & var(u_T) \end{bmatrix} = \sigma^2 \mathbf{I}_T$$

$$R^2 = \frac{S_{yy} - S_{\widehat{u}\widehat{u}}}{S_{yy}} = \frac{S_{\widehat{y}\widehat{y}}}{S_{yy}}$$

$$\begin{aligned} E(\widehat{\beta}_1) &= \beta_1 \\ E(\widehat{\beta}_2) &= \beta_2 \\ E(\widehat{\alpha}) &= \alpha \end{aligned}$$

Variance-covariance matrix of the estimators

$$C(\widehat{\beta}) = \begin{bmatrix} var(\widehat{\alpha}) & cov(\widehat{\alpha}, \widehat{\beta}_1) & \cdots & cov(\widehat{\alpha}, \widehat{\beta}_K) \\ cov(\widehat{\beta}_1, \widehat{\alpha}) & var(\widehat{\beta}_1) & \cdots & cov(\widehat{\beta}_1, \widehat{\beta}_K) \\ \vdots & \vdots & \ddots & \vdots \\ cov(\widehat{\beta}_K, \widehat{\alpha}) & cov(\widehat{\beta}_K, \widehat{\beta}_1) & \cdots & var(\widehat{\beta}_K) \end{bmatrix} = \sigma^2 \left(\mathbf{X}'\mathbf{X}\right)^{-1} \ .$$

Prediction

$$\widehat{y}_0 = \mathbf{x}_0'\widehat{\beta} \ .$$

$$var(\widehat{y}_0 - y_0) = \sigma^2 \left[1 + \mathbf{x}_0' \left(\mathbf{X}'\mathbf{X}\right)^{-1} \mathbf{x}_0\right] \ .$$

Hypothesis Testing (F-Test)

$$F = \frac{\left(S^0_{\widehat{u}\widehat{u}} - S_{\widehat{u}\widehat{u}}\right)/L}{S_{\widehat{u}\widehat{u}}/\left(T - K - 1\right)} \sim F_{(L, T-K-1)}$$

4 Violation of Assumptions

Assumption A1:

Omitting a relevant variable:

$$E(\widehat{\widetilde{\beta}}_1) = \beta_1 + (\mathbf{X}_1'\mathbf{X}_1)^{-1}\,\mathbf{X}_1'\mathbf{X}_2\beta_2 \neq \beta_1$$

$$\widetilde{\sigma}^2 = E[(X_2\beta_2)^2] + \sigma^2$$

Using an irrelevant variable:

$$E(\widehat{\beta}) = E(\beta)$$

$$\widetilde{\sigma}^2 = \sigma^2$$

Adjusted coefficient of determination $\overline{R}^2$

$$\begin{aligned} \overline{R}^2 &= 1 - \frac{S_{\widehat{u}\widehat{u}}\,/(T-K-1)}{S_{yy}\,/(T-1)} \\ &= 1 - \left(1 - R^2\right)\frac{T-1}{T-K-1} \end{aligned}$$

Other criteria

$$\begin{aligned} AIC &= \ln\left(\frac{S_{\widehat{u}\widehat{u}}}{T}\right) + \frac{2(K+1)}{T} \\ SC &= \ln\left(\frac{S_{\widehat{u}\widehat{u}}}{T}\right) + \frac{(K+1)\ln T}{T} \\ PC &= \frac{S_{\widehat{u}\widehat{u}}\left[1 + (K+1)\,/T\right]}{T-K-1} \end{aligned}$$

Other diagnostic tools: t-test, F-test, non-nested F-test, J-test.

Assumption A2:

Nonlinear regression functions

$$\begin{aligned}
\ln y_t &= \alpha + \beta \ln x_t + u_t && (logarithmic) \\
y_t &= \alpha + \beta \ln x_t + u_t && (semi-logarithmic) \\
\ln y_t &= \alpha + \beta x_t + u_t && (exponential) \\
\ln y_t &= \alpha + \beta\,(1/x_t) + u_t && (log-inverse) \\
y_t &= \alpha + \beta\,(1/x_t) + u_t && (inverse) \\
y_t &= \alpha + \beta_1 x_t + \beta_2 x_t^2 + u_t && (quadratic)
\end{aligned}$$

Zaremba's Box-Cox-Test

$$l = \frac{T}{2}\left|\ln\left(\frac{S_{\widehat{u}\widehat{u}}/\tilde{y}^2}{S^*_{\widehat{u}\widehat{u}}}\right)\right| \sim \chi^2_{(1)} ,$$

$$where \quad S^*_{\widehat{u}\widehat{u}} = Sum\, of\, residual\, squares\, of\, the\, model\, with\, \ln y_t^*$$

Regression Specification Error Test (RESET)

$$F = \frac{\left(S'_{\widehat{u}\widehat{u}} - S^*_{\widehat{u}\widehat{u}}\right)/L}{S^*_{\widehat{u}\widehat{u}}/\left(T - K^* - 1\right)} \sim F_{(L,T-K^*-1)}$$

Another diagnostic tool: R^2

Assumption A3:

Diagnostic tools: F-test, t, t-test

Prognostic Chow-test:

$$F = \frac{(S^*_{\widehat{u}\widehat{u}} - S^I_{\widehat{u}\widehat{u}})/T_{II}}{S^I_{\widehat{u}\widehat{u}}/(T_I - K - 1)} .$$

Assumption B2:

Goldfeld-Quandt Test

$$F = \frac{S^{II}_{\widehat{u}\widehat{u}}\,/(T_{II} - K - 1)}{S^{I}_{\widehat{u}\widehat{u}}\,/(T_I - K - 1)},$$

where $S^I_{\widehat{u}\widehat{u}}$ and $S^{II}_{\widehat{u}\widehat{u}}$ are the sum of residual squares of groups I and II.

White-Test

$$R^2 T \sim \chi^2_{(v)} ,$$

Breusch-Pagan-Test

$$g_t = \frac{\widehat{u}_t^2}{\widehat{\sigma}_t^2} \qquad BP = \frac{S_{\widehat{g}\widehat{g}}}{2}$$

or alternatively

$$R^2 T \sim \chi^2_{(v)} ,$$

where ν=Number of slope parameters of the auxiliary model and R^2 of the auxiliary model.

Assumption B3:
AR(1)-Process :

$$\begin{aligned}
u_t &= \rho u_{t-1} + e_t \ , \qquad -1 < \rho < 1 \\
E(u_t) &= \sum_{j=0}^{\infty} \rho^j E(e_{t-j}) = 0 \\
var(u_t) &= \frac{\sigma_e^2}{1-\rho^2} \equiv \sigma^2 \\
cov(u_t, u_{t-\tau}) &= \rho^\tau \left(\frac{\sigma_e^2}{1-\rho^2} \right) = \rho^\tau \sigma^2 \neq 0
\end{aligned}$$

Estimator for ρ

$$\widehat{\rho} = \frac{\sum_{t=2}^{T} \widehat{u}_{t-1} \widehat{u}_t}{\sum_{t=2}^{T} \widehat{u}_{t-1}^2}.$$

Durbin-Watson Test

$$\begin{aligned}
d &= \frac{\sum_{t=2}^{T} (\widehat{u}_t - \widehat{u}_{t-1})^2}{\sum_{t=1}^{T} \widehat{u}_t^2} \\
&\approx 2(1-\widehat{\rho})
\end{aligned}$$

Durbin's h-Test

$$h = \widehat{\rho} \sqrt{\frac{T}{1 - T \cdot \widehat{var}(\widehat{\beta}_2)}}$$

Breusch-Godfrey-Test

$$BG = TR^2 \sim \chi^2_{(K)}$$

5 Panel Estimation

Panel Estimation

Pooled Model

$$y_{i,t} = \alpha + \beta_1 x_{1i,t} + \beta_2 x_{2i,t} + ... + \beta_K x_{Ki,t} + u_{i,t}$$

Fixed-Effects-Model

$$y_{i,t} = \alpha_i + \beta_1 x_{1i,t} + \beta_2 x_{2i,t} + ... + \beta_K x_{Ki,t} + u_{i,t}$$

Random-Effects-Model

$$y_{i,t} = \beta_1 x_{1i,t} + \beta_2 x_{2i,t} + ... + \beta_K x_{Ki,t} + \varepsilon_{i,t}$$

$$\varepsilon_{i,t} = u_{i,t} + a_i$$