Self-Entry-Level Test

Institute of Economics
Institute for Statistics and Econometrics
Kiel Universtiy, Germany

Dear Candidate,

the following questions are not an examination in the usual sense, but merely a voluntary test. The main purpose of this test is to help YOU to find out whether you might be a successful participant in our master's degree programme. The knowledge required to answer these questions - without much looking-up and without any (!) learning - is what we expect you to bring with you, as part of the skills acquired in your bachelor studies.
In the application process for any of our master's degree programmes, we can only evaluate you based on your grades and credit points. We cannot take into account your answers to these questions since we do not know the amount of help you might have used. But if you fail to answer 40% of the test questions easily and decide to come to Kiel anyway, the probability of you leaving the programme after the first examinations without success is increased - wasting valuable time and money.

So, please be honest to yourself when answering the following questions!

Questions of the following sections are relevant for the single programmes:

Economics:
Mathematics, Econometrics, Statistics, Microeconomics, and Macroeconomics
Quantitative Economics:
Mathematics, Econometrics, Statistics, Microeconomics, and Macroeconomics
Quantitative Finance:
Mathematics, Econometrics, and Statistics
Environmental and Resource Economics:
Mathematics, Econometrics, and Microeconomics

Mathematical Questions

1. Determine the solution to the following integrals:

   (a) $\int 2^x + e^{x/8} + \cos(2x)dx$

   (b) $\int \sqrt{4x}^{-2/3} + \frac{4}{x^4}dx, x > 0$

   (c) $\int_1^{10} \exp(-\ln(\frac{a}{x}))dx, a > 0$

   (d) $\int_0^b \frac{1}{\exp(2x)}dx, b > 0$

   (e) $\int_0^\pi \cos(x)dx$

2. Determine the solution to the following integrals by partial integration:

   (a) $\int_{1/2}^1 \ln(4x^2)dx$

   (b) $\int x^2\theta \exp(-\theta x)dx, \theta > 0, x > 0$

3. Obtain the first and second order derivatives for the following functions at t=0:

   (a) $f(t) = (pe^t + 1 - p)^n, n \in N, p \in (0, 1)$

   (b) $f(t) = \exp(\mu t + \frac{1}{2}\sigma^2 t^2), \sigma > 0$

4. Obtain the first order derivatives for the following functions:

   (a) $f(t) = \ln(\sqrt[3]{(t-a)(t+a)}), |t| > |a|$

   (b) $f(t) = \frac{1}{(t-1)^a}, t > 1$

   (c) $f(t) = \cos(\frac{4}{\sqrt{t}}), t > 0$

   (d) $f(t) = \sin(\cos(t))$

   (e) $f(t) = \ln(\ln(t)), t > 1$

5. Find the limit for $n \in \mathbb{N}$ or $x \in \mathbb{R}$, if it exists:

   (a) $\lim\limits_{n\to\infty} \left(\frac{\left(\frac{1}{2}\right)^n + 2 \cdot 4^n}{5 \cdot 4^n - 2 \cdot 3^n} \right)$

   (b) $\lim\limits_{n\to\infty} \left(\frac{\sqrt[3]{n^4} + n^2 - 1}{(n - 2\sqrt{n})^2 + 2} \right)$

   (c) $\lim\limits_{n\to\infty} \frac{(-1)^n 2n}{2n^3 - n^2}$

   (d) (using de l'Hôpital's rule) $\lim\limits_{x\to\infty} x^2 e^{-x}$

6. Let $f(x) = 2^{1-x}$.

   (a) Find the derivative of $f(x)$ up to order 3.

(b) Give the n-th derivative $f^{(n)}(x)$ for $n \in \mathbb{N}_0$.

(c) Find the Taylor polynomial $T_f^2(x)$ of order 2 in $x_0 = 0$.

(d) Give the Taylor series $T_f^\infty(x)$ in $x_0 = 0$.

7. Consider the following Cobb-Douglas production function as given:

$$f(K, L) = aK^bL^c, a, b, c > 0; K > 0, L > 0$$

The partial derivatives are given by $f'_K(K, L) = abK^{b-1}L^c$ and $f'_L(K, L) = acK^bL^{c-1}$.

(a) Calculate the partial differential df_K with respect to K in (K_0, L_0)=(1,4).

(b) Calculate the total differential df in $(K_0, L_0) = (1, 4)$ for $c = \frac{1}{2}$.

8. Use the Lagrangian method to determine the critical points of $f(x, y) = x^2 + y^2$ subject to the constraint $y = \frac{1}{x}, x > 0$.

9. Use the Lagrangian method to solve the critical points of the Cobb-Douglas utility function $U(x, y) = ax^by^{1-b}, x, y, a > 0, 0 < b < 1$ subject to the budget restriction $g(x, y) = 2x + 3y = 10$.

10. Use the Lagrangian method to determine the critical points of the utility function $U(x_1, x_2) = (x_1-1)^{0.4}(x_2-2)^{0.8}, x_1 > 1, x_2 > 2$ subject to the budget constraint $g(x_1, x_2) = x_1+2x_2 = 11$.

11. Consider $f(x, y) = \exp(-\frac{2}{x} + y)$ with $x \neq 0$.

(a) Determine the first and second order partial derivatives.

(b) Calculate the total differential df in $(x_0, y_0) = (2, 1)$.

12. Consider $f(x, y) = x^2 + y^2 + 4$ and $g(x, y) = 2x^3 + 2y^3 + x^2 + y^2$ with $x, y \neq 0$. Use the Lagrangian method to solve for the critical points of f subject to the constraint that $g(x, y) = 0$.

13. Calculate the determinants of the following matrices:

(a) $\mathbf{A} = \begin{pmatrix} 1 & -3 & 4 \\ -4 & 1 & 3 \\ 2 & -2 & 3 \end{pmatrix}$

(b) $\mathbf{B} = \begin{pmatrix} 1 & -3 & 4 \\ 3 & 2 & 0 \\ 3 & 1 & 2 \end{pmatrix}$

(c) $\mathbf{C} = \begin{pmatrix} 2a & -1 & 1 \\ 3a & 0 & 3 \\ 0 & 1 & 1 \end{pmatrix}$

(d) $\mathbf{D} = \begin{pmatrix} -1 & 2 & 3 & 4 \\ 2 & 0 & 0 & 4 \\ 1 & 1 & 2 & 1 \\ 4 & 3 & -1 & 7 \end{pmatrix}$

14. Determine the eigenvalues of the following matrices:

(a) $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix}$

(b) $\mathbf{B} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

(c) $\mathbf{C} = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 8 \end{pmatrix}$

(d) $\mathbf{D} = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}$

(e) $\mathbf{E} = \begin{pmatrix} 2 & -2 & 3 \\ 0 & 3 & -2 \\ 0 & -1 & 3 \end{pmatrix}$

15. Let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 4 & -6 \\ -2 & 3 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix},$$

$$\mathbf{D} = \begin{pmatrix} -2 & 7 \\ 5 & -1 \end{pmatrix}, \quad \mathbf{E} = \begin{pmatrix} 1 & 1 \end{pmatrix}.$$

Perform the following operations, if possible:

(a) $\mathbf{A} \cdot \mathbf{A}$

(b) $\mathbf{A} \cdot \mathbf{B}$

(c) $\mathbf{A} \cdot \mathbf{C}$

(d) $\mathbf{A} \cdot \mathbf{D}$

(e) $\mathbf{E}^T \cdot \mathbf{A} \cdot \mathbf{E}$

(f) $\mathbf{E} \cdot \mathbf{A} \cdot \mathbf{E}^T$

16. Find the inverses of the following matrices:

(a) $\mathbf{A} = \begin{pmatrix} 1 & -3 \\ -3 & 5 \end{pmatrix}$

(b) $\mathbf{B} = \begin{pmatrix} 1 & 3 & 2 \\ 3 & 5 & 0 \\ 0 & 5 & 7 \end{pmatrix}$

(c) $\mathbf{C} = \begin{pmatrix} 1 & 2 & 4 \\ 8 & 6 & 3 \\ 8 & 6 & 2 \end{pmatrix}$

Mathematical Questions: RESULTS

1. (a) $= 2^x \frac{1}{\ln(2)} + 8e^{x/8} + 0.5\sin(2x) + c$

   (b) $= 0.9449x^{2/3} - \frac{4}{3}x^{-3} + c$

   (c) $= \frac{99}{2a}$

   (d) $= 0.5(1 - e^{-2b})$

   (e) $= 0$

2. (a) $= 0.3863$

   (b) $= -e^{-\theta x}(x^2 + \frac{2x}{\theta} + \frac{2}{\theta^2})$

3. (a) $f'(t)|_{t=0} = np$, $f''(t)|_{t=0} = n^2p^2 + np(1-p)$

   (b) $f'(t)|_{t=0} = \mu$, $f''(t)|_{t=0} = \mu^2 + \sigma^2$

4. (a) $= \frac{2t}{3(t^2-a^2)}$

   (b) $= -a(t-1)^{-a-1}$

   (c) $= 2\sin(\frac{4}{\sqrt{t}})t^{-3/2}$

   (d) $= -\cos(\cos(t))\sin(t)$

   (e) $= \frac{1}{t\ln(t)}$

5. (a) $= 0.4$

   (b) $= 1$

   (c) $= 0$

   (d) $= 0$

6. (a) $f'(x) = -2^{1-x}\ln(2)$, $f''(x) = 2^{1-x}(\ln(2))^2$, $f'''(x) = -2^{1-x}(\ln(2))^3$

   (b) $= 2^{1-x}(-\ln(2))^n$

   (c) $= 2 - 2\ln(2) + (\ln(2))^2$

   (d) $= \sum_{k=0}^{\infty} \frac{(-\ln(2)x)^k}{k!}$

7. (a) $= ab4^c\Delta K$

   (b) $= 2ab\Delta K + \frac{a}{4}\Delta L$

8. $x = y = \pm 1, \lambda = \pm 2$

9. $x = 5b, y = \frac{10}{3}(1-b)$

10. $x_1 = 3, x_2 = 4, \lambda = 0.4595$

11. (a) $f'_x(x,y) = 2f(x,y)x^{-2}, f'_y(x,y) = f(x,y)$
$f''_{xx}(x,y) = 4f(x,y)(x^{-4} - x^{-3}), f''_{yy}(x,y) = f(x,y), f''_{xy}(x,y) = f'_x(x,y)$

(b) $= 0.5\Delta x + \Delta y$

12. $x = y = -0.5, \lambda = 2$

13. (a) $= -21$

(b) $= 10$

(c) $= 0$

(d) $= 108$

14. (a) $\lambda_{1,2} = \pm 2$

(b) $\lambda_1 = 3, \lambda_2 = 1$

(c) $\lambda_1 = 8, \lambda_2 = 7, \lambda_3 = 6$

(d) $\lambda_{1,2,3} = 0$

(e) $\lambda_1 = 2, \lambda_{2,3} = 3 \pm \sqrt{2}$

15. (a) $= \begin{pmatrix} 5 & 10 \\ 10 & 20 \end{pmatrix}$

(b) $= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

(c) $= \begin{pmatrix} 8 & 5 \\ 16 & 10 \end{pmatrix}$

(d) $= \begin{pmatrix} 8 & 5 \\ 16 & 10 \end{pmatrix}$

(e) not defined

(f) $= 9$

16. (a) $= -\frac{1}{4}\begin{pmatrix} 5 & 3 \\ 3 & 1 \end{pmatrix}$

(b) $= \begin{pmatrix} 17.5 & -5.5 & -5 \\ -10.5 & 3.5 & 3 \\ 7.5 & -2.5 & -2 \end{pmatrix}$

(c) $= \begin{pmatrix} -0.6 & 2 & -1.8 \\ 0.8 & -3 & 2.9 \\ 0.0 & 1 & -1.0 \end{pmatrix}$

Econometric Questions

1. What is a cumulative distribution function, a probability distribution function (for discrete random variables), and a probability density function (for continuous random variables)?

2. What is an expected value, a variance and a standard deviation of a random variable Y?

3. Characterize a joint and a conditional probability distributions.

4. What is the normal distribution, the t distribution, the F distribution, and the χ^2 distribution?

5. What means convergence in probability and convergence in distribution?

6. What is simple random sampling? What means that the observations $Y_1, ..., Y_n$ of a random sample are independently and identically distributed (i.i.d.)?

7. How is the sample average $\bar{Y}$ computed?

8. Show that if $Y_1, ..., Y_n$ are i.i.d. with mean μ_Y and variance σ_Y^2 then:

   (a) $\bar{Y}$ is an unbiased and consistent estimator of the population mean,

   (b) the sampling distribution of $\bar{Y}$ has mean μ_Y and variance $\sigma_{\bar{Y}}^2 = \sigma_Y^2/n$,

   (c) the law of large numbers says that $\bar{Y}$ converges in probability to μ_Y, and

   (d) the central limit theorem says that the standardized version of $\bar{Y}$,$(\bar{Y} - \mu_Y)/\sigma_{\bar{Y}}$, has a standard normal distribution when n is large.

9. What is the difference between population and sample regression line?

10. What is this: $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$?

11. Explain the use of $\mathbf{P_X} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ and $\mathbf{M_X} = \mathbf{I_n} - \mathbf{P_X}$.

12. Show that if (1) the regression errors, u_i, have a mean of zero conditional on the regressors X_i, (2) the sample observations are i.i.d. random draws from the population, (3) the regressors are not linearly dependent, and (4) large outliers are unlikely, then the OLS estimators of the population model $Y_i = X_i\beta + u_i$ are unbiased, consistent, and asymptotically normally distributed.

13. What is the purpose and meaning of R^2, adjusted R^2, and the standard error of the regression (SER)?

14. Show that the squared SER can be written a $s_{\hat{u}}^2 = \frac{\mathbf{U'M_XU}}{n-k-1}$.

15. How can single and joint hypotheses be tested using t and F-statistics? What is the p-value?

16. Construct a 95% confidence interval for a regression coefficient.

17. What is the difference between traditional and heteroscedasticity-consistent standard errors?

18. State the Gauss-Markov theorem.

19. Show that if the assumptions of question 12 hold and the regression errors are additionally homoscedastic and normally distributed, then:

    (a) $\mathbf{U}|\mathbf{X} \sim \mathcal{N}(\mathbf{0}_n, \sigma_u^2 \mathbf{I}_n)$.

    (b) $\hat{\boldsymbol{\beta}}|\mathbf{X} \sim \mathcal{N}(\boldsymbol{\beta}, \sigma_u^2(\mathbf{X}'\mathbf{X})^{-1})$.

    (c) $\frac{n-k-1}{\sigma_u^2} \times \hat{\sigma_u^2} \sim \chi^2_{n-k-1}$, and

    (d) $t_j = \frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}_u \sqrt{[(\mathbf{X}'\mathbf{X})^{-1}]_{jj}}} \sim t(n-k-1)$

20. Why can multicollinearity matter?

21. What is an omitted variable bias?

22. Explain other forms of regressor endogeneity such as the choice of an incorrect functional form, measurement error, and simultaneous causality.

23. Interpret the regression coefficients for different functional forms: linear-linear, log-linear, linear-log, and log-log.

24. What is the use of interaction terms?

Econometric Questions: RESULTS

1. The **cumulative probability distribution function** of a random variable evaluated at a particular value is the probability that the random variable is less than or equal to that particular value.

   The **probability distribution** of discrete random variables is the list of all possible values of the variable and the probability that each value will occur. These probabilities sum to 1.

   Because a continuous random variable can take on a continuum of possible values, the probability distribution used for discrete variables, which lists the probability of each possible value of the random variable, is not suitable for continuous variables. Instead, the probability is summarized by the **probability density function**. The area under the probability density function between any two points is the probability that the random variable falls between those two points.

2. The **expected value** of a random variable Y, denoted $E(Y)$, is the long-run average value of the random variable over many repeated trials or occurrences. The expected value of a discrete random variable is computed as a weighted average of the possible outcomes of that random variable, where the weights are the probabilities of that outcome. The expected value of Y is also called the expectation of Y or the mean of Y and is denoted μ_Y.

   The variance and standard deviation measure the dispersion or the "spread" of a probability distribution. The **variance** of a random variable Y, denoted $var(Y)$, is the expected value of the square of the deviation of Y from its mean: $var(Y) = E[(Y - \mu_Y)^2]$. Because the variance involves the square of Y, the units of the variance are the units of the square of Y, which makes the variance awkward to interpret. It is therefore common to measure the spread by the **standard deviation**, which is the square root of the variance and is denoted σ_Y. The standard deviation has the same unit as Y.

3. The **joint probability distribution** of two discrete random variables, say X and Y, is the probability that the random variables simultaneously take on certain values, say x and y. The probabilities of all possible (x, y) combinations sum to 1. The joint probability distribution can be written as the function $Pr(X = x, Y = y)$.

   The distribution of a random variable Y conditional on another random variable X taking on a specific value is called the **conditional distribution** of Y given X. The conditional probability that Y takes on the value y when X takes on the value x is written $Pr(Y = y|X = x)$.

4. A continuous random variable with a **normal distribution** has the familiar bell-shaped probability density. The function defining the normal probability density is given by:

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{y-\mu}{\sigma})^2}$$

The normal density with mean μ and variance σ^2 is symmetric around its mean and has 95% of its probability between $\mu - 1.96\sigma$ and $\mu + 1.96\sigma$. The Normal distribution with mean μ and variance σ^2 is expressed concisely as $N(\mu, \sigma^2)$. The standard normal distribution is the normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$ and is denoted by $N(0, 1)$.

The **Student t distribution** with m degrees of freedom is defined to be the distribution of the ratio of a standard normal random variable, divided by the square root of an independently distributed chi-squared random variable with m degrees of freedom divided by m.

The **F distribution** with m and n degrees of freedom, denoted $F_{m,n}$ is defined to be the distribution of the ratio of a chi-squared random variable with degrees of freedom m divided by m, to an independently distributed chi-squared random variable with degrees of freedom n, divided by n.

The **chi-squared distribution** is the distribution of the sum of m squared independent standard normal random variables. The distribution depends on m, which is called the degrees of freedom of the chi-squared distribution. A chi-squared distribution with m degrees of freedom is denoted by χ^2_m.

5. Let $S_1, S_2, ..., S_n$ be a sequence of random variables. For example, S_n could be the sample average $\overline{Y}$ of a sample of n observations of the random variable Y. The sequence of random variables $\{S_n\}$ is said to **converge in probability** to a limit, μ (that is, $S_n \xrightarrow{p} \mu$), if the probability that S_n is within $\pm\delta$ of μ tends to 1 as $n \to \infty$, as long as the constant δ is positive. That is, $S_n \xrightarrow{p} \mu$ if and only if $Pr(|S_n - \mu| \geq \delta) \to 0$ as $n \to \infty$ for every $\delta > 0$.

   If the distributions of a sequence of random variables converge to a limit as n goes to infinity, then the sequence of random variables is said to **converge in distribution**. The central limit theorem says that, under general conditions, the standardized sample average converges in distribution to a normal random variable. Let $F_1, F_2, ..., F_n$ be a sequence cumulative distribution functions corresponding to a sequence of random variables, $S_1, S_2, ..., S_n$. For example, S_n might be the standardized sample average, $(\overline{Y} - \mu_{\overline{Y}})/\sigma_{\overline{Y}}$. Then the sequence of random variables S_n is said to converge in distribution to S (denoted $S_n \xrightarrow{d} S$) if the distribution functions $\{F_n\}$ converge to F, the distribution of S. That is, $S_n \xrightarrow{d} S$ if and only if $\lim_{n\to\infty} F_n(t) = F(t)$ where the limit holds at all points t at which the limiting distribution F is continous.

6. **Random sampling** is one of the most popular types of random or probability sampling. In a simple random sample, n objects are selected at random from a population and each member of the population is equally likely to be included in the sample. The value of the random variable Y for the i^{th} randomly drawn object is denoted Y_i. Because each object is equally likely to be drawn and the distribution of Y_i is the same for all i, the random variables $Y_1, ..., Y_n$ are independently and identically distributed (i.i.d.); that is, the distribution of Y_i is the same for all $i = 1, ..., n$ and Y_1 is distributed independently of $Y_2, ..., Y_n$ and so forth.

7. The **sample average** or sample mean, $\overline{Y}$, of the n observations $Y_1, ..., Y_n$ is obtained as follows:

$$\overline{Y} = \frac{1}{n}(Y_1 + Y_2 + ... + Y_n) = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

8. (a) Let $\hat{\mu}_Y$ denote some estimator of μ_Y, such as $\bar{Y}$ or Y_1. The estimator $\hat{\mu}_Y$ is **unbiased** if $E(\hat{\mu}_Y) = \mu_Y$, where $E(\hat{\mu}_Y)$ is the mean of the sampling distribution of $\hat{\mu}_Y$; otherwise, $\hat{\mu}_Y$ is biased.

   Another desirable property of an estimator μ_Y is that, when the sample size is large, the uncertainty about the value of μ_Y arising from random variations in the sample is very small. Stated more precisely, a desirable property of $\hat{\mu}_Y$ is that the probability that it lies within a small interval of the true value μ_Y approaches 1 as sample size increases, that is, $\hat{\mu}_Y$ is **consistent** for μ_Y.

   (b) The **mean of** $\bar{Y}$ is given by:

$$E(\overline{Y}) = \frac{1}{n}\sum_{i=1}^{n} E(Y_i) = \mu_Y.$$

   The **variance of** $\bar{Y}$ is given as follows:

$$\begin{aligned}
\sigma^2_{\bar{Y}} &= var\left(\frac{1}{n}\sum_{i=1}^{N} E(Y_i)\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{N} var(Y_i) + \frac{1}{n^2}\sum_{i=1}^{N}\sum_{j=1, j\neq i}^{N} cov(Y_i, Y_j) \\
&= \frac{\sigma^2_Y}{n}
\end{aligned}$$

   where $Y_1, ..., Y_n$ are i.i.d. and Y_i and Y_j are independently distributed for $i \neq j$, so $cov(Y_i, Y_j) = 0$.

(c) The law of large numbers states that, under general conditions, $\bar{Y}$ will be near μ_Y with very high probability when n is large. This is sometimes called the "law of averages." When a large number of random variables with the same mean are averaged together, the large values balance the small values and their sample average is close to their common mean. The sample average $\overline{Y}$ converges in probability to μ_Y (or, equivalently, $\overline{Y}$ is consistent for μ_Y) if the probability that $\overline{Y}$ is in the range $(\mu_Y - c)$ to $(\mu_Y + c)$ becomes arbitrarily close to 1 as n increases for any constant $c > 0$. The convergence of $\overline{Y}$ to μ in probability is written, $\overline{Y} \xrightarrow{p} \mu_Y$. The **law of large numbers** says that if Y_i, $i = 1, .., n$ are are independently and identically distributed with $E(Y_i) = \mu_Y$ and if large outliers are unlikely (technically if $var(Y_i) = \sigma_Y^2 < \infty$, then $\overline{Y} \xrightarrow{p} \mu_Y$.

(d) Recall that the mean of $\bar{Y}$ is μ_Y and its variance is $\sigma_{\bar{Y}}^2 = \sigma_Y^2/n$. According to the **central limit theorem**, when n is large, the distribution of $\bar{Y}$ is approximately $N(\mu_Y, \sigma_{\bar{Y}}^2)$. The distribution of $\bar{Y}$ is *exactly* $N(\mu_Y, \sigma_{\bar{Y}}^2)$ when the sample is drawn from a population with the normal distribution $N(\mu_Y, \sigma_Y^2)$. Therefore, the distribution of the standardized version of $\bar{Y}$, $(\bar{Y} - \mu_Y)/\sigma_{\bar{Y}}$ is well approximated by a standard normal distribution $N(0, 1)$ when n is large.

9. The **population regression line** is given by $Y_i = \beta_0 + \beta_1 X_i + u_i$. This is the relationship that holds between Y and X on average over the population. Thus, if you knew the value of X, according to this population regression line you would predict that the value of the dependent variable, Y, is $\beta_0 + \beta_1 X$. Since the coefficients β_0 and β_1 of the population regression line are unknown we must use sample data to estimate them. This can be achieved using the ordinary least squares (OLS) estimators for β_0 denoted by $\hat{\beta}_0$ and the OLS estimator for β_1 denoted by $\hat{\beta}_1$. The OLS regression line, also called **sample regression line**, is the straight line constructed using the OLS estimators: $\hat{\beta}_0 + \hat{\beta}_1 X$. The predicted value of Y_i given X_i, based on the OLS regression line is $Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$. The residual for the i^{th} observation is the difference between Y_i and its predicted value: $\hat{u}_i = Y_i - \bar{Y}_i$. Concluding, one can say that the OLS estimators, $\hat{\beta}_0$ and $\hat{\beta}_1$ are sample counterparts of the population coefficients, β_0 and β_1. Similarly, the OLS regression line $\hat{\beta}_0 + \hat{\beta}_1 X$ is the sample counterpart of the population regression line $\beta_0 + \beta_1 X$, and the OLS residuals $\hat{u}_i$ are sample counterparts of the population errors u_i.

10. The formula $\boldsymbol{\hat{\beta} = (X'X)^{-1}X'y}$ **calculates the OLS estimators** $\hat{\beta}_0$ and $\hat{\beta}_1$ for our true unknown population coefficients β_0 and β_1, where $(X'X)$ is non-singular matrix. The OLS estimator minimizes the sum of squared prediction mistakes, $\sum_{i=1}^{n}(Y_i - b_0 - b_1 X_{1i} - \cdots - b_k X_{ki})^2$. The above formula for the OLS estimator is obtained by taking the derivative of the sum of squared prediction mistakes with respect to each element of the coefficient vector, setting these derivatives to zero, and solving for the estimator $\hat{\beta}$.

11. The algebra of OLS in the multivariate model relies on the two symmetric n x n matrices, $\boldsymbol{P_X = X(X'X)^{-1}X'}$ and $\boldsymbol{M_X = I_n - P_X}$. A matrix C is idempotent if C is square and $CC = C$. Because $P_X = P_X P_X$ and $M_X = M_X M_X$ and because P_X and M_X are symmetric, P_X and M_X are symmetric idempotent matrices. The matrices P_X and M_X can be used to decompose an n-dimensional vector Z into two parts: a part that is spanned by the columns of X and a part orthogonal to the columns of X. In other words, $P_X Z$ is the projection of Z onto the space spanned by the columns of X, $M_X Z$ is the part of Z orthogonal to the columns of X, and $Z = P_X Z + M_X Z$.

12. The least squares estimator is **unbiased** in every sample. To show this, write

$$b = (X'X)^{-1}X'y = (X'X)^{-1}X'(X\beta + \epsilon) = \beta + (X'X)^{-1}X'\epsilon$$

Now, take expectations, iterating X;

$$E[b|X] = \beta + E[(X'X)^{-1}X'\epsilon|X]$$

By Assumption (1), the second term is 0, so

$$E[b|X] = \beta$$

Therefore,

$$E[b] = E_X \{[b|X]\} = E_X\beta = \beta$$

The interpretation of this result is that for any particular set of observations, X, the least squares estimator has expectation β. Therefore, when we average this over the possible values of X we find the unconditional mean is β as well.

To show **consistency** we write least squares estimator as follows:

$$b = \beta + \left(\frac{X'X}{n}\right)^{-1}\left(\frac{X'\epsilon}{n}\right).$$

If Q^{-1} exists, then

$$\operatorname{plim} b = \beta + Q^{-1}\operatorname{plim}\left(\frac{X'\epsilon}{n}\right)$$

because the inverse is a continuous function of the original matrix. We require the probability limit of the last term. Let

$$\frac{1}{n}X'\epsilon = \frac{1}{n}\sum_{i=1}^{n} x_i\epsilon_i = \frac{1}{n}\sum_{i=1}^{n} w_i = \bar{w}$$

Then,

$$\operatorname{plim} b = \beta + Q^{-1}\operatorname{plim}\bar{w}$$

We have $\text{Var}[\bar{w}] = E[\text{Var}[\bar{w}|X]] + \text{Var}[E[\bar{w}|X]]$. The second term is zero because $E[\epsilon_i|x_i]$=0. To obtain the first, we use $E[\epsilon_i'\epsilon_i|X] = \sigma^2 I$, so We have $\text{Var}[\bar{w}] = E[\text{Var}[\bar{w}|X]] + \text{Var}[E[\bar{w}|X]]$. The second term is zero because $E[\epsilon_i|x_i]$=0. To obtain the first, we use $E[\epsilon_i'\epsilon_i|X] = \sigma^2 I$, so

$$\text{Var}[\bar{w}|X] = E[\bar{w}'\bar{w}|X] = \frac{1}{n}X'E[\epsilon_i'\epsilon_i|X]X\frac{1}{n} = \left(\frac{\sigma^2}{n}\right)\left(\frac{X'X}{n}\right)$$

Therefore,

$$\text{Var}[\bar{w}] = \left(\frac{\sigma^2}{n}\right) E\left(\frac{X'X}{n}\right)$$

The variance will collapse to zero if the expectation in parentheses is (or converges to) a constant matrix, so that the leading scalar will dominate the product as n increases. It then follows that

$$\lim_{n\to\infty} \text{Var}[\bar{w}] = 0 \cdot\ Q = 0$$

Since the mean of $\bar{w}$ is identically zero and its variance converges to zero, $\bar{w}$ converges in mean square to zero, so plim $\bar{w} = 0$. Therefore

$$\text{plim}\,\frac{X'\epsilon}{n} = 0,$$

so,

$$\text{plim}\, b = \beta + Q^{-1} \cdot 0 = \beta$$

This result establishes that under four Assumptions, b is a consistent estimator of β in the linear regression model.

Finally, we need to show that the OLS estimator of the population model is **asymptotically normally distributed**. We start with:

$$\sqrt{n}(b - \beta) = \left(\frac{X'X}{n}\right)\left(\frac{1}{\sqrt{n}}\right) X'\epsilon$$

Since the inverse matrix is a continuous function of the original matrix, $\text{plim}(X'X/n)^{-1} = Q^{-1}$. Therefore, if the limiting distribution of the random vector exists, then that limiting distribution is the same as that of

$$\left[\text{plim}\left(\frac{X'X}{n}\right)^{-1}\right]\left(\frac{1}{\sqrt{n}}\right) X'\epsilon = Q^{-1}\left(\frac{1}{\sqrt{n}}\right) X'\epsilon$$

Thus, we must establish the limiting distribution of

$$\left(\frac{1}{\sqrt{n}}\right) X'\epsilon = \sqrt{n}(\bar{w} - E[\bar{w}])$$

where $E[\bar{w}] = 0$. We can use the multivariate Lindeberg-Feller version of the central limit theorem to obtain the limiting distribution of $\sqrt{n}\bar{w}$. Using that formulation, $\bar{w}$ is the average of n independent random vectors $w_i = x_i\epsilon_i$, with means 0 and variances

$$\text{Var}\left[x_i'\epsilon_i\right] = \sigma^2 E\left[x_i'x_i\right] = \sigma^2 Q_i$$

The variance of $\sqrt{n}\bar{w}$ is

$$\sigma^2 \bar{Q}_n = \sigma^2 \left(\frac{1}{n}\right)\left[Q_1 + Q_2 + \cdots + Q_n\right]$$

As long as the sum is not dominated by any particular term and the regressors are well behaved, which in this case, the following holds:

$$\lim_{n\to\infty} \sigma^2 \bar{Q}_n = \sigma^2 \bar{Q}$$

Therefore, we may apply the Lindeberg-Feller central limit theorem to the vector $\sqrt{n}\bar{w}$. We now have the elements we need for a formal result. If $[x_i\epsilon_i]$, $i = 1, \ldots\ldots, n$ are independent vectors distributed with mean 0 and variance $\sigma^2 Q_i < \infty$ then

$$\left(\frac{1}{\sqrt{n}}\right) X'\epsilon \xrightarrow{d} N[0, \sigma^2 Q].$$

It then follows that

$$Q^{-1}\left(\frac{1}{\sqrt{n}}\right) X'\epsilon \xrightarrow{d} N[Q^{-1}0, Q^{-1}(\sigma^2 Q)Q^{-1}].$$

Combining terms,

$$\sqrt{n}(b - \beta) \xrightarrow{d} N[0, \sigma^2 Q].$$

the asymptotic standard normal distribution of b.

13. The **standard error of the regression (SER)** estimates the standard deviation of the error term u_i. Thus the SER is a measure of the spread of the distribution of Y around the regression line. In multiple regression, the SER is

$$SER = s_{\hat{u}} = \sqrt{s_{\hat{u}}^2} \text{ where } s_{\hat{u}}^2 = \frac{1}{n-k-1}\sum_{i=1}^{n} \hat{u}_i^2 = \frac{SSR}{n-k-1}$$

where SSR is the sum of squared residuals, $SSR = \sum_{i=1}^{n} \hat{u}_i^2$

The regression $\boldsymbol{R^2}$ is the fraction of the sample variance Y_i explained by (or predicted by) the regressors. Equivalently, the R^2 is 1 minus the fraction of the variance of Y_i *not* explained by the regressors. The mathematical definition of the R^2 is the same as for the regression with a single regressor:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{SSR}{TSS'}$$

where the explained sum of squares is $ESS = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$ and the total sum of squares is $TSS = \sum_{i=1}^{n}(Y_i - \bar{Y})^2$

Because the R^2 increases when a new variable is added to the model, an increase in the R^2 when adding an new variable does not necessarily mean that the added variable actually improves the fit of the model. In this sense, the R^2 gives an inflated estimate of how well the regression fits the data. One way to correct for this is to deflate or reduce the R^2 by some factor, and this is what the **adjusted $\boldsymbol{R^2}$** , or $\boldsymbol{\bar{R}^2}$, does. The adjusted R^2, or $\bar{R}^2$, is a modified version of the R^2 that does not necessarily increase when a new regressor is added. The $\bar{R}^2$ is calculated as follows:

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1}\frac{SSR}{TSS'} = 1 - \frac{s_{\hat{u}}^2}{s_Y^2}$$

The difference between this formula and the one of R^2 is that the ratio of the sum of squared residuals to the total sum of squares is multiplied by the factor $(n-1)/(n-k-1)$. As the second expression in the formula of $\bar{R}^2$ shows, this means that the adjusted R^2 is 1 minus the ratio of the sample variance of the OLS residuals (with the degree-of-freedom correction in SER formula) to the sample variance of Y.

14. Recall that $P_X = X(X'X)^{-1}X'$, where $P_X = P_X P_X$ and $M_X = I_n - P_X$, where $M_X = M_X M_X$. Further recall the OLS predicted values $\hat{Y} = X\hat{\beta}$ and the OLS residuals, $\hat{U} = Y - \hat{Y}$. We now can rewrite $\hat{Y}$ and $\hat{U}$ as follows:

$$\hat{Y} = X\hat{\beta} = X(X'X)^{-1}X'Y = P_X Y$$

$$\hat{U} = Y - \hat{Y} = Y - X\hat{\beta} = Y - X(X'X)^{-1}X'Y = (I_n - X(X'X)^{-1}X')Y = M_X Y =$$

$$M_X(X\beta + U) = M_X X\beta + M_X U = (I_n - X(X'X)^{-1}X')X + M_X U = M_X U$$

Hence, **the squared SER**, $s_{\hat{u}}^2$, can be written as:

$$s_{\hat{u}}^2 = \frac{1}{n-k-1}\sum_{i=1}^{n}\hat{u}_i^2 = \frac{1}{n-k-1}\hat{U}'\hat{U} = \frac{1}{n-k-1}U'M_X U$$

where the final equality follows because $\hat{U}'\hat{U} = (M_X U)'(M_X U) = U'M_X M_X U = U'M_X U$ (because M_X is symmetric and idempotent).

15. The **t-Statistic** is a statistic used to perform a hypothesis test. In general the t-statistic has the form

$$t = \frac{\text{estimator-hypothesized value}}{\text{standard error of the estimator}}.$$

The null and alternative hypothesis need to be stated precisely before they can be tested. As an example, assume we want to test hypotheses about the slope β_1. The null hypothesis and the two-sided alternative hypothesis are

$$H_0 : \beta_1 = \beta_{1,0} \quad vs. \quad H_1 : \beta_1 \neq \beta_{1,0} \quad (two - sided - alternative)$$

One could also use a one-sided alternative hypothesis here and test against $H_1 : \beta_1 > or < \beta_{1,0}$. The first step is to compute the standard error of $\hat{\beta}_1$, $SE(\hat{\beta}_1)$. The standard error of $\hat{\beta}_1$ is an estimator of $\sigma_{\hat{\beta}_1}$ the standard deviation of the sampling distribution of $\hat{\beta}_1$. Specifically,

$$SE(\hat{\beta}_1) = \sqrt{\hat{\sigma}^2_{\hat{\beta}_1}}$$

where

$$\hat{\sigma}^2_{\hat{\beta}_1} = \frac{1}{n} \times \frac{\frac{1}{n-2}\sum_{i=1}^{n}(X_i - \bar{X})^2 \hat{u}_i^2}{[\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2]^2}$$

The second step is to compute the t-statistic,

$$t = \frac{\hat{\beta}_1 - \beta_{1,0}}{SE(\hat{\beta}_1)}$$

The third step is to compute the **p-value**, the probability of observing a value of $\hat{\beta}_1$ at least as different from $\beta_{1,0}$ as the estimate actually computed ($\hat{\beta}_1^{act}$), assuming that the null hypothesis is correct. Stated mathematically,

$$\begin{aligned} p\text{-value} &= Pr_{H_0}\left[\left|\hat{\beta}_1 - \beta_{1,0}\right| > \left|\hat{\beta}_1^{act} - \beta_{1,0}\right|\right] \\ &= Pr_{H_0}\left[\left|\frac{\hat{\beta}_1 - \beta_{1,0}}{SE(\hat{\beta}_1)}\right| > \left|\frac{\hat{\beta}_1^{act} - \beta_{1,0}}{SE(\hat{\beta}_1)}\right|\right] \\ &= Pr_{H_0}(|t| > |t^{act}|) \end{aligned}$$

where Pr_{H_0} denotes the probability computed under the null hypothesis. Because $\hat{\beta}_1$ is approximately normally distributed in large samples, under the null hypothesis the t-statistic is approximately distributed as a standard normal random variable, so in large samples,

$$p\text{-value} = Pr(|Z| > |t^{act}|) = 2\Phi(-|t^{act}|)$$

A p-value of less than 5% provides evidence against the null hypothesis in the sense that, under the null hypothesis, the probability of obtaining a value of $\hat{\beta}_1$ at least as far from the null as that actually observed is less than 5%. If so, the null hypothesis is rejected at the 5% significance level. Alternatively, the hypothesis can be tested at the 5% significance level simply by comparing the absolute value of the t-statistic to 1.96, the critical value for a two-sided test, and rejecting the null hypothesis at the 5% level if $|t^{act}| > 1.96$.

The **F-Statistic** is used to test joint hypotheses. When the joint null hypothesis has the two restrictions that $\beta_1 = 0$ and $\beta_2 = 0$, the F-statistic combines the two t-statistics t_1 and t_2 using the formula

$$F = \frac{1}{2}\left(\frac{t_1^2 + t_2^2 - 2\hat{\rho}_{t_1,t_2} t_1 t_2}{1 - \hat{\rho}^2_{t_1,t_2}}\right)$$

where $\hat{\rho}_{t_1,t_2}$ is an estimator of the correlation between the two t-statistics. If we want to test for more restrictions, say q, we need to use the following approach. Consider a joint hypothesis that is linear in the coefficients and imposes q restrictions, where $q \geq k+1$. Each of these q restrictions can involve one or more of the regression coefficients. This joint null hypothesis can be written in matrix notation as

$$R\beta = r$$

where R is a $q \times (k+1)$ nonrandom matrix with full row rank and r is nonrandom $q \times 1$ vector. The number of rows of R is q, which is the number of restrictions being imposed under the null hypothesis. The heteroskedasticity-robust F-statistic testing the joint hypothesis in matrix form is

$$F = (R\hat{\beta} - r)'[R\hat{\Sigma}_{\hat{\beta}}R']^{-1}(R\hat{\beta} - r)/q$$

16. A 95% two-sided **confidence interval** for β_1 is an interval that contains the true value of β_1 with a 95% probability; that is, it contains the true value of β_1 in 95% of all possible randomly drawn samples. Equivalently, it is the set of values of β_1 that cannot be rejected by a 5% two-sided hypothesis test. When the sample size is large, it is constructed as

$$95\% \text{ confidence interval for } \beta_1 = [\hat{\beta}_1 - 1.96SE(\hat{\beta}_1), \hat{\beta}_1 + 1.96SE(\hat{\beta}_1)].$$

17. The error term u_i is homoskedastic if the variance of the conditional distribution of u_i given X_i, $var(u_i|X_i = x)$, is constant for $i = 1, ..., n$ and in particular does not depend on x. Otherwise the error term is heteroskedastic. The **homoskedasticity-only standard error** of $\hat{\beta}_1$ is $SE(\hat{\beta}_1) = \sqrt{\tilde{\sigma}^2_{\hat{\beta}_1}}$, where $\tilde{\sigma}^2_{\hat{\beta}_1}$ is the homoskedasticity-only estimator of the

variance of $\hat{\beta}_1$:

$$\tilde{\sigma}^2_{\hat{\beta}_1} = \frac{s^2_{\hat{u}}}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

The **heteroskedasticity-robust standard errors** can be obtained via

$$\hat{\sigma}^2_{\hat{\beta}_1} = \frac{1}{n} \times \frac{\frac{1}{n-2}\sum_{i=1}^{n}(X_i - \bar{X})^2\hat{u}_i^2}{[\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2]^2}$$

Because homoskedasticity is a special case of heteroskedasticity, the estimator $\hat{\sigma}^2_{\hat{\beta}_1}$ of the variances of $\hat{\beta}_1$ given above produces valid statistical inferences whether the errors are heteroskedastic or homoskedastic. Thus hypothesis tests and confidence intervals based on those standard errors are valid whether or not the errors are heteroskedastic. This is not the case for $\tilde{\sigma}^2_{\hat{\beta}_1}$ which can only be used under homoskedasticity.

18. If the three least squares assumptions (1. The error term u_i has conditional mean zero given X_i: $E(u_i|X_i) = 0$, 2. (X_i, Y_i), $i = 1, ..., n$, are i.i.d. draws from their joint distribution, 3. Large outliers are unlikely: X_i and Y_i have nonzero finite fourth moments) hold *and* if errors are homoskedastic, then the OLS estimator $\hat{\beta}_1$ is the **B**est (most efficient) **L**inear conditionally **U**nbiased **E**stimator **(BLUE)**.

19. For the following proofs we make use of the following six assumptions:
(1) $E(u_i|X_i) = 0$ (u_i has conditional mean zero)
(2) (X_i, Y_i), $i = 1, ..., n$, are independently and identically distributed (i.i.d.) draws from their joint distribution
(3) X_i and u_i have nonzero finite fourth moments.
(4) X has full column rank (there is no perfect multicollinearity)
(5) $var(u_i|X_i) = \sigma_u^2$ (homoskedasticity)
(6) The conditional distribution of u_i given X_i is normal (normal errors)

    (a) The first and second assumptions imply that $E(u_i|X) = E(u_i|X_i) = 0$ and that $cov(u_i, u_j|X) = E(u_i u_j|X) = E(u_i u_j|X_i, X_j) = E(u_i|X_i)E(u_j|X_j) = 0$ for $i \neq j$. The first, second, and fifth assumptions imply that $E(u_i^2|X) = E(u_i^2|X_i) = \sigma_u^2$. Combining these results, we have that under Assumptions (1) and (2), $E(U|X) = 0_n$, and under Assumptions (1), (2) and (5), $E(UU'|X) = \sigma_u^2 I_n$, where 0_n is the n-dimensional vector of zeros and I_n is the $n \times n$ identity matrix. Similarly, the first, second, fifth, and sixth assumptions imply that the conditional distribution of the n-dimensional random vector U, conditional on X is the multivariate normal distribution. That is, under Assumptions (1), (2), (5) and (6), **the conditional distribution of U given X is $N(0_n, \sigma_u^2 I_n)$.**

(b) Because $\hat{\beta} = \beta + (X'X)^{-1}X'U$ and because the distribution of U conditional on X is, by assumption, $N(0_n, \sigma_u^2 I_n)$, the conditional distribution of $\hat{\beta}$ given X is multivariate normal with mean β. The covariance matrix of $\hat{\beta}$, conditional on X, is $\sigma_{\hat{\beta}|X} = E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'|X] = E[(X'X)^{-1}X'UU'X(X'X)^{-1}|X] = (X'X)^{-1}X'(\sigma_u^2 I_n)X(X'X)^{-1}$ $= \sigma_u^2(X'X)^{-1}$. Accordingly, under all six assumptions, **the finite-sample conditional distribution of $\hat{\beta}$ given X is $\hat{\beta} \sim N(\beta, \sigma_u^2(X'X)^{-1})$.**

(c) If all six assumptions hold, then $s_{\hat{u}}^2 = \hat{\sigma}_u^2$ has an exact sampling distribution that is proportional to a chi-squared distribution with $n-k-1$ degrees of freedom:

$$s_{\hat{u}}^2 \sim \frac{\sigma_u^2}{n-k-1} \times \chi_{n-k-1}^2$$

The proof starts with the following equation:

$$s_{\hat{u}}^2 = \frac{1}{n-k-1}\sum_{i=1}^{n}\hat{u}_i^2 = \frac{1}{n-k-1}\hat{U}'\hat{U} = \frac{1}{n-k-1}\hat{U}'M_xU.$$

Because U is normally distributed conditional on X and because M_x is a symmetric idempotent matrix, the quadratic form $U'M_xU/\sigma_u^2$ has an **exact chi-squared distribution with degrees of freedom equal to the rank of M_x**. The rank of M_x is $n-k-1$. Thus $U'M_xU/\sigma_u^2$ has an exact χ_{n-k-1}^2 distribution, from which the given equation follows.

(d) If (i) Z has a standard normal distribution, (ii) W has a chi-squared distribution with m degrees of freedom distribution, and (iii) Z and W are independently distributed, then the random variable $Z/\sqrt{(W/m)}$ has a t-distribution with m degrees of freedom. We start the proof with the following Equation where $s_{\hat{u}} = \hat{\sigma}_u$:

$$t_j = \frac{\hat{\beta}_j - \beta_{j,0}}{s_{\hat{u}}\sqrt{[(X'X)^{-1}]_{jj}}}$$

To put t_j in the above mentioned form, we rewrite the equation as

$$t_j = \frac{(\hat{\beta}_j - \beta_{j,0})/\sqrt{\sigma_u^2[(X'X)^{-1}]_{jj}}}{\sqrt{W/(n-k-1)}}$$

where $W = (n-k-1)/(s_{\hat{u}}^2/\sigma_u^2)$, and let $Z = (\hat{\beta}_j - \beta_{j,0})/\sqrt{\sigma_u^2[(X'X)^{-1}]_{jj}}$ and $m = n-k-1$. With these definitions, $t_j = Z/\sqrt{W/m}$. Thus, to prove the statement given in the question, we must show (i) through (iii) for these definitions of Z, W and m.

(i) An implication of the results of exercise (b) is that, under the null hypothesis, $Z = (\hat{\beta}_j - \beta_{j,0})/\sqrt{\sigma_u^2[(X'X)^{-1}]_{jj}}$ has an exact standard normal distribution, which shows (i).

(ii) From exercise (c) we know that W has a chi-squared distribution with $n-k-1$ degrees of freedom which shows (ii).
(iii) To show (iii), it must be shown that $\hat{\beta}_j$ and $s_{\hat{u}}^2$ are independently distributed. We already know that $\hat{\beta}-\beta=(X'X)^{-1}X'U$ and $s_{\hat{u}}^2=(M_xU)'(MxU)/(n-k-1)$. Thus $\hat{\beta}-\beta$ and $s_{\hat{u}}^2$ are independent if $(X'X)^{-1}X'U$ and M_xU are independent. Both $(X'X)^{-1}X'U$ and M_xU are linear combinations of U, which has a $N(0_{n\times 1},\sigma_u^2 I_n)$ distribution conditional on X. But because $M_X X(X'X)^{-1}=0_{n\times(k+1)}$, it follows that $(X'X)^{-1}X'U$ and M_xU are independently distributed. Consequently under all six assumptions above $\hat{\beta}_j$ and $s_{\hat{u}}^2$ are independently distributed, which shows (iii) and thus completes the proof.

20. Perfect **multicollinearity** arises when one of the regressors is a perfect linear combination of the other regressors. In case of perfect multicollinearity it is impossible to compute the OLS estimator. Imperfect multicollinearity arises when one of the regressors is very highly correlated - but not perfectly correlated - with the other regressors. Unlike perfect multicollinearity, imperfect multicollinearity does not prevent estimation of the regression, nor does it imply a logical problem with the choice of the regressors. However, it does mean that one or more regression coefficients could be estimated imprecisely.

21. **Omitted variable bias** is the bias in the OLS estimator that arises when the regressor, X, is correlated with an omitted variable. For omitted variable bias to occur, two conditions must be met:
(1) X is correlated with the omitted variable.
(2) The omitted variable is a determinant of the dependent variable, Y.

22. **Functional form misspecification** arises when the functional form of the estimated regression function differs from the functional form of the population regression function. If the functional form is misspecified, then the estimator of the partial effect of a change in one of the variables will, in general, be biased. Fuctional form misspecification often can be detected by plotting the data and the estimated regression function, and it can be corrected by using a different functional form.

    Errors-in-variables in the OLS estimator arises when an independent variable is measured imprecisely. This bias depends on the nature of the measurement error and persists even if the sample size is large. If the measured variable equals the actual value plus a mean-zero, independently distributed measurement error, then the OLS estimator in a regression with a single right-hand variable is biased toward zero, and its probability limit is given by:

$$\hat{\beta}_1 \xrightarrow{p} \frac{\sigma_X^2}{\sigma_X^2+\sigma_w^2}\beta_1$$

    The **measurement error** in Y is different from measurement error in X (Error-in-variable). If Y has a classical measurement error, then this measurement error increases the variance of the regression and of $\hat{\beta}_1$S but does not induce bias in $\hat{\beta}_1$.

Simultaneous causality bias, also called simultaneous equations bias, arises in a regression of Y on X when, in addition to the causal link of interest from X to Y, there is a causal link from Y to X. This reverse causality makes X correlated with the error term in the population regression of interest.

23. Logarithms can be used to transform the dependent variable Y, an independent variable X, or both (but the variable being transformed must be positive). The following table summarizes these four cases and the interpretation of the regression coefficient β_1. In each case, β_1 can be estimated by applying OLS after taking the logarithm of the dependent and/or independent variable.

Case	Regression Specification	Intepretation of β_1
linear-linear	$Y_i = \beta_0 + \beta_1 X_i + u_i$	A change in X by one unit ($\Delta X = 1$) is associated with a change in Y of β_1.
log-linear	$ln(Y_i) = \beta_0 + \beta_1 X_i + u_i$	A change in X by one unit ($\Delta X = 1$) is associated with a $100\beta_1\%$ change in Y.
linear-log	$Y_i = \beta_0 + \beta_1 ln(X_i) + u_i$	A 1% change in X is associated with a change in Y of $0.01\beta_1$.
log-log	$ln(Y_i) = \beta_0 + \beta_1 ln(X_i) + u_i$	A 1% change in X is associated with a $\beta_1\%$ change in Y, so β_1 is the elasticity of Y with respect to X.

24. Consider the population regression of Y_i against two binary (only taking on values 0 or 1) variables D_{1i} and D_{2i}. The population linear regression of Y_i on these two variables is

$$Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + u_i$$

The specification of the model has an important limitation: The effect of D_{1i} in this specification, holding constant D_{2i}, is the same for both possible values of D_{2i}. There is, however, no reason that this must be so. Phrased mathematically, the effect on Y_i of D_{1i}, holding D_{2i} constant, could depends on the value of D_{2i}. Although this specification does not allow for this interaction between D_{1i} and D_{2i}, it is easy to modify the specification so that it does by introducing another regressor, the product of the two binary variables

$D_{1i} \times D_{2i}$. The resulting regression is

$$Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_3 (D_{1i} \times D_{2i}) + u_i$$

The new regressor, the product $D_{1i} \times D_{2i}$, is called **interaction term** or interacted regressor. The interaction term allows the population effect on Y_i of Changing D_{1i} (from $D_{1i} = 0$ to $D_{1i} = 1$) to depend on D_{2i}.

Statistical Questions

1. Assume that the random variable X follows an unknown distribution with $E(X) = 2$ and $Var(X) = 4$. A random variable Y also follows an unknown distribution with $E(Y) = -2$, and $Var(Y) = 9$. Additionally, $\rho_{XY} = 0.5$. Now consider the random variable Z with:

$$Z = \omega X + (1 - \omega)Y, \omega \in (0, 1]$$

   (a) Calculate the expected value and the variance of Z.

   (b) Determine the value of ω such that the variance of Z is minimized.

   (c) Now assume that $\rho_{XY} = 0$. What consequences for the dependence between Y and X do you suspect? What follows if X and Y follow a Normal distribution?

2. Assume that the continuous random variable has an expected value of $E(X) = 2$ and variance $Var(x) = 1$. What conclusions do you draw about the probability $P(0 < X < 4)$?

3. Consider the random variable Y having probability density function:

$$f_Y(y) = \tfrac{3}{2}(y - 0.5y^2), 0 < y < 2$$

   (a) Calculate the expected value and the variance of Y.

   (b) Now consider the second random variable Z having the same probability density function as Y. Correlation between Y and X is given by $3Var(Y)$. Calculate the variance of the random variable T given by $T = 2Y + Z$

4. Consider an urn containing $N = 10$ balls, of which $M = 3$ are red. From this urn you draw $n = 3$ times one ball without replacement. Calculate the probability of at least drawing 2 red balls.

5. Consider an identically and independently distributed random sample, $(x_1, ..., x_n)$, from an exponentially-distributed population having probability density function $f(x) = \frac{1}{\theta}e^{-\frac{x}{\theta}}$.

   (a) Set up the log likelihood function and calculate the Maximum-Likelihood-Estimator $\widehat{\theta_{ML}}$ for θ.

   (b) Check for biasedness of the ML-estimator and calculate the Mean-Squared-Error (MSE). Is $\widehat{\theta_{ML}}$ MSE-consistent?

6. Consider the following results of a horse race:

Horse	A	B	C	D	$\sum$
Number of wins	10	16	12	22	60

(a) For a significance level of $\alpha = 0.05$ check the hypothesis, whether the number of wins is uniformly distributed across the 4 horses.

(b) Now consider the number of wins of horse D. For a significance level of $\alpha = 0.01$ check the hypothesis of a Bernoulli-distribution with parameter $\rho = 1/3$ ($d = 1$, if horse D won the race, $d = 0$ otherwise)

(c) Calculate the probability of horse D winning at most twice in 10 races under the assumption of a Bernoulli-distribution from part b) with $\rho = 1/3$

7. The random variable X has the following probability density function:

$$f(x) = \begin{cases} 4/\mathrm{x}^5 & \text{if } x > 1 \\ 0 & \text{else} \end{cases}$$

(a) Determine the first order non-central moment and the second order central moment.

(b) Determine the probability density function of the random variable $Z = ln(X^2)$

8. Entrepreneur Meyer produces wooden slats whose lengths are to be of 1 meter. He wants to increase productivity by installing a new machine. In order to control for the quality of the slats from the new machine, $n_1 = 101$ randomly chosen slats are measured. The mean length is 99cm and the sample variance is 30 cm^2. Assume the wooden slats to be normally distributed.

(a) Test the null hypothesis that the mean length coincides with the scheduled value. Calculate the p-value of the conducted test. Assume a significance value of $\alpha = 0.05$

(b) Now assume that the entrepreneur wants to be statistically sure that the mean length of the wooden slats produced by the new machine undercuts 100cm. Formulate the null and the alternative hypotheses, conduct the test and calculate the p-value. *Hint: Assume a large sample size.*

9. A student has no time to prepare for a multiple choice test of 20 questions. He decides to have a guess at all the questions. Each question has 5 possible answers.

(a) How is the random variable distributed, that gives the number of correct answers?

(b) How many questions will the student answer correctly on average?

(c) The test is passed if 10 questions are answered correctly. What is the probability of the student to pass the test?

(d) How should the threshold for passing the test be defined, if the student's chance to pass the test by guessing the answers, should be larger than 5%?

10. After the game ended in a tie the penalty shoot-out takes place. Each team has 5 shoots and the team which scores more often than the other wins. Assume that the single shoots are independent from each other and each player scores with a probability of 0.8. What is the probability that after 10 shots (5 shots per team) the game is decided?

11. Which distributions do the following random variables X possess?

    (a) X_1= Number of items of a rarely sold product demanded in one day

    (b) X_2= Time passing between two worst-case scenarios in nuclear energy generation

    (c) X_3= Number of telephone calls in a large call centre within one hour

    (d) X_4= Number of correct guesses in the lottery "6 out of 49"

12. Assume your little sister does handicrafts and the necklace consists of 50 parts which have on average a length of 2 cm and a standard deviation of 0.2 cm. Which distribution does the length of the whole necklace follow?

13. A bank knows that the mean indebtedness of its student customers at the end of the academic year is distributed approximately normally with a mean of 300 and a standard deviation of 100. What is the probability that a student chosen at random will have a debt:

    (a) Less than 100

    (b) Between 100 and 200

    (c) Over 450

    (d) A debt of 400

14. A dating agency usually claims in its advertising that the mean age of the members on its list is lower than 40. As a check, it takes a sample of 30 members and obtains a sample mean age of 37 years and a sample variance of 25 years. Test whether its claim is justified with:

    (a) $\alpha = 0.05$

    (b) $\alpha = 0.01$

15. A trade union believes in the basis of its membership records that at least four out of every five union members in the financial services sector are members of its union. A smaller rival union denies this statement and wants to show that the proportion is in fact smaller than the one claimed by the larger union. It finds that the proportion belonging to the larger rival union is 0.72. Conduct the associated test using

    (a) $\alpha = 0.05$

    (b) $\alpha = 0.01$

Statistical Questions: RESULTS

1. a) $E(Z) = 4\omega - 2$; $VAR(Z) = 7\omega^2 - 12\omega + 9$
   b) $\omega = 6/7$
   c) $\rho = 0$ implies that there is no linear relationship among x and y. If x and y follow a bivariate normal distribution, $\rho = 0$ implies that x and y are independent.
   $Z \sim N(4\omega - 2, 13\omega^2 - 18\omega + 9)$

2. $P(0 < X < 4) = 0.75$

3. a) $E(Y) = 1$; $VAR(Y) = 0.2$
   b) $VAR(T) = 1.48$

4. $P(X \geq 2) = 11/60$

5. a) $\ln(L) = n\ln(1/\theta) - 1/\theta \sum_i X_i$; $\hat{\theta}_{ML} = \bar{x}$
   b) the estimator is unbiased and MSE-consistent
   c) weglassen

6. a) reject H_0, accept H_1
   b) don't reject H_0
   c) $P(X \leq 2) = 0.299$

7. $E(X) = 1.333$; $E(X^2) = 2 \rightarrow E[(X - \mu)^2] = 0.223$

8. a) don't reject H_0; pvalue=0.0702
   b) reject H_0, accept H_1; pvalue=0.0333

9. a) $X \sim Bin(n = 20, p = 1/5)$
   b) $E(X) = 4$
   c) $P(\text{'pass the test'}) = 0.0026$
   d) the threshold should be 'at least 6 correct answers'.

10. $P(\text{'decision'}) = 0.8855$

11. a) poisson distribution
    b) exponential distribution
    c) normal distribution
    d) hypergeometric distribution.

12. $N(100\,\text{cm}; 2\,\text{cm}^2)$

13. a) $P(X < 100) = 0.0228$
    b) $P(100 < X < 200) = 0.1359$
    c) $P(X > 450) = 0.0668$
    d) $P(X = 450) = 0$

14. a) reject H_0, accept H_1
    b) reject H_0, accept H_1
15. a) reject H_0, accept H_1
    b) don't reject H_0

Macroeconomic Questions

Multiple-Choice Questions

1. The effectiveness of monetary policy depends on the LM-curve and does not depend on the IS-curve.

   (a) True

   (b) False

2. The effectiveness of fiscal policy depends on the IS-curve and does not depend on the LM-curve.

   (a) True

   (b) False

3. An expansionary fiscal policy raises the trade surplus.

   (a) True

   (b) False

4. Higher inflationary expectations decrease the nominal interest rate.

   (a) True

   (b) False

5. On the long run the growth in money supply significantly affects the price levels.

   (a) True

   (b) False

6. A fast adjustment on the money market implies that the economy always moves along

   (a) The IS-curve

   (b) The LM-curve

   (c) The IS- as well as the LM-curve

   (d) Neither the IS- nor the LM curve

7. An increase of government demand raises the

   (a) Interest rates

   (b) Investments

   (c) Interest rates and investment

(d) Neither interest rates nor investment

8. A high interest rate sensitivity of money demand raises

   (a) The multiplier of fiscal policy

   (b) The multiplier of monetary policy

   (c) The multipliers of fiscal as well as monetary policy

   (d) Neither the multiplier of fiscal nor of monetary policy

9. Compared to a closed economy the simple multiplier of an open economy is

   (a) Larger

   (b) Smaller

   (c) The same

   (d) a), b) and c) hold

10. An increase in government spending causes an increase in investment

    (a) In the short run

    (b) In the long run

    (c) In the short as well as In the long run

    (d) Neither in the short nor in the long run

11. Unfavorable supply shocks

    (a) Increase prices and output

    (b) Reduce prices and increase output

    (c) Increase prices and reduce output

    (d) Reduce prices and output

Goods market

Consider the following goods market model

$$\begin{aligned}
Y &= Y^s = Y^d \\
Y^d &= C + I^a + G \\
C &= a + b \cdot Y^v \qquad (a > 0,\ 0 < b < 1) \\
Y^v &= Y - T \\
T &= t \cdot Y \qquad (0 < t < 1)
\end{aligned}$$

where Y = income, Y^s = goods supply, Y^d = goods demand, Y^v = disposable income, C = consumption, I^a = exogenous/autonomous investments, G = government spending, T = taxes, t = tax rate.

1. Derive the equilibrium level of income.
2. Derive the investment and government spending multiplier (dY/dI and dY/dG).
3. Why does the multiplier process converge?
4. Derive the government spending multiplier in case the government spending is completely tax financed ($dG = dT$). Compare your result to the multiplier derived in exercise 2.

IS/LM model

Consider the following IS/LM model for the closed economy (in reduced form):

$$
\begin{aligned}
Y &= C((1-t)Y) + I(i) + G \\
M/\overline{P} &= L(Y, i)
\end{aligned}
$$

where Y = income, C = consumption, I = investment, G = government spending, M = money supply, L = money demand, i = interest rate, $\overline{P}$ = fixed price level.

1. Analyze graphically (in a i/Y diagram) the effects of an increase in
   (a) government spending ($dG > 0$).
   (b) the money supply ($dM > 0$).

   Describe verbally the adjustment process.
2. Derive the government spending multiplier. Why is it smaller than (or equal to) the elementary multiplier with fixed investments? For which special cases is it equal to the elementary multiplier?

IS/LM/Z (Mundell-Fleming) model

Consider the IS/LM/Z model of the small open economy:

$$
\begin{aligned}
Y &= C((1-t)Y) + I(i) + G + A(Y, Y_a, e) \\
\frac{M}{\overline{P}} &= L(Y, i) \\
Z &= 0 = A^n(Y, Y_a, e) + K(i, i_a)
\end{aligned}
$$

where Y = (domestic) income, C = consumption, I = investment, G = government spending, M = money supply, L = money demand, i = interest rate, i_a = foreign interest rate, $\overline{P}$ = fixed price level, A = net exports, Y_a = foreign income, Z = balance of payments surplus, A^n = current account surplus, K = capital account surplus, e = exchange rate.

1. Why is a current account surplus associated with a capital account deficit in the case of flexible exchange rates? Why is this not necessarily the case under fixed exchange rates?
2. Assume flexible exchange rates. Analyze verbally and graphically (in a i/Y diagram) the effects of an increase in
   (a) government spending ($dG > 0$).
   (b) the money supply ($dM > 0$).
3. Show graphically that under flexible exchange rates and in the limit case of perfect capital mobility, fiscal policy is inefficient.
4. Why is monetary policy inefficient in case of a fixed exchange rate regime?

AD/AS model

Consider the flex-price model of a closed economy with wage rigidity:

$$\begin{aligned} N &= N^d(W/P) < N^s(W/P) \\ Y &= Y^s = Y(N, \overline{K}) \\ Y^d &= C((1-t)Y) + I(i) + G \\ M/P &= L(Y, i) \end{aligned}$$

where Y = income, C = consumption, I = investment, G = government spending, M = money supply, L = money demand, i = interest rate, P = flexible price level, N = employed labor, N^d = labor demand, N^s = labor supply, W = wage, $\overline{K}$ = fixed capital stock, t = income tax rate. Wages are inflexible downwards.

1. Explain the Keynes effect.
2. Analyze verbally and graphically (in a P/Y diagram) the effects of an increase in
   (a) government spending ($dG > 0$).
   (b) the money supply ($dM > 0$).
3. Assume that wages are completely flexible (Classical version) such that $N = N^d(W/P) = N^s(W/P)$. Show that fiscal and monetary policies are inefficient with respect to output in this case.

(Static) New Keynesian Macroeconomics

Consider the static approximation of the baseline New Keynesian model:

$$
\begin{aligned}
x &= a - b \cdot r + \varepsilon_1 \\
x &= y - \bar{y} \\
r &= i - \pi^e \\
i &= r^n + \pi^T + k_\pi(\pi - \pi^T) + k_x \cdot x \\
r^n &= \frac{a}{b} \\
\pi &= \pi^e + \delta \cdot x + \varepsilon_2
\end{aligned}
$$

where x = output gap, y = output, $\overline{y}$ = flex-price output, r = real interest rate, i = nominal interest rate, π^e = inflation expectations, r^n = natural real interest rate, π^T = inflation target level of the central bank, ε_1 = demand shock, ε_2 = supply shock. Assume that the inflation expectations are equal to the targeting rule of the central bank.

1. Explain the Taylor principle in the interest rule of Taylor-type.
2. Name the differences to the IS/LM-AS/AD model.
3. Analyze verbally and graphically (in a π/x diagram) the effects of a
   (a) positive demand shock.
   (b) positive supply shock.
   (c) anticipated and credible increase in the target level of the central bank ($\pi^e = \pi^T$).
   (d) unanticipated increase in the target level ($d\pi^T > 0$, $d\pi^e = 0$).

Dynamic overshooting model of Dornbusch-type

Consider the dynamic Dornbusch-type model of a small open economy:

$$
\begin{aligned}
y &= (a_0 + a_1 y - a_2(i - E(\dot{p}))) + g + (b_0 - b_1 y + b_2 y^* - b_3 \tau) \\
\tau &= p - (p^* + e) \\
m - p &= l_0 + l_1 y - l_2 i \\
i &= i^* + E(\dot{e}) \\
\dot{p} &= \pi + \delta(y - \overline{y})
\end{aligned}
$$

with: y = real income or real output, i = nominal interest rate, $i - \dot{p}$ = real interest rate, $\dot{p}$ = inflation rate, g = government expenditure, τ = terms of trade, e = flexible nominal exchange rate in price notation, $\dot{e}$ = rate of change of the exchange rate, m = money supply (exogenously given), $\dot{m}$ = growth rate of money supply (control variable of the central bank), $\overline{y}$ = natural or long-term output level, π = augmentation term of the Phillips curve. Foreign variables (y^*, i^*, p^*) are denoted by a superscript star. A dot above a variable stands for the derivative of this variable with respect to time t, e.g. $\dot{p} = dp/dt$.

The model can be reduced into a two-dimensional system of the form

$$\begin{pmatrix} a_2 & \lambda/\delta \\ l_2 & l_2 - l_1/\delta \end{pmatrix} \begin{pmatrix} \dot{\tau} \\ \dot{m}^r \end{pmatrix} = \begin{pmatrix} b_3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \tau \\ m^r \end{pmatrix} + \begin{pmatrix} \lambda \cdot \overline{y} - a_0 - b_0 - g - b_2 \cdot y^* + a_2(i^* - \dot{p}^*) \\ -l_0 - l_1 \cdot \overline{y} + l_2(i^* - \dot{p}^*) + l_2 \cdot \dot{m} \end{pmatrix}$$

with $\lambda = 1 - a_1 + b_1 > 0$ and $m^r = m - p$. Assume that $l_2 < l_1/\delta$, implying that the system matrix has one stable and one unstable root. In a τ/m^r-diagram, the stable saddle arm has a negative slope and the unstable arm has a positive slope.

1. Assume that the terms of trade are forward-looking and the real money stock is predetermined. Why does this assumption ensure the uniqueness and stability of the model solution?

2. Compute the steady state values of τ and m^r.

3. Analyze graphically in a τ/m^r-diagram the effects of $d\dot{m} = 1$.

Dynamic New Keynesian Macroeconomics – Microfoundation

The representative household seeks to maximize the objective function

$$E_t \sum_{k=0}^{\infty} \beta^k U(C_{t+k}, N_{t+k})$$

with period utility function

$$U(C_t, N_t) = \frac{C_t^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\eta}}{1+\eta}$$

where C_t is the quantity consumed, N_t is the hours of work, P_t is the price of the consumption good. Each period the household faces the following budget constraint

$$(1 + i_{t-1})B_{t-1} + W_t N_t - T_t^n = P_t C_t + B_t$$

where W_t is the nominal wage, B_t is the quantity of one-period, nominally riskless bonds, i_t is the corresponding nominal interest rate and T^n captures lump-sum taxes and dividends.

1. Derive the Euler equation by solving the household's optimization problem.

2. Log-linearize the Euler equation around the steady state.

Macroeconomic Questions: RESULTS

Goods market

1. $Y_0 = \frac{1}{1-b(1-t)}(a + I^a + G)$
2. $\frac{dY}{dI^a} = \frac{dY}{dG} = \frac{1}{1-b(1-t)}$
3. Due to taxes and savings.
4. $\frac{dY}{dG}\Big|_{dT=dG} = 1$ (Haavelmo-Theorem)

IS/LM model

1. (a) $G\uparrow \rightarrow Y\uparrow, i\uparrow$

   (b) $M\uparrow \rightarrow Y\uparrow, i\downarrow$
2. $\frac{dY}{dG} = \frac{1}{1-C_{Y^v}(1-t)-I_i L_y/L_i}$. $I_i = 0$ or $L_y \rightarrow -\infty$ (liquidity trap) give the elementary multiplier of the goods market model.

IS/LM/Z (Mundell-Fleming) model

1. e flexible implies $Z = 0$ and thus $A^n(Y, Y_a, e) = -K(i, i_a)$.
2. (a) $G\uparrow \rightarrow Y\uparrow, i\uparrow, e\uparrow$ (if capital mobility is sufficiently small)
   $G\uparrow \rightarrow Y\uparrow, i\uparrow, e\downarrow$ (if capital mobility is sufficiently large, but not perfect)

   (b) $M\uparrow \rightarrow Y\uparrow, i\downarrow, e\uparrow$
3. Under perfect capital mobility: $dA = -dG$ (complete crowding out).
4. Under fixed exchange rate: Decrease in reserves completely neutralizes the effects of the expansionary monetary policy.

AD/AS model

1. $P\downarrow \rightarrow \frac{M}{P}\uparrow \rightarrow i\downarrow \rightarrow Y\uparrow$
2. (a) $G\uparrow \rightarrow Y\uparrow, i\uparrow, P\uparrow$

   (b) $M\uparrow \rightarrow Y\uparrow, i\downarrow, P\uparrow$
3. Monetary policy: $dM = dP = dW \rightarrow dN = 0 \rightarrow dY = 0$.
   Fiscal policy: $dG = -dI \rightarrow dY = 0$ (investment crowding out).

(Static) New Keynesian Macroeconomics

1. $k_\pi > 1$.

2. Most important differences: Goods demand depends on real interest rate. Control variables of the central banks is the nominal interest rate. Supply side is described by an inflation equation (Phillips curve), which depends on expected future inflation.

3. (a) $\varepsilon_1 \uparrow \rightarrow x \uparrow, \pi \uparrow, i \uparrow$

   (b) $\varepsilon_1 \uparrow \rightarrow x \downarrow, \pi \uparrow, i \uparrow$

   (c) $\pi^T = \pi^e \uparrow \rightarrow dx = 0, \pi \uparrow, i \uparrow$

   (d) $\pi^T \uparrow (d\pi^e = 0) \rightarrow dx \uparrow, \pi \uparrow, i \downarrow$

Dynamic overshooting model of Dornbusch-type

1. Since system matrix has one stable and one unstable eigenvalue (cf. Blanchard-Kahn-conditions).

2.

$$\begin{pmatrix} \overline{\tau} \\ \overline{m}^r \end{pmatrix} = -\begin{pmatrix} \frac{1}{b_3}[\lambda \cdot \overline{y} - a_0 - b_0 - g - b_2 \cdot y^* + a_2(i^* - \dot{p}^*)] \\ -l_0 - l_1 \cdot \overline{y} + l_2(i^* - \dot{p}^*) + l_2 \cdot \dot{m} \end{pmatrix}$$

3. $d\overline{\tau} = 0$, $d\overline{m} < 0$. $\dot{\tau} \geq 0, \dot{m}^r \leq 0 \, \forall \, t > 0$.

Dynamic New Keynesian Macroeconomics – Microfoundation

1. $C_t^{-\sigma} = \beta(1 + i_t) E_t \left(\frac{P_t}{P_{t+1}}\right) E_t C_{t+1}^{-\sigma}$

2. $c_t = E_t c_{t+1} - \frac{1}{\sigma}(i_t - E_t \pi_{t+1} + \log \beta)$ with $\pi_t = \log P_t - \log P_{t-1}$ and $c_t = \log C_t - \log \overline{C}$. $\overline{C}$ is the steady state of C_t.

Microeconomic Questions

Firm Theory

1. Define the concept of a production function.

2. Define the concept of an isoquant.

3. What is the marginal product of a particular input?

4. Define the concept of increasing (constant, decreasing) returns to scale.

5. Consider the following production function

$$f(x_1, x_2) = x_1^a x_2^b, a, b > 0$$

   (a) Determine the marginal product of each factor.

   (b) Determine the technical rate of substitution of input 1 for input 2.

   (c) Show that the technical rate of substitution equals the slope of the isoquant.

   (d) Under what circumstances does this production function exhibit increasing, constant or decreasing returns to scale?

   (e) Let the output be denoted by y, the output price by p. Input prices for factors 1 and 2 are given by w_1 and w_2, respectively. Write down the profit equation.

6. A firm produces an output good q using one input good x according to the production function $f(x) = 4\sqrt{x}$. The output price is given by p, the input price is given by w. Determine the profit maximizing factor demand function and the supply function.

7. What does it mean if two production factors are perfect substitutes (perfect complements)?

8. Consider the following production functions. For which functions are the production factors perfect substitutes or perfect complements?

   (a) $f(x_1, x_2) = x_1 + x_2$

   (b) $f(x_1, x_2) = \sqrt{x_1} + \sqrt{x_2}$

   (c) $f(x_1, x_2) = \sqrt{ax_1 + bx_2}$

   (d) $f(x_1, x_2) = min\{a\sqrt{x_1}, b\sqrt{x_2}\}$

   (e) $f(x_1, x_2) = [x_1^\rho + x_2^\rho]^{\frac{1}{\rho}}$

9. Define the concept of a profit function and of a cost function.

Consumer Theory

1. Explain what it means if a utility function represents a certain preference ordering. Explain the concept of an indifference curve.

2. Sweetie only consumes chocolate bars and lollies. The amounts of chocolate bars and lollies are given by x_C and x_L, respectively. Her preferences can be represented by the following utility function:

$$U(x_C, x_L) = ln(x_C) + ln(x_L)$$

   Prices for chocolate bars and lollies are given by p_C and p_L, respectively. Sweetie's income is denoted by m.

   (a) Write down her budget constraint.

   (b) Determine the utility maximizing consumption bundle.

   (c) Assume prices are given by $p_C = p_L = 2$. Sweetie's income is given by m=500. How many chocolate bars and how many lollies does Sweetie consume in the optimum?

   (d) Assume the price of chocolate bars decreases to $p'_C = 1$. Determine the new consumption bundle. Explain why the change in optimal consumption (the total effect of the price change) can be decomposed into a substitution effect and an income effect

3. Oskar's preferences for pears and apples can be represented by the following utility function:

$$U(x_1, x_2) = 2x_1 + x_2$$

   where x_1 denotes the amount of pears and x_2 the amount of apples he consumes. The price of a pear is given by $p_1 = 0.5$, the price of an apple is given by $p_2 = 0.2$.

   (a) How many pears and apples will Oskar consume if he has 3 Euros?

   (b) Determine the marginal rate of substitution of pears for apples and interpret your result.

4. Explain the difference between a normal and an inferior good.

Markets

1. Assume the demand for bikes is given by

$$Q^D = 50 - \tfrac{3}{2}P$$

   Supply is given by

$$Q^S = \tfrac{1}{2}P.$$

   (a) Determine equilibrium price and quantity.

   (b) Determine the price elasticity of supply and demand in equilibrium and interpret your results.

2. Assume the demand for strawberries (in kg) in Schleswig-Holstein is given by

$$Q^D = 2000 - 300P.$$

   Supply is given by

$$Q^S = 200P.$$

   (a) Determine equilibrium price and quantity.

   (b) Assume a subsidy s=1 per kilogram strawberries is introduced. Determine the new equilibrium (i.e. quantity, producers' price P^S and consumers' price P^c).

3. A monopolist with cost function $C(y) = 6y + 28$ is facing the inverse demand function $P(y) = 30 - 2y$. Here, y denotes his level of output. Determine monopoly price and quantity, profit of the monopolist as well as producers' and consumers' surplus.

Microeconomic Questions: RESULTS

Firm Theory

1. We may look at a firm that produces one output y using two inputs. The production function $f(x_1, x_2)$, with $x_1, x_2 \geq 0$, measures the maximum possible output for given input quantities x_1, x_2.

2. The isoquant is the set of all input bundles that produce exactly y units of output.

3. The marginal product of input 1 is the additional amount of output we get if we increase input 1 by a marginal unit while keeping input 2 constant:

$$MP_1 = \frac{\partial f(x_1, x_2)}{x_1} \geq 0$$

4. What happens to output if we scale the amount of all inputs up by some constant t?

   - Constant returns to scale:

$$f(t \cdot x_1, t \cdot x_2) = t \cdot f(x_1, x_2) = t \cdot y \;\forall\; t \geq 0 \tag{1}$$

   (If we double all inputs, output also doubles.)

   - Increasing returns to scale:

$$f(t \cdot x_1, t \cdot x_2) > t \cdot f(x_1, x_2) = t \cdot y \;\forall\; t \geq 1 \tag{2}$$

   (If we double all inputs, output more than doubles.)

   - Decreasing returns to scale:

$$f(t \cdot x_1, t \cdot x_2) < t \cdot f(x_1, x_2) = t \cdot y \;\forall\; t \geq 1 \tag{3}$$

   (If we double all inputs, output less than doubles.)

5. (a) $MP_1 = ax_1^{a-1}x_2^b, MP_2 = bx_1^a x_2^{b-1}$

   (b) $TRS_{1,2} = -\frac{ax_2}{bx_1}$

   (c) We look at a vector of (small) input changes (dx_1, dx_2). The resulting change in output is approximated by:

$$dy = \frac{\partial f(x_1, x_2))}{\partial x_1} dx_1 + \frac{\partial f(x_1, x_2)}{\partial x_2} dx_2$$

Along an isoquant, output stays constant:

$$0 = \frac{\partial f(x_1, x_2)}{\partial x_1} dx_1 + \frac{\partial f(x_1, x_2)}{\partial x_2} dx_2$$

$$\Rightarrow \frac{dx_2}{dx_1} = -\frac{\frac{\partial f(x_1,x_2))}{\partial x_1}}{\frac{\partial f(x_1,x_2))}{\partial x_2}} = TRS_{1,2}$$

(d) $a + b > 1, a + b = 1, a + b < 1$, respectively.

(e) $\Pi = py - w_1x_1 - w_2x_2 = px_1^a x_2^b - w_1x_1 - w_2x_2$

6. $x(p, w) = \frac{4p^2}{w^2}, y(p, w) = \frac{8p}{w}$

7. Perfect complements are input goods that are always used together in fixed proportions. For example, one needs always two tires and one frame to produce one bicycle.

   Two input goods are perfect substitutes if the firm is able to substitute one good for the other at a constant rate.

8. (a) Perfect substitutes
   (b) Neither nor
   (c) Perfect substitutes
   (d) Perfect complements
   (e) Neither nor

9. The profit function yields maximal profits for any given vector of input and output prices: $\Pi(p, w_1, w_2)$.

   The cost function measures the minimum cost of producing a given output level for some fixed input prices: $C(y, w_1, w_2)$

Consumer Theory

1. A utility function represents a preference ordering if it assigns a higher number to more-preferred bundles than to less-preferred bundles:
   $(x_1, x_2) \succeq (y_1, y_2) \Leftrightarrow u(x_1, x_2) \geq u(y_1, y_2)$.

   An indifference curve is the set of bundles of goods that yield a certain utility level $\bar{U}$. Thus, the consumer is indifferent between all the bundles of this set.

2. (a) $m = p_C x_C + p_L x_L$
   (b) $x_C^* = \frac{1}{2}\frac{m}{p_C}, x_L^* = \frac{1}{2}\frac{m}{p_L}$
   (c) She consumes 125 chocolate bars and 125 lollies.

(d) $x_C^{'} = 250, x_L^{'} = 125$

Relative prices change such that chocolate bars become relatively cheaper (substitution effect) and real income increases (income effect).

3. (a) Oskar consumes 15 apples and no pear.

   (b) $MRS_{1,2} = 2$. Oskar is willing to give up 2 apples for one additional pear. Apples and pears are perfect substitutes.

4. How does a consumer's demand for a good change as his income changes?

   If good i is a normal good, then the demand for it increases when income increases: $\frac{\partial x_i}{\partial m} > 0$.

   If good j is an inferior good, then the demand for it decreases when income increases: $\frac{\partial x_j}{\partial m} < 0$

Markets

1. (a) $P^* = 25, Q^* = 12.5$

   (b) Demand elasticity: $\epsilon_{Q^D,P} = -3$. In equilibrium, an increase in the price of one per cent leads to a decrease in demand of three per cent.

   Supply elasticity: $\epsilon_{Q^S,P} = 1$. In equilibrium, an increase in the price of one per cent leads to an increase in supply of one per cent.

2. (a) $P^* = 4, Q^* = 800$

   (b) $P^c = 3.6, P^S = 4.6, Q^* = 920$

3. Price: $p^M = 18$, Quantity: $y^M = 6$, Profit: $\Pi^M = 44$,
   Producer's surplus: $PS^M = 72$, $CS^M = 36$

Recommended literature:

Students interested in applying for the Master's degree programme of Kiel University should dispose of knowledge in the fields of microeconomics, macroeconomics, mathematics, statistics, and econonometrics comparable to the contents of the following literature (or comparable work):

Microeconomics:

- Varian, H. R. (2010): Intermediate Microeconomics: A Modern Approach, 8th edition, W. W. Norton & Company, New York.
- Pindyck, R. S., Rubinfeld, D.L. (2005): Microeconomics, 6th edition, Pearson Prentice Hall, Upper Saddle River, N.J.

Macroeconomics:

- Blanchard, O., Macroeconomics, Prentice Hall
- Mankiw, N., Macroeconomics, Worth Publishers
- Sydsæter, K. et al., Further Mathematics for Economic Analysis, Prentice Hall

Mathematics:

- Sydsæter, K., P. Hammond, Essential Mathematics for Economic Analysis, Prentice Hall

Statistics:

- Mc Clave, J., P. Benson, T. Sincich, Statistics for Business and Economics, Pearson

Econometrics:

- J.H. Stock and M.M. Watson (2012) Introduction to Econometrics, 3.Ed., Pearson (International Edition), Chapters 1-9,17, and 17.1-5.