# 2. The multiple regression model

## 2.1 Model Specification

In an **econometric single equation model**, an endogenous variable y is dependent on one or more exogenous variables. In general, an econometric single equation model can be presented as a **multiple regression model** composed of a systematic and an unsystematic component. For linear models, the systematic component is given by a linear combination of k exogenous variables $x_1, x_2, ..., x_k$ and the unsystematic component is given by a disturbance variable u.

**The multiple regression model**:

(2.1a) $y_t = \beta_1 + \beta_2 \cdot x_{2t} + ... + \beta_k \cdot x_{kt} + u_t, \quad t = 1,2,...n$

The variable $x_1$ is a pseudo variable relating to the absolute term that for all t (= periods or cross-sectional units) takes on the value 1:

$$x_{1t} = 1 \text{ for all } t$$

More compact model form:

(2.1b) $y_t = \sum_{j=1}^{k} \beta_j \cdot x_{jt} + u_t = \mathbf{x}'_t \boldsymbol{\beta} + u_t$

with $\underset{1xk}{\mathbf{x}'_t} = (1 \quad x_{2t} \quad \ldots \quad x_{kt})$ and $\underset{1xk}{\boldsymbol{\beta}'} = (\beta_1 \quad \beta_2 \quad \ldots \quad \beta_k)$

Regression coefficients: ß$_1$, ß$_2$, ..., ß$_k$

absolute term (intercept): ß$_1$
Slopes: ß$_2$, ..., ß$_k$

**Assumptions** about the unobservable **stochastic disturbance** u:

**A1:**
The expected value of disturbance variables is in all periods equal to 0, i.e., the disturbance term has no systematic effect on the endogenous variable y:

(2.2) $E(u_t) = 0$ for all t = 1, ... n

**A2:**
The variance of the disturbance term u is constant over time (homoscedasticity):

(2.3) $\mathrm{Var}(u_t) = E[u_t - \underbrace{E(u_t)}_{=0}]^2 = E(u_t^2) = \sigma^2$ for all t = 1, ... n

**A3:**
Between the disturbance terms at different points in time there are no linear dependencies, meaning their covariance is 0. The disturbance term is therefore uncorrelated (absence of autocorrelation):

(2.4) $\mathrm{Cov}(u_t, u_{t-j}) = E\{[u_t - \underbrace{E(u_t)}_{=0}][u_{t-j} - \underbrace{E(u_{t-j})}_{=0}]\} = E(u_t u_{t-j}) = 0$ for J≠0

Equation (2.1a) applies to all observations at the different times t = 1, ... n:

$$\begin{array}{ccccccc} y_1 & = & \beta_1 + & \beta_2 x_{21} + & \beta_3 x_{31} + & \cdots + & \beta_k x_{k1} + & u_1 \\ y_2 & = & \beta_1 + & \beta_2 x_{22} + & \beta_3 x_{32} + & \cdots + & \beta_k x_{k2} + & u_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_n & = & \beta_1 + & \beta_2 x_{2n} + & \beta_3 x_{3n} + & \cdots + & \beta_k x_{kn} + & u_n \end{array}$$

or compact

$$(2.5) \qquad \underset{n\text{x}1}{\mathbf{y}} = \underset{n\text{x}k}{\mathbf{X}} \; \underset{k\text{x}1}{\boldsymbol{\beta}} + \underset{n\text{x}1}{\mathbf{u}}$$

with

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & x_{21} & \cdots & x_{k1} \\ 1 & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{2n} & \cdots & x_{kn} \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}$$

**Assumptions** about the **disturbance term** u in a compact form:

**A1:**
The vector of the expected values of the disturbance term u is equal to the zero vector **0**:

(2.6) $$\mathrm{E}(\mathbf{u}) = \underset{nx1}{\mathbf{0}}$$

in more detail: $$\mathrm{E}(\mathbf{u}) = \mathrm{E}\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} \mathrm{E}(u_1) \\ \mathrm{E}(u_2) \\ \vdots \\ \mathrm{E}(u_n) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{0}$$

**A2:**
The variance-covariance matrix of the error terms is a scalar matrix (= scalar multiple of the identity matrix **I**):

(2.7) $$\mathrm{Cov}(\mathbf{u}) = \mathrm{E}(\mathbf{u}\mathbf{u}') = \sigma^2 \underset{nxn}{\mathbf{I}}$$

The outer product **uu**' (product of column and row vector):

$$\mathbf{uu}' = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \left(u_1, u_2, \ldots, u_n\right) = \begin{pmatrix} u_1^2 & u_1 u_2 & \cdots & u_1 u_n \\ u_2 u_1 & u_2^2 & \cdots & u_2 u_n \\ \vdots & \vdots & \cdots & \vdots \\ u_n u_1 & u_n u_2 & \cdots & u_n^2 \end{pmatrix}$$

Formation of expectation:

$$\mathrm{Cov}(\mathbf{u}) = \mathrm{E}(\mathbf{uu}') = \begin{pmatrix} \mathrm{E}(u_1^2) & \mathrm{E}(u_1 u_2) & \cdots & \mathrm{E}(u_1 u_n) \\ \mathrm{E}(u_2 u_1) & \mathrm{E}(u_2^2) & \cdots & \mathrm{E}(u_2 u_n) \\ \vdots & \vdots & \cdots & \vdots \\ \mathrm{E}(u_n u_1) & \mathrm{E}(u_n u_2) & \cdots & \mathrm{E}(u_n^2) \end{pmatrix} = \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & & \\ \vdots & & \ddots & \\ 0 & & & \sigma^2 \end{pmatrix}$$

Factoring out σ²:

$$\mathrm{Cov}(\mathbf{u}) = \mathrm{E}(\mathbf{uu}') = \sigma^2 \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{pmatrix} = \sigma^2 \mathbf{I}$$

● The vector of the expected values and the variance-covariance matrix of the endogenous variable y

- Vector of the expected values of y:

(2.8) $$\mathrm{E}(\mathbf{y}) = \mathrm{E}(\mathbf{X\beta} + \mathbf{u}) = \mathrm{E}(\mathbf{X\beta}) + \underbrace{\mathrm{E}(\mathbf{u})}_{=\mathbf{0}} = \mathbf{X\beta}$$

The random vector of the endogenous variables y equals **Xß**,

- Variance-covariance matrix of y:

(2.9) $$\mathrm{E}\{[\mathbf{y} - \mathrm{E}(\mathbf{y})][\mathbf{y} - \mathrm{E}(\mathbf{y})]'\} = \mathrm{E}[\underbrace{(\mathbf{y} - \mathbf{X\beta})}_{=\mathbf{u}}\underbrace{(\mathbf{y} - \mathbf{X\beta})'}_{=\mathbf{u}'}] = \mathrm{E}(\mathbf{uu}') = \sigma^2\,\mathbf{I}$$

The variance-covariance matrix of the endogenous variables y is equal to the variance-covariance matrix of the disturbance variables u.

## 2.2 Method of Least Squares (OLS method)

Ordinary least squares method [ordinary least squares (OLS)]:

Determination of estimation values for the k regression coefficients$_1$, ß$_2$, ..., ß$_k$ of the vector **ß**

**Least squares criterion**:

Minimizing the sum of the squared error terms:

(2.10a) $$S = \sum_{t=1}^{n} u_t^2 = \mathbf{u}'\mathbf{u}$$

or

(2.10b) $$S = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

(2.11) $$S = \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$$

With $\left(\underset{1xk}{\boldsymbol{\beta}'}\ \underset{kxn}{\mathbf{X}'}\ \underset{nx1}{\mathbf{y}}\right) = \mathbf{y}'\mathbf{X}\boldsymbol{\beta}$ we get

(2.12) $$S = \mathbf{y}'\mathbf{y} - 2\mathbf{y}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$$

Necessary condition for a minimum of S

(2.13) $$\frac{\partial S}{\partial \boldsymbol{\beta}} = -2\mathbf{X'y} + 2(\mathbf{X'X})\boldsymbol{\beta} \overset{!}{=} \mathbf{0}$$

Normal equation system:

(2.14) $$(\mathbf{X'X})\boldsymbol{\beta} = \mathbf{X'y}$$

OLS estimator of **ß**:

(2.15) $$\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'y}$$

Conditions for the existence of the inverse (**X'X**) $^{-1}$ :

- rank(**X**) = rank(**X'X**) = k < n

- No linear dependence between the **X**-Columns

- Regression equation resulting from OLS estimation:

(2.16) $$y_t = \hat{\beta}_1 + \hat{\beta}_2 \cdot x_{2t} + ... + \hat{\beta}_k \cdot x_{kt} + \hat{u}_t, \quad t = 1,2,...n$$

$\hat{\beta}_1, \hat{\beta}_2, ....., \hat{\beta}_k$: estimated regression coefficients

$\hat{u}_t$: residual (= estimated value of the disturbance)

- Regression values (estimated values of the dependent variable y):

(2.17) $$\hat{y}_t = \hat{\beta}_1 + \hat{\beta}_2 \cdot x_{2t} + ... + \hat{\beta}_k \cdot x_{kt}, \quad t = 1,2,...n$$

- vector $\hat{\mathbf{y}}$ of regression values:

(2.18) $$\hat{\mathbf{y}} = \mathbf{X} \cdot \hat{\boldsymbol{\beta}}$$ With $\hat{\mathbf{y}} = (\hat{y}_1 \;\; \hat{y}_2 \;\; ... \;\; \hat{y}_n)'$

- Residuals:

(2.19) $$\hat{u}_t = y_t - \hat{y}_t, \quad t = 1,2,...,n \;,$$

- Vector of the residuals:

(2.20) $$\hat{\mathbf{u}} = \mathbf{y} - \hat{\mathbf{y}}$$ With $\hat{\mathbf{u}} = (\hat{u}_1 \;\; \hat{u}_2 \;\; ... \;\; \hat{u}_n)'$

Special case: **Simple regression model** (K = 2)

(2.21) $$y_t = \beta_1 + \beta_2 \cdot x_t + u_t$$

Estimated regression equation:

(2.22) $$\hat{y}_t = \hat{\beta}_1 + \hat{\beta}_2 \cdot x_t$$

$\hat{\beta}_1$: intercept (absolute term)

$\hat{\beta}_2$ : slope

Determining the inverse (**X'X**)$^{-1}$

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{|\mathbf{X}'\mathbf{X}|}(\mathbf{X}'\mathbf{X})^{adj}$$

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \sum x_t \\ \sum x_t & \sum x_t^2 \end{bmatrix}$$

determinant: $$|\mathbf{X}'\mathbf{X}| = n \cdot \sum x_t^2 - \left(\sum x_t\right)^2$$

The adjoint matrix $(\mathbf{X'X})^{adj}$ consists of the adjoints

$$XX_{ij} = (-1)^{i+j} \left| XX_{ij} \right|$$

$$\begin{aligned} XX_{11} &= \sum x_t^2 \\ XX_{12} &= -\sum x_t \\ XX_{21} &= -\sum x_t \\ XX_{22} &= n \end{aligned}$$

$$(\mathbf{X'X})^{adj} = \begin{bmatrix} XX_{11} & XX_{21} \\ XX_{12} & XX_{22} \end{bmatrix} = \begin{bmatrix} \sum x_t^2 & -\sum x_t \\ -\sum x_t & n \end{bmatrix}$$

inverse of **X'X**:

$$(\mathbf{X'X})^{-1} = \frac{1}{n \cdot \sum x_t^2 - (\sum x_t)^2} \begin{bmatrix} \sum x_t^2 & -\sum x_t \\ -\sum x_t & n \end{bmatrix}$$

Product of the transpose of **X** and the vector **y:**

$$\mathbf{X'y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum y_t \\ \sum x_t y_t \end{bmatrix}$$

OLS estimation vector $\hat{\boldsymbol{\beta}}$:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \frac{1}{n\sum x_t^2 - (\sum x_t)^2}\begin{bmatrix} \sum x_t^2 \sum y_t - \sum x_t \sum x_t y_t \\ n\sum x_t y_t - \sum x_t \sum y_t \end{bmatrix}$$

intercept (absolute term):

$$(2.23) \quad \hat{\beta}_1 = \frac{\sum x_t^2 \sum y_t - \sum x_t \sum x_t y_t}{n\sum x_t^2 - (\sum x_t)^2}$$

Slope:

$$(2.24) \quad \hat{\beta}_2 = \frac{n\sum x_t y_t - \sum x_t \sum y_t}{n\sum x_t^2 - (\sum x_t)^2}$$

Relationship between slope measure and covariance of X and Y:

$$(2.25) \quad \hat{\beta}_2 = \frac{s_{xy}}{s_x^2}$$

with $s_{xy} = \frac{1}{n}\sum(x_t - \bar{x})(y_t - \bar{y})$ and $s_x^2 = \frac{1}{n}\sum(x_t - \bar{x})^2$

Relationship between intercept and slope:

$$\text{(2.26)} \quad \hat{\beta}_1 = \frac{\sum y_t}{n} - \hat{\beta}_2 \frac{\sum x_t}{n} = \bar{y} - \hat{\beta}_2 \cdot \bar{x}$$

course of the regression line through the centroid $(\bar{x}, \bar{y})$

$$\bar{y} = \hat{\beta}_1 + \hat{\beta}_2 \cdot \bar{x}$$

## Example: Macroeconomic consumption function

The disposable income $Y^v$ is considered to be a crucial influencing factor of private consumption C. The regression model of the macroeconomic consumption function

$$(2.27) \qquad C_t = C_0 + c_1 \cdot Y_t^v + u_t$$

can be estimated with the OLS method under the inclusion of error terms $u_t$ that meet the standard assumptions A1-A3. $u_t$ includes all unsystematic influencing factors impacting on aggregate consumption.

The parameter $c_1$ is known from macroeconomics as the marginal propensity to consume:

$$\frac{dC_t}{dY_t^v} = c_1$$

It indicates the change in consumer spending, which is induced by an increase or decrease in disposable income by 1 unit. The intecept $C_0$ on the other hand indicates autonomous consumption, which is independent of disposable income.

The consumption function for the Federal Republic of Germany for the period from 1994 to 2012 is now to be estimated with the OLS method with the use of annual data.

| Year | Disposable Income bln. € (at current prices) | Price index Private Consumption (2010=100) | Private Consumption bln. € (in prices of 2010) |
|---|---|---|---|
| 1994 | 917,37 | 79,1 | 1032,75 |
| 1995 | 958,14 | 80,5 | 1066,47 |
| 1996 | 990,08 | 81,6 | 1088,64 |
| 1997 | 1024,89 | 83,2 | 1110,82 |
| 1998 | 1050,81 | 84,0 | 1130,14 |
| 1999 | 1080,15 | 84,5 | 1161,86 |
| 2000 | 1114,59 | 85,7 | 1195,04 |
| 2001 | 1177,85 | 87,4 | 1233,43 |
| 2002 | 1208,30 | 88,6 | 1240,58 |
| 2003 | 1248,88 | 89,6 | 1264,51 |
| 2004 | 1291,88 | 91,0 | 1283,61 |
| 2005 | 1338,80 | 92,5 | 1306,98 |
| 2006 | 1390,33 | 93,9 | 1339,54 |
| 2007 | 1443,28 | 96,1 | 1356,73 |
| 2008 | 1519,54 | 98,6 | 1389,62 |
| 2009 | 1516,28 | 98,9 | 1391,55 |
| 2010 | 1578,89 | 100,0 | 1433,16 |
| 2011 | 1663,83 | 102,1 | 1487,66 |
| 2012 | 1734,64 | 104,1 | 1521,59 |

Source: Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung, Annual Report 2013/14

Let $y_t$ denote the real value of private consumption ($C_t$) and $x_t$ the real value of disposable income $\left(Y_t^V\right)$ in year t. The observation matrix **X** and the vector **y** of the dependent variables are then given by:

$$\underset{19\,x\,2}{X} = \begin{bmatrix} 1 & 1159{,}39 \\ 1 & 1190{,}61 \\ 1 & 1212{,}71 \\ 1 & 1231{,}47 \\ 1 & 1251{,}21 \\ 1 & 1278{,}66 \\ 1 & 1300{,}70 \\ 1 & 1347{,}78 \\ 1 & 1363{,}26 \\ 1 & 1394{,}62 \\ 1 & 1419{,}00 \\ 1 & 1448{,}13 \\ 1 & 1480{,}52 \\ 1 & 1502{,}37 \\ 1 & 1541{,}25 \\ 1 & 1533{,}14 \\ 1 & 1578{,}89 \\ 1 & 1630{,}14 \\ 1 & 1666{,}05 \end{bmatrix} \quad \text{und} \quad \underset{19\,x\,1}{y} = \begin{bmatrix} 1032{,}75 \\ 1066{,}47 \\ 1088{,}64 \\ 1110{,}82 \\ 1130{,}14 \\ 1161{,}86 \\ 1195{,}04 \\ 1233{,}43 \\ 1240{,}58 \\ 1264{,}51 \\ 1283{,}61 \\ 1306{,}98 \\ 1339{,}54 \\ 1356{,}73 \\ 1389{,}62 \\ 1391{,}55 \\ 1433{,}16 \\ 1487{,}66 \\ 1521{,}59 \end{bmatrix}$$

Work table for determining the OLS estimators of $\beta_1$ and $\beta_2$

| $t$ | $x_t$ | $y_t$ | $x_t^2$ | $x_t \cdot y_t$ |
|---|---|---|---|---|
| 1 | 1159,39 | 1032,75 | 1344185,17 | 1197360,02 |
| 2 | 1190,61 | 1066,47 | 1417552,17 | 1269749,85 |
| 3 | 1212,71 | 1088,64 | 1470665,54 | 1320204,61 |
| 4 | 1231,47 | 1110,82 | 1516518,36 | 1367941,51 |
| 5 | 1251,21 | 1130,14 | 1565526,46 | 1414042,47 |
| 6 | 1278,66 | 1161,86 | 1634971,40 | 1485623,91 |
| 7 | 1300,70 | 1195,04 | 1691820,49 | 1554388,53 |
| 8 | 1347,78 | 1233,43 | 1816510,93 | 1662392,29 |
| 9 | 1363,26 | 1240,58 | 1858477,83 | 1691233,09 |
| 10 | 1394,62 | 1264,51 | 1944964,94 | 1763510,94 |
| 11 | 1419,00 | 1283,61 | 2013561,00 | 1821442,59 |
| 12 | 1448,13 | 1306,98 | 2097080,50 | 1892676,95 |
| 13 | 1480,52 | 1339,54 | 2191939,47 | 1983215,76 |
| 14 | 1502,37 | 1356,73 | 2257115,62 | 2038310,45 |
| 15 | 1541,25 | 1389,62 | 2375451,56 | 2141751,83 |
| 16 | 1533,14 | 1391,55 | 2350518,26 | 2133440,97 |
| 17 | 1578,89 | 1433,16 | 2492893,63 | 2262801,99 |
| 18 | 1630,14 | 1487,66 | 2657356,42 | 2425094,07 |
| 19 | 1666,05 | 1521,59 | 2775722,60 | 2535045,02 |
| $\sum$ | 26529,90 | 24034,68 | 37472832,36 | 33960226,83 |

Absolute term:

$$\hat{c}_0 = \frac{\sum x_t^2 \sum y_t - \sum x_t \sum x_t y_t}{n\sum x_t^2 - \left(\sum x_t\right)^2} = \frac{37472832{,}36 \cdot 24034{,}68 - 26529{,}90 \cdot 33960226{,}83}{19 \cdot 37472832{,}36 - \left(26529{,}90\right)^2}$$

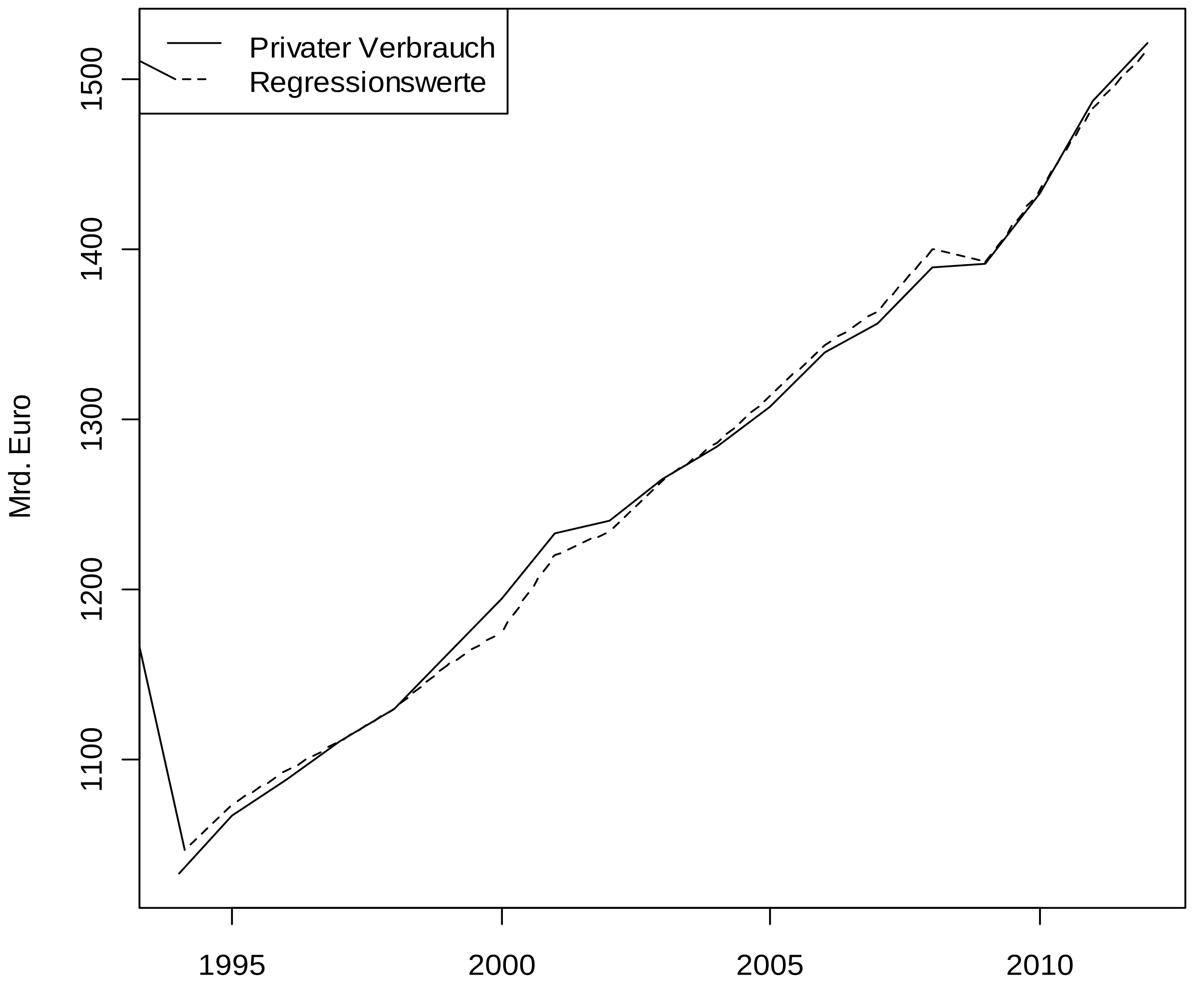
$$= \frac{-313887342{,}31}{8148220{,}83} = -38{,}52$$

Slope estimate:

$$\hat{c}_1 = \frac{n \cdot \sum x_t y_t - \sum x_t \sum y_t}{n\sum x_t^2 - \left(\sum x_t\right)^2} = \frac{19 \cdot 33960226{,}83 - 26529{,}90 \cdot 24034{,}68}{19 \cdot 37472832{,}36 - \left(26529{,}90\right)^2}$$

$$= \frac{7606652{,}86}{8148220{,}83} = 0{,}934$$

Figure: Consumption Function



## Example: money demand function

The transaction volume and the opportunity costs of holding money are regarded as the decisive factors for the demand for money. While the transaction volume is generally measured by the gross domestic product y, the opportunity costs are captured by the returns on alternative assets, mostly characterized by a representative rate of interest r. In the absence of money illusion, the demand for money m is then the function

$$m = f(y, r)$$

where m and y are real values obtained by relating the nominal money supply M and the nominal domestic product Y to the price level P:

$$m = \frac{M}{P} \text{ und } \quad y = \frac{Y}{P}$$

In the log-linear approach that is often used in empirical investigations of the demand for money, the econometric money demand function is then

(2.28) $$\ln m_t = \beta_1 + \beta_2 \ln y_t + \beta_3 \ln r_t + u_t$$

with the error term $u_t$, which is required to meet the standard assumptions.

In this model the regression coefficients $\beta_2$ and $\beta_3$ can immediately be interpretated as income and interest elasticities. Since the need for money increases with increasing transaction volume, in economic theory a positive income elasticity is assumed. On the other hand, a negative interest elasticity is to be expected because the opportunity costs of holding money increase with rising interest rates.

In the econometric estimation of the money demand function, the growth of the amount of money M1, which consists of the cash and the deposits holdings of the public, is used as the to be explained variable. The representative rate of interest is captured by EURIBOR (euro interbank offered rate) indicating the short-term interest rate. As a deflator for the determination of the real money supply, the price index for gross domestic product to the base 2010 is used. During the observation period from 1994 to 2012, this results in the following database:

| Year | Money demand $M_1$ bln. € | GDP bln. € (in prices of 2010) | Short-term interest rate % | GDP deflator (2010=100) |
|---|---|---|---|---|
| 1994 | 360,90 | 1782,2 | 5,3 | 87,7 |
| 1995 | 378,80 | 1848,5 | 4,5 | 89,5 |
| 1996 | 405,00 | 1875,0 | 3,3 | 90,0 |
| 1997 | 456,80 | 1912,6 | 3,3 | 90,3 |
| 1998 | 465,00 | 1959,7 | 3,5 | 90,8 |
| 1999 | 513,40 | 2000,2 | 2,9 | 91,0 |
| 2000 | 557,60 | 2047,5 | 4,3 | 90,4 |
| 2001 | 574,60 | 2101,9 | 4,2 | 91,4 |
| 2002 | 601,60 | 2132,2 | 3,3 | 92,7 |
| 2003 | 583,50 | 2147,5 | 2,3 | 93,7 |
| 2004 | 631,90 | 2195,7 | 2,1 | 94,7 |
| 2005 | 655,40 | 2224,4 | 2,1 | 95,3 |
| 2006 | 725,80 | 2313,9 | 3,0 | 95,6 |
| 2007 | 760,00 | 2428,5 | 4,2 | 97,2 |
| 2008 | 789,70 | 2473,8 | 4,2 | 97,9 |
| 2009 | 1031,90 | 2374,5 | 1,2 | 99,1 |
| 2010 | 1116,80 | 2496,2 | 0,8 | 100,0 |
| 2011 | 1171,50 | 2592,6 | 1,3 | 100,8 |
| 2012 | 1373,10 | 2643,9 | 0,5 | 102,1 |

Sources: Monthly reports of the Deutsche Bundesbank (various issues); Annual Report 2013/4 of the Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung

Since the money demand function has a logarithmic-linear shape, the logarithms of the observed values are first of all to be determined:

| t | $\ln m_t$ | $\ln y_t$ | $\ln r_t$ |
|---|---|---|---|
| 1 | 5,75761268 | 7,48560384 | 1,66770682 |
| 2 | 5,82592289 | 7,52212978 | 1,5040774 |
| 3 | 5,89906655 | 7,53636394 | 1,19392247 |
| 4 | 6,02206799 | 7,55621885 | 1,19392247 |
| 5 | 6,04575449 | 7,58054668 | 1,25276297 |
| 6 | 6,14676449 | 7,60100245 | 1,06471074 |
| 7 | 6,22262483 | 7,62437482 | 1,45861502 |
| 8 | 6,26377301 | 7,65059698 | 1,43508453 |
| 9 | 6,32398055 | 7,66490959 | 1,19392247 |
| 10 | 6,30427078 | 7,67205965 | 0,83290912 |
| 11 | 6,39467982 | 7,69425618 | 0,74193734 |
| 12 | 6,43731308 | 7,7072425 | 0,74193734 |
| 13 | 6,54243705 | 7,74668969 | 1,09861229 |
| 14 | 6,60469809 | 7,79502906 | 1,43508453 |
| 15 | 6,65065447 | 7,81351071 | 1,43508453 |
| 16 | 6,92986801 | 7,77254217 | 0,18232156 |
| 17 | 7,01822273 | 7,82252485 | -0,22314355 |
| 18 | 7,07401524 | 7,86041651 | 0,26236426 |
| 19 | 7,24585979 | 7,88001038 | -0,69314718 |
| $\sum$ | 121,708 | 145,986 | 17,779 |

$$\mathbf{X} = \begin{bmatrix} 1 & \ln y_1 & \ln r_2 \\ 1 & \ln y_2 & \ln r_2 \\ \vdots & \vdots & \vdots \\ 1 & \ln y_n & \ln r_n \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} \ln m_1 \\ \ln m_2 \\ \vdots \\ \ln m_3 \end{bmatrix}$$

OLS estimation of the regression coefficients $\beta_1$. $\beta_2$ and $\beta_3$

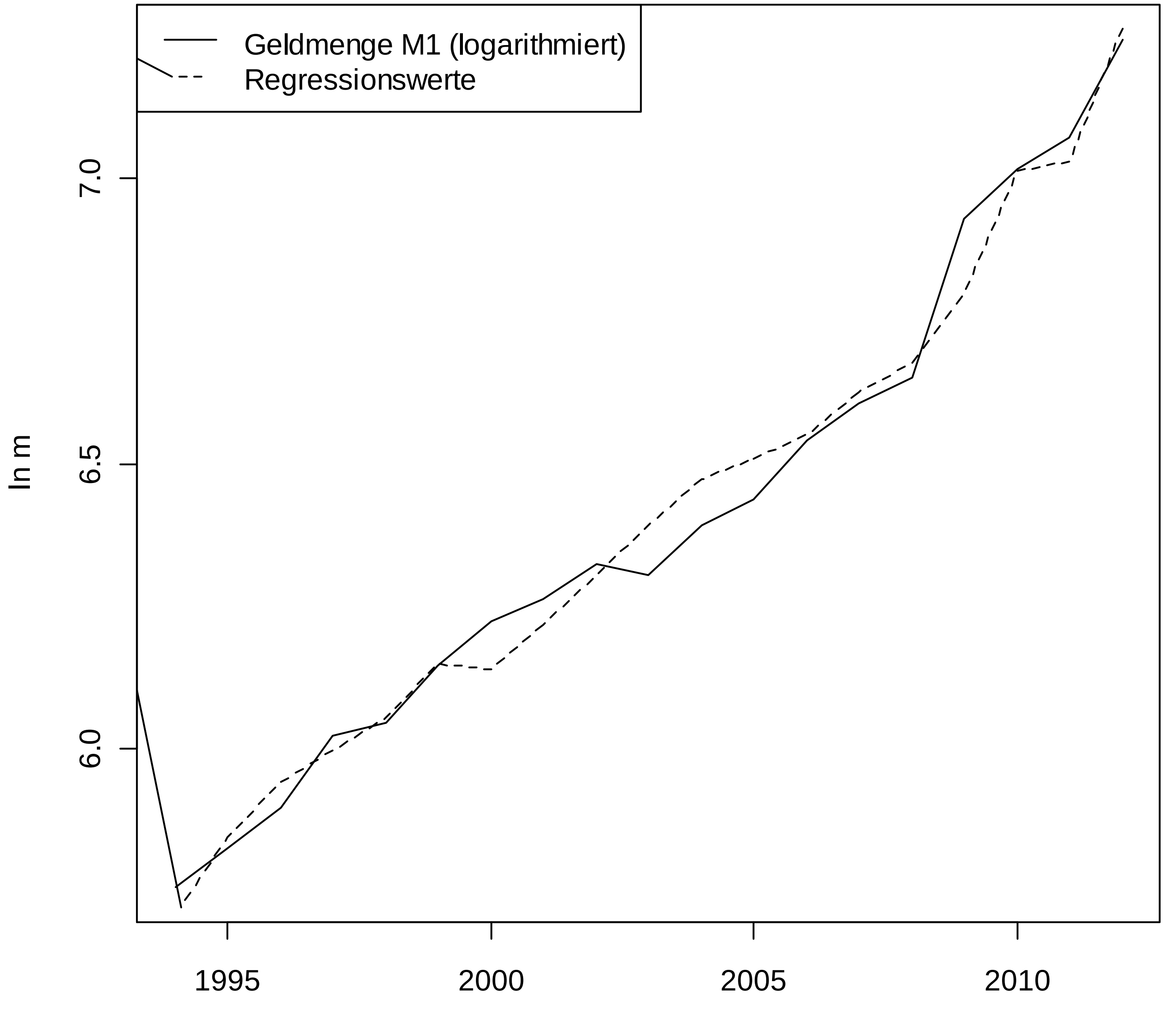
$$\underset{3x3}{X'X} = \begin{bmatrix} n & \sum \ln y_t & \sum \ln r_t \\ \sum \ln y_t & \sum (\ln y_t)^2 & \sum \ln y_t \cdot \ln r_t \\ \sum \ln r_t & \sum \ln r_t \cdot \ln y_t & \sum (\ln r_t)^2 \end{bmatrix} = \begin{bmatrix} 19,000 & 145,986 & 17,779 \\ 145,986 & 1121,937 & 135,657 \\ 17,779 & 135,657 & 23,963 \end{bmatrix}$$

$$\underset{3x1}{X'y} = \begin{bmatrix} \sum \ln m_t \\ \sum \ln y_t \cdot \ln m_t \\ \sum \ln r_t \cdot \ln m_t \end{bmatrix} = \begin{bmatrix} 121,708 \\ 936,038 \\ 109,843 \end{bmatrix} \cdot \quad \underset{3x3}{(X'X)^{-1}} = \begin{bmatrix} 451,821 & -57,873 & -7,591 \\ -57,873 & 7,416 & 0,956 \\ -7,591 & 0,956 & 0,260 \end{bmatrix}$$

OLS estimation vector:

$$\underset{3x1}{\hat{\beta}} = \begin{bmatrix} \hat{\beta}_1 & = & -15,017 \\ \hat{\beta}_2 & = & 2,811 \\ \hat{\beta}_3 & = & -0,189 \end{bmatrix}$$

Figure: Money demand function



● Implicit restrictions of the least squares estimation

Vector of the residuals: $\hat{u}$

(2.29) $$\hat{\mathbf{u}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$$

Multiplying (2.29) by the transpose of **X**:

(2.30) $$\mathbf{X}'\hat{\mathbf{u}} = \mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} \underset{\substack{\text{wegen} \\ (2.14)}}{=} \mathbf{0}$$

Therefore the following k linear constraints apply to the n OLS residuals:

$$\mathbf{X}'\hat{\mathbf{u}} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{k1} & x_{k2} & \cdots & x_{kn} \end{bmatrix} \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_n \end{bmatrix} = \begin{bmatrix} \sum\limits_{t=1}^{n} \hat{u}_t \\ \sum\limits_{t=1}^{n} x_{2t}\hat{u}_t \\ \vdots \\ \sum\limits_{t=1}^{n} x_{kt}\hat{u}_t \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

1. Restriction:

Sum and mean of the residuals are equal to 0:

$$\sum_{t=1}^{n} \hat{u}_t = 0 \quad \text{and (2.31)} \quad \frac{1}{n}\sum_{t=1}^{n} \hat{u}_t = \bar{\hat{u}} = 0$$

Restrictions 2 to k:

$$\text{(2.32)} \quad \sum_{t=1}^{n} x_{jt}\hat{u}_t = 0 \quad \text{für } j = 2,\ldots k$$

They imply that the covariance between the OLS residuals and the exogenous variables are equal to 0:

$$s_{\hat{u},x} = \frac{1}{n}\sum_{t=1}^{n}(\hat{u}_t - \bar{\hat{u}})(x_{jt} - \bar{x}_j) = \frac{1}{n}\sum_{t=1}^{n}\hat{u}_t(x_{jt} - \bar{x}_j) = \frac{1}{n}\sum_{t=1}^{n}\hat{u}_t x_{jt} - \bar{x}_j \frac{1}{n}\sum_{t=1}^{n}\hat{u}_t$$

With $\bar{x}_j = \frac{1}{n}\sum_{t=1}^{n} x_{jt}$

it follows by (2.31) and (2.32) that

$$s_{\hat{u},x} = \frac{1}{n}\sum_{t=1}^{n} \hat{u}_t \cdot x_{jt} = 0$$