Lecture 11: Lineare (Regressions-) Modelle

11.1 Flussdiagramm

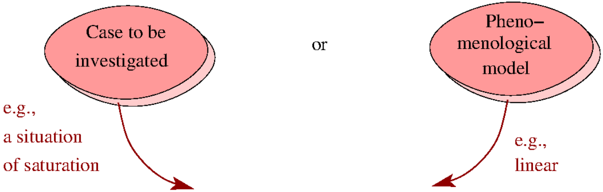
11.2 Modellspezifikation

- 11.2.1 Funktionale Spezifikation
- 11.2.2 Statistische Spezifikation
- 11.2.3 Datenspezifikation

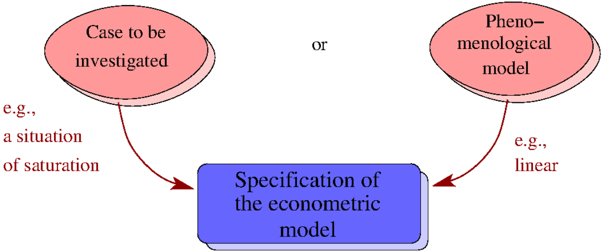
11.3 Ordinary Least Squares (OLS) Schätzung

11.1 Flussdiagramm

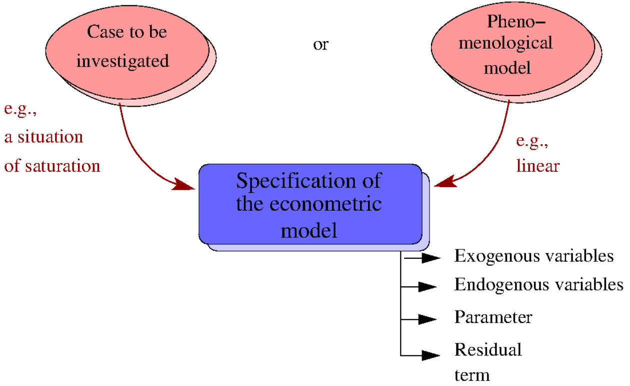
11.1 Flussdiagramm



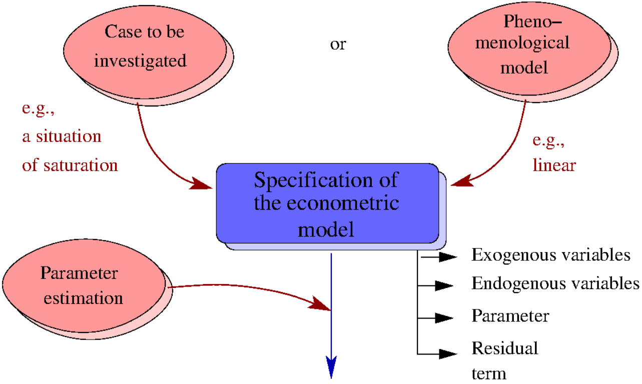
11.1 Flussdiagramm



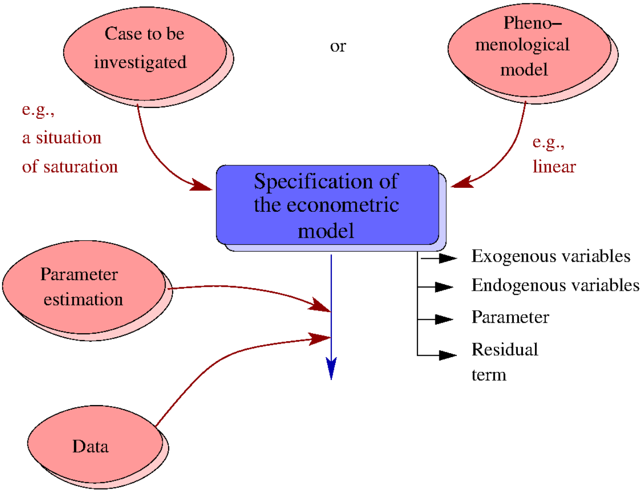
11.1 Flussdiagramm



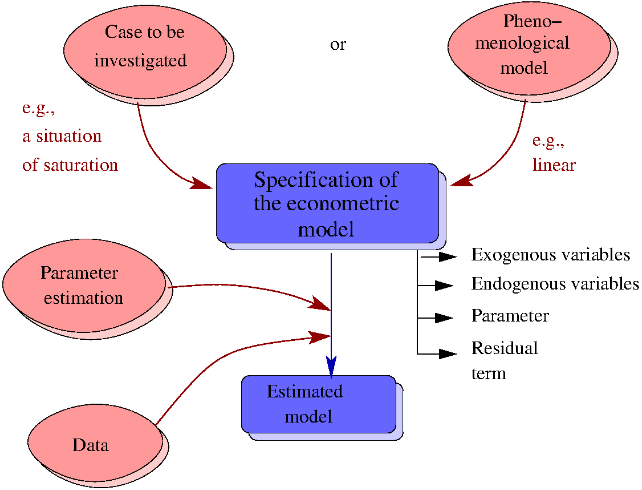
11.1 Flussdiagramm



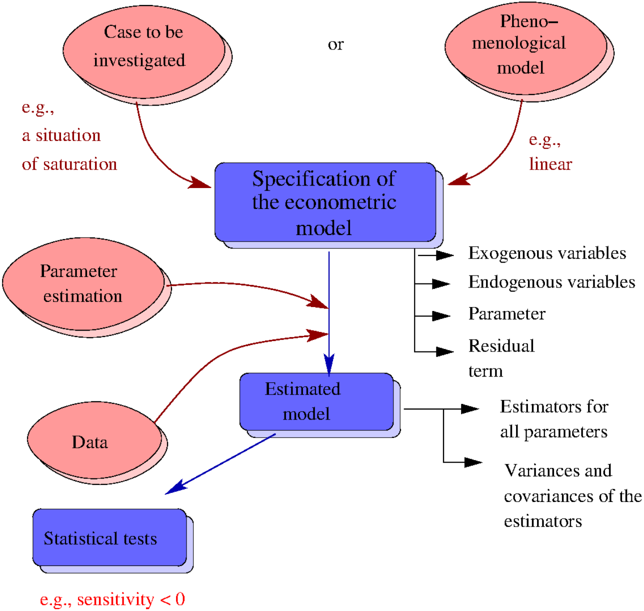
11.1 Flussdiagramm



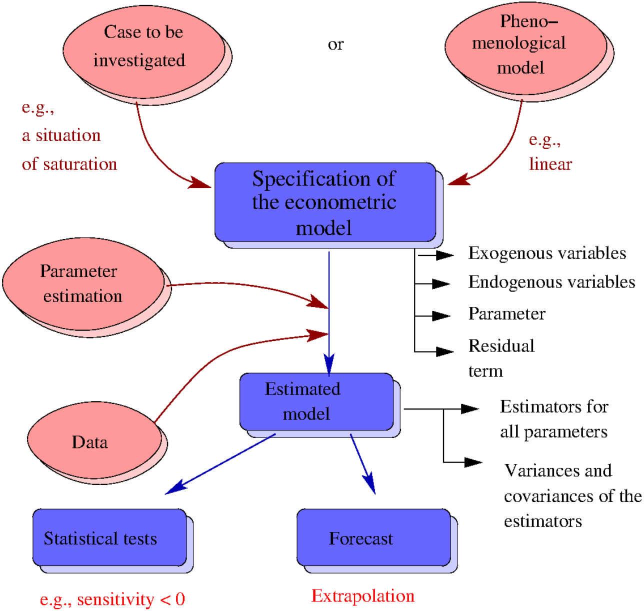
11.1 Flussdiagramm



11.1 Flussdiagramm



11.1 Flussdiagramm



11.2 Modellspezifikation

Modellspezifikation: *vollständige strukturelle Spezifikation* des Modells und seine Konsistenz mit den verfügbaren Daten. Es gibt drei Ebenen:

- **Funktionale Spezifikation**: Exogene und endogene Variable und die funktionale Form, wie sie im Modell auftauchen. Inspesondere die Transformation der exogenen Variablen $\tilde{x}$ in lineare **Faktoren** $x_j = g_j(\tilde{x})$ durch feste, i.A. nichtlineare Funktionen $g_j(.)$
- **Statistische Spezifikation**: Wie sind z.B. *Fehlerterme* verteilt und miteinander korreliert?
- **Datenspezifikation**: Passt das Modell zu den Daten? Gibt es eine ausreichende Zahl an Datensätzen? Sind alle exogenen *und* endogenen Variablen vorhanden?

11.2 Modellspezifikation

Modellspezifikation: *vollständige strukturelle Spezifikation* des Modells und seine Konsistenz mit den verfügbaren Daten. Es gibt drei Ebenen:

- **Funktionale Spezifikation**: Exogene und endogene Variable und die funktionale Form, wie sie im Modell auftauchen. Inspesondere die Transformation der exogenen Variablen $\tilde{\boldsymbol{x}}$ in lineare **Faktoren** $x_j = g_j(\tilde{\boldsymbol{x}})$ durch feste, i.A. nichtlineare Funktionen $g_j(.)$
- **Statistische Spezifikation**: Wie sind z.B. *Fehlerterme* verteilt und miteinander korreliert?
- **Datenspezifikation**: Passt das Modell zu den Daten? Gibt es eine ausreichende Zahl an Datensätzen? Sind alle exogenen *und* endogenen Variablen vorhanden?

11.2 Modellspezifikation

Modellspezifikation: *vollständige strukturelle Spezifikation* des Modells und seine Konsistenz mit den verfügbaren Daten. Es gibt drei Ebenen:

- **Funktionale Spezifikation**: Exogene und endogene Variable und die funktionale Form, wie sie im Modell auftauchen. Inspesondere die Transformation der exogenen Variablen $\tilde{\boldsymbol{x}}$ in lineare **Faktoren** $x_j = g_j(\tilde{\boldsymbol{x}})$ durch feste, i.A. nichtlineare Funktionen $g_j(.)$
- **Statistische Spezifikation**: Wie sind z.B. *Fehlerterme* verteilt und miteinander korreliert?
- **Datenspezifikation**: Passt das Modell zu den Daten? Gibt es eine ausreichende Zahl an Datensätzen? Sind alle exogenen *und* endogenen Variablen vorhanden?

11.2 Modellspezifikation

Modellspezifikation: *vollständige strukturelle Spezifikation* des Modells und seine Konsistenz mit den verfügbaren Daten. Es gibt drei Ebenen:

- **Funktionale Spezifikation**: Exogene und endogene Variable und die funktionale Form, wie sie im Modell auftauchen. Inspesondere die Transformation der exogenen Variablen $\tilde{\boldsymbol{x}}$ in lineare **Faktoren** $x_j = g_j(\tilde{\boldsymbol{x}})$ durch feste, i.A. nichtlineare Funktionen $g_j(.)$
- **Statistische Spezifikation**: Wie sind z.B. *Fehlerterme* verteilt und miteinander korreliert?
- **Datenspezifikation**: Passt das Modell zu den Daten? Gibt es eine ausreichende Zahl an Datensätzen? Sind alle exogenen *und* endogenen Variablen vorhanden?

WARNUNG

Fehlspezifizierte Modelle führen zu Fehlern aller Art, von unproblematisch bis desaströs

- irrelevant: Einige Fehlspezifikationen werden automatisch durch die Kalibrierungsprozedur "entdeckt", z.B. in dem sie "Null/Null" Fehler oder singuläre Matrizen produziert
- mild: keine automatische Detektierung durch die Kalibrierungsmethode, aber milde Folgen: Schätzer ist nach wie vor unverzerrt/effizient, aber die induktive Statistik führt zu unkorrekten Ergebnissen
- mittel: Immer noch unverzerrte Schätzer, aber dieser ist nicht mehr effizient und es gibt starke inferentielle Fehler, insbesondere wird eine zu hohe Signifikanz vorgegaukelt
- desaströs: ie Ergebnisse sind in einer unvorhersehbaren Weise verzerrt

Junk in, junk out!

Es gibt Lügen, verdammte Lügen und **Statistik!**

WARNUNG

Fehlspezifizierte Modelle führen zu Fehlern aller Art, von unproblematisch bis desaströs

- irrelevant: Einige Fehlspezifikationen werden automatisch durch die Kalibrierungsprozedur "entdeckt", z.B. in dem sie "Null/Null" Fehler oder singuläre Matrizen produziert
- mild: keine automatische Detektierung durch die Kalibrierungsmethode, aber milde Folgen: Schätzer ist nach wie vor unverzerrt/effizient, aber die induktive Statistik führt zu unkorrekten Ergebnissen
- mittel: Immer noch unverzerrte Schätzer, aber dieser ist nicht mehr effizient und es gibt starke inferentielle Fehler, insbesondere wird eine zu hohe Signifikanz vorgegaukelt
- desaströs: ie Ergebnisse sind in einer unvorhersehbaren Weise verzerrt

Junk in, junk out!

Es gibt Lügen, verdammte Lügen und **Statistik!**

WARNUNG

Fehlspezifizierte Modelle führen zu Fehlern aller Art, von unproblematisch bis desaströs

- irrelevant: Einige Fehlspezifikationen werden automatisch durch die Kalibrierungsprozedur "entdeckt", z.B. in dem sie "Null/Null" Fehler oder singuläre Matrizen produziert
- mild: keine automatische Detektierung durch die Kalibrierungsmethode, aber milde Folgen: Schätzer ist nach wie vor unverzerrt/effizient, aber die induktive Statistik führt zu unkorrekten Ergebnissen
- mittel: Immer noch unverzerrte Schätzer, aber dieser ist nicht mehr effizient und es gibt starke inferentielle Fehler, insbesondere wird eine zu hohe Signifikanz vorgegaukelt
- desaströs: ie Ergebnisse sind in einer unvorhersehbaren Weise verzerrt

Junk in, junk out!

Es gibt Lügen, verdammte Lügen und **Statistik!**

WARNUNG

Fehlspezifizierte Modelle führen zu Fehlern aller Art, von unproblematisch bis desaströs

- irrelevant: Einige Fehlspezifikationen werden automatisch durch die Kalibrierungsprozedur "entdeckt", z.B. in dem sie "Null/Null" Fehler oder singuläre Matrizen produziert
- mild: keine automatische Detektierung durch die Kalibrierungsmethode, aber milde Folgen: Schätzer ist nach wie vor unverzerrt/effizient, aber die induktive Statistik führt zu unkorrekten Ergebnissen
- mittel: Immer noch unverzerrte Schätzer, aber dieser ist nicht mehr effizient und es gibt starke inferentielle Fehler, insbesondere wird eine zu hohe Signifikanz vorgegaukelt
- desaströs: ie Ergebnisse sind in einer unvorhersehbaren Weise verzerrt

Junk in, junk out!

Es gibt Lügen, verdammte Lügen und **Statistik!**

WARNUNG

Fehlspezifizierte Modelle führen zu Fehlern aller Art, von unproblematisch bis desaströs

- irrelevant: Einige Fehlspezifikationen werden automatisch durch die Kalibrierungsprozedur "entdeckt", z.B. in dem sie "Null/Null" Fehler oder singuläre Matrizen produziert
- mild: keine automatische Detektierung durch die Kalibrierungsmethode, aber milde Folgen: Schätzer ist nach wie vor unverzerrt/effizient, aber die induktive Statistik führt zu unkorrekten Ergebnissen
- mittel: Immer noch unverzerrte Schätzer, aber dieser ist nicht mehr effizient und es gibt starke inferentielle Fehler, insbesondere wird eine zu hohe Signifikanz vorgegaukelt
- desaströs: ie Ergebnisse sind in einer unvorhersehbaren Weise verzerrt

Junk in, junk out!

Es gibt Lügen, verdammte Lügen und **Statistik!**

WARNUNG

Fehlspezifizierte Modelle führen zu Fehlern aller Art, von unproblematisch bis desaströs

- irrelevant: Einige Fehlspezifikationen werden automatisch durch die Kalibrierungsprozedur "entdeckt", z.B. in dem sie "Null/Null" Fehler oder singuläre Matrizen produziert
- mild: keine automatische Detektierung durch die Kalibrierungsmethode, aber milde Folgen: Schätzer ist nach wie vor unverzerrt/effizient, aber die induktive Statistik führt zu unkorrekten Ergebnissen
- mittel: Immer noch unverzerrte Schätzer, aber dieser ist nicht mehr effizient und es gibt starke inferentielle Fehler, insbesondere wird eine zu hohe Signifikanz vorgegaukelt
- desaströs: ie Ergebnisse sind in einer unvorhersehbaren Weise verzerrt

Junk in, junk out!

Es gibt Lügen, verdammte Lügen und **Statistik!**

WARNUNG

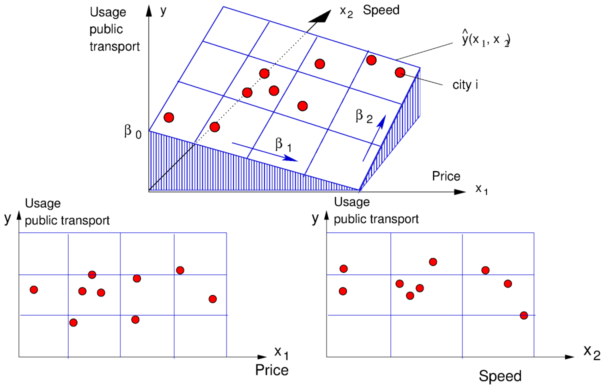
Fehlspezifizierte Modelle führen zu Fehlern aller Art, von unproblematisch bis desaströs

- irrelevant: Einige Fehlspezifikationen werden automatisch durch die Kalibrierungsprozedur "entdeckt", z.B. in dem sie "Null/Null" Fehler oder singuläre Matrizen produziert
- mild: keine automatische Detektierung durch die Kalibrierungsmethode, aber milde Folgen: Schätzer ist nach wie vor unverzerrt/effizient, aber die induktive Statistik führt zu unkorrekten Ergebnissen
- mittel: Immer noch unverzerrte Schätzer, aber dieser ist nicht mehr effizient und es gibt starke inferentielle Fehler, insbesondere wird eine zu hohe Signifikanz vorgegaukelt
- desaströs: ie Ergebnisse sind in einer unvorhersehbaren Weise verzerrt

Junk in, junk out!

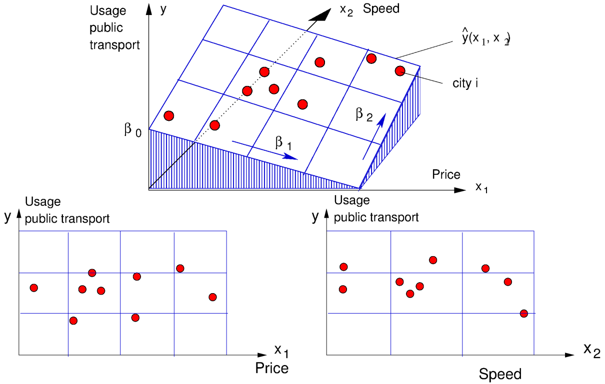
Es gibt Lügen, verdammte Lügen und **Statistik!**

11.2.1 Funktionale Spezifikation 1: relevante Faktoren



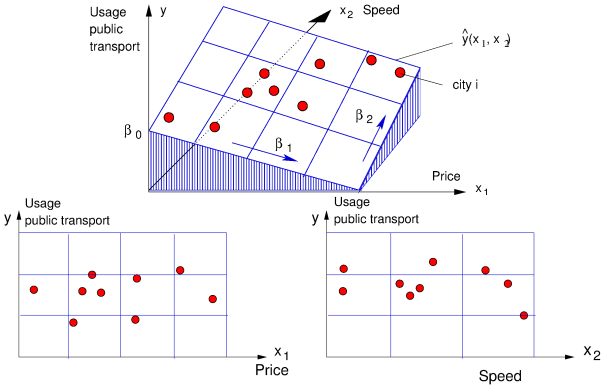
- Berücksichtige alle relevanten Faktoren (oben), lasse keinen aus! (unten)
- **Folgen fehlender Faktoren:** i.A. Verzerrung: **"junk in, junk out"**
- **Folgen überflüssiger Faktoren:** **keine Verzerrung, aber unnötig hohe Schätzfehler**
- **Lösung:** Tests, z.B. **F-test**: *ökonometrische Expertise nötig!*

11.2.1 Funktionale Spezifikation 1: relevante Faktoren



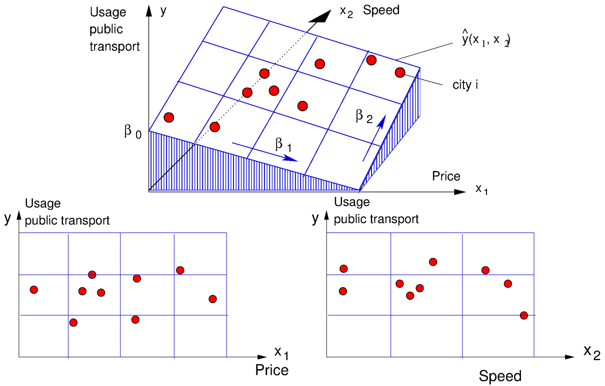
- Berücksichtige alle relevanten Faktoren (oben), lasse keinen aus! (unten)
- **Folgen fehlender Faktoren:** i.A. Verzerrung: **"junk in, junk out"**
- **Folgen überflüssiger Faktoren:** **keine Verzerrung, aber unnötig hohe Schätzfehler**
- **Lösung:** Tests, z.B. **F-test**: *ökonometrische Expertise nötig!*

11.2.1 Funktionale Spezifikation 1: relevante Faktoren



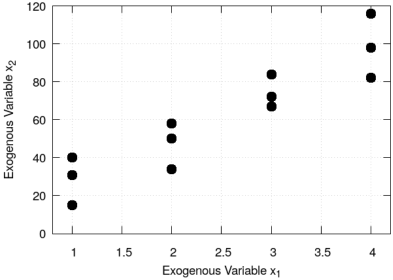
- Berücksichtige alle relevanten Faktoren (oben), lasse keinen aus! (unten)
- **Folgen fehlender Faktoren:** i.A. Verzerrung: **"junk in, junk out"**
- **Folgen überflüssiger Faktoren:** **keine Verzerrung, aber unnötig hohe Schätzfehler**
- **Lösung:** Tests, z.B. **F-test**: *ökonometrische Expertise nötig!*

11.2.1 Funktionale Spezifikation 1: relevante Faktoren

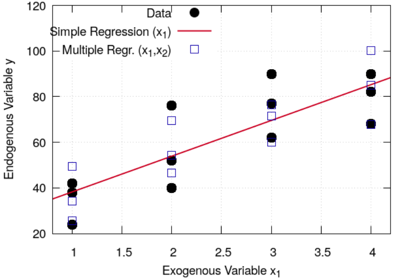


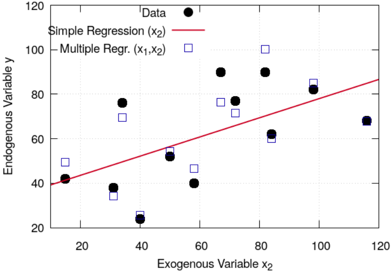
- Berücksichtige alle relevanten Faktoren (oben), lasse keinen aus! (unten)
- **Folgen fehlender Faktoren:** i.A. Verzerrung: **"junk in, junk out"**
- **Folgen überflüssiger Faktoren:** **keine Verzerrung, aber unnötig hohe Schätzfehler**
- **Lösung:** Tests, z.B. **F-test**: *ökonometrische Expertise nötig!*

Beispiel: Auslastung von Hotels

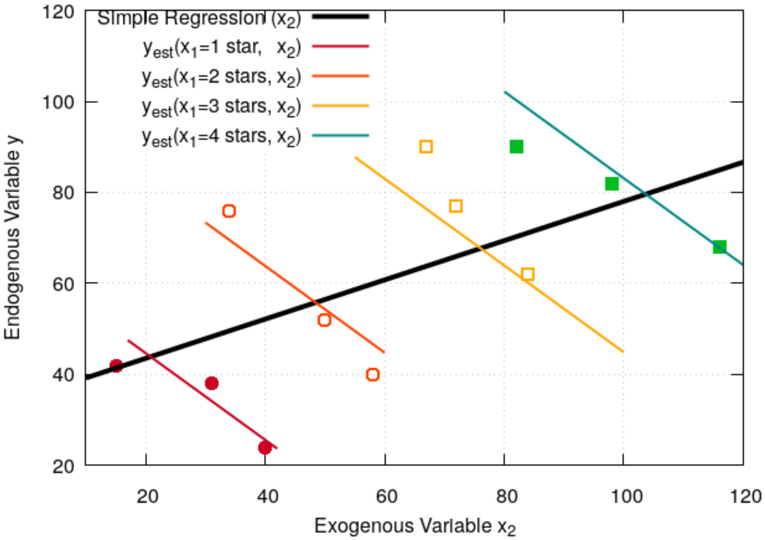


- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$ mit den Faktoren $x_0 = 1$, x_1: Proxy für Qualität [# stars]; x_2: Preis [€/Nacht].
- Die exogenen Faktoren sind (nicht vollständig) korreliert:
- Endogene Variable: Auslastung [%]
- Nachfrage ist positiv mit Qualität und Preis (!) korreliert

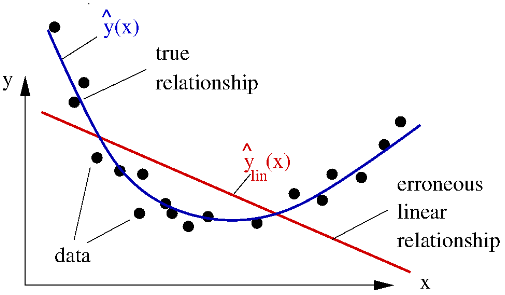




Effekt der Korrelationen zwischen den exogenen Variablen

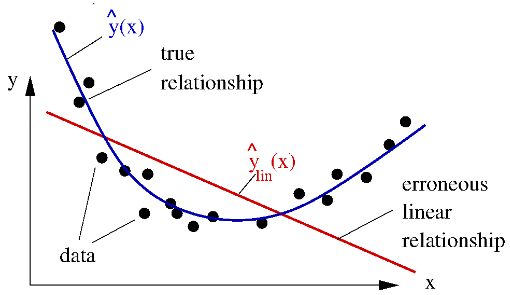


Funktionale Spezifikation 2: Linearität



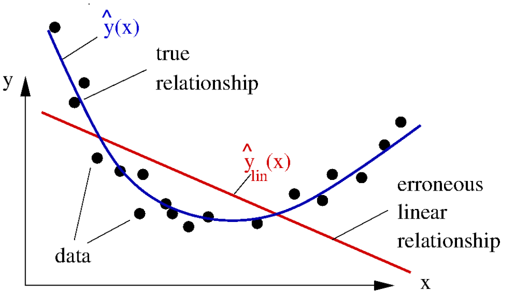
- Das Modell sollte linear sein (hier nicht erfüllt)
- **Konsequenz:** **"junk in, junk out"**
- **Lösung:** Nichtlineare Transformation der exogenen Variablen in **Faktoren**, bezüglich derer das Modell linear ist, hier z.B. $x'_0 = 1, x'_1 = 1/x, x'_2 = x^2$ or $x'_0 = 1, x'_1 = x, x'_2 = x^2$.

Funktionale Spezifikation 2: Linearität



- Das Modell sollte linear sein (hier nicht erfüllt)
- **Konsequenz: "junk in, junk out"**
- **Lösung:** Nichtlineare Transformation der exogenen Variablen in **Faktoren**, bezüglich derer das Modell linear ist, hier z.B. $x'_0 = 1, x'_1 = 1/x, x'_2 = x^2$ or $x'_0 = 1, x'_1 = x, x'_2 = x^2$.

Funktionale Spezifikation 2: Linearität



- Das Modell sollte linear sein (hier nicht erfüllt)
- **Konsequenz: "junk in, junk out"**
- **Lösung:** Nichtlineare Transformation der exogenen Variablen in **Faktoren**, bezüglich derer das Modell linear ist, hier z.B. $x'_0 = 1, x'_1 = 1/x, x'_2 = x^2$ or $x'_0 = 1, x'_1 = x, x'_2 = x^2$.

Beispiel: Kraftstoffverbrauch

Für einen konstanten Wirkungsgrad chemische $\rightarrow$ mechanische Energie ist der für die Fortbewegung benötigte Verbrauch y (pro 100 km) proportional zum *Fahrwiderstand* mit den additiven Anteilen

- Rollreibung: Unabhängig von der Geschwindigkeit $\tilde{x}_1$, proportional zur Masse $\tilde{x}_2$.
- Luftwiderstand: proportional zur quadrierten Geschwindigkeit $\tilde{x}_1^2$, unabhängig von der Masse
- Steigung: proportional zum Produkt aus Masse und Steigung $\tilde{x}_3$

Zusätzlich gibt es eine Basisverbrauchsrate ($\approx 0.5\,\mathrm{Liter/h}$) für Leerlauf, Heizung, Lüftung, Licht etc $\Rightarrow$ Anteil proportional zu 1/Geschwindigkeit $\Rightarrow$ Model

$$y(x) = \sum_{j=1}^{4} \beta_j x_j + \epsilon, \quad x_1 = \tilde{x}_2, \quad x_2 = \tilde{x}_1^2, \quad x_3 = \tilde{x}_2\tilde{x}_3, \quad x_4 = \frac{1}{\tilde{x}_1}$$

Beispiel: Kraftstoffverbrauch

Für einen konstanten Wirkungsgrad chemische $\rightarrow$ mechanische Energie ist der für die Fortbewegung benötigte Verbrauch y (pro 100 km) proportional zum *Fahrwiderstand* mit den additiven Anteilen

- Rollreibung: Unabhängig von der Geschwindigkeit $\tilde{x}_1$, proportional zur Masse $\tilde{x}_2$.
- Luftwiderstand: proportional zur quadrierten Geschwindigkeit $\tilde{x}_1^2$, unabhängig von der Masse
- Steigung: proportional zum Produkt aus Masse und Steigung $\tilde{x}_3$

Zusätzlich gibt es eine Basisverbrauchsrate ($\approx 0.5\,\mathrm{Liter/h}$) für Leerlauf, Heizung, Lüftung, Licht etc $\Rightarrow$ Anteil proportional zu 1/Geschwindigkeit $\Rightarrow$ Model

$$y(\boldsymbol{x}) = \sum_{j=1}^{4} \beta_j x_j + \epsilon, \quad x_1 = \tilde{x}_2, \quad x_2 = \tilde{x}_1^2, \quad x_3 = \tilde{x}_2\tilde{x}_3, \quad x_4 = \frac{1}{\tilde{x}_1}$$

Beispiel: Kraftstoffverbrauch

Für einen konstanten Wirkungsgrad chemische → mechanische Energie ist der für die Fortbewegung benötigte Verbrauch y (pro 100 km) proportional zum *Fahrwiderstand* mit den additiven Anteilen

- Rollreibung: Unabhängig von der Geschwindigkeit $\tilde{x}_1$, proportional zur Masse $\tilde{x}_2$.
- Luftwiderstand: proportional zur quadrierten Geschwindigkeit $\tilde{x}_1^2$, unabhängig von der Masse
- Steigung: proportional zum Produkt aus Masse und Steigung $\tilde{x}_3$

Zusätzlich gibt es eine Basisverbrauchsrate ($\approx 0.5\,\mathrm{Liter/h}$) für Leerlauf, Heizung, Lüftung, Licht etc ⇒ Anteil proportional zu 1/Geschwindigkeit ⇒ Model

$$y(\boldsymbol{x}) = \sum_{j=1}^{4} \beta_j x_j + \epsilon, \quad x_1 = \tilde{x}_2, \quad x_2 = \tilde{x}_1^2, \quad x_3 = \tilde{x}_2 \tilde{x}_3, \quad x_4 = \frac{1}{\tilde{x}_1}$$

Beispiel: Kraftstoffverbrauch

Für einen konstanten Wirkungsgrad chemische $\rightarrow$ mechanische Energie ist der für die Fortbewegung benötigte Verbrauch y (pro 100 km) proportional zum *Fahrwiderstand* mit den additiven Anteilen

- Rollreibung: Unabhängig von der Geschwindigkeit $\tilde{x}_1$, proportional zur Masse $\tilde{x}_2$.
- Luftwiderstand: proportional zur quadrierten Geschwindigkeit $\tilde{x}_1^2$, unabhängig von der Masse
- Steigung: proportional zum Produkt aus Masse und Steigung $\tilde{x}_3$

Zusätzlich gibt es eine Basisverbrauchsrate ($\approx 0.5\,\mathrm{Liter/h}$) für Leerlauf, Heizung, Lüftung, Licht etc $\Rightarrow$ Anteil proportional zu 1/Geschwindigkeit $\Rightarrow$ Model

$$y(\boldsymbol{x}) = \sum_{j=1}^{4} \beta_j x_j + \epsilon, \quad x_1 = \tilde{x}_2, \quad x_2 = \tilde{x}_1^2, \quad x_3 = \tilde{x}_2\tilde{x}_3, \quad x_4 = \frac{1}{\tilde{x}_1}$$

Beispiel: Kraftstoffverbrauch

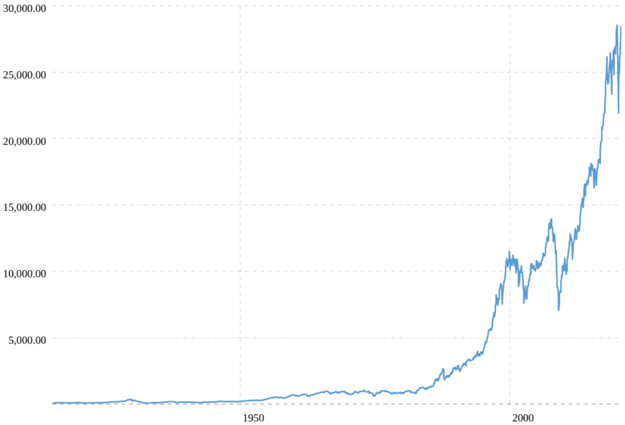
Für einen konstanten Wirkungsgrad chemische $\rightarrow$ mechanische Energie ist der für die Fortbewegung benötigte Verbrauch y (pro 100 km) proportional zum *Fahrwiderstand* mit den additiven Anteilen

- Rollreibung: Unabhängig von der Geschwindigkeit $\tilde{x}_1$, proportional zur Masse $\tilde{x}_2$.
- Luftwiderstand: proportional zur quadrierten Geschwindigkeit $\tilde{x}_1^2$, unabhängig von der Masse
- Steigung: proportional zum Produkt aus Masse und Steigung $\tilde{x}_3$

Zusätzlich gibt es eine Basisverbrauchsrate ($\approx 0.5\,\mathrm{Liter/h}$) für Leerlauf, Heizung, Lüftung, Licht etc $\Rightarrow$ Anteil proportional zu 1/Geschwindigkeit $\Rightarrow$ Model

$$y(\boldsymbol{x}) = \sum_{j=1}^{4} \beta_j x_j + \epsilon, \quad x_1 = \tilde{x}_2, \quad x_2 = \tilde{x}_1^2, \quad x_3 = \tilde{x}_2 \tilde{x}_3, \quad x_4 = \frac{1}{\tilde{x}_1}$$

Transformation der endogenen Variable I

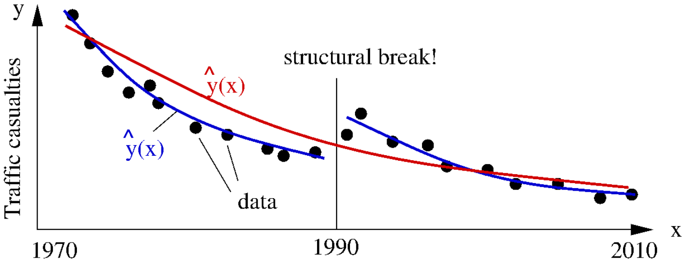


Eine Zeittransformation $\tilde{x} \to x = \exp(\tilde{x})$ würde zwar linearisieren, aber die Fluktuationen (Residualterme) sind nicht mehr i.i.d ($\Rightarrow$ statistische Spezifikation)

Transformation der endogenen Variable II

Transformation der *endogenen* Variable $y \to u = \ln(y)$, $x = \tilde{x}$ gibt hingegen ein korrekt spezifiziertes Modell $u(x) = \beta_0 + \beta_1 x + \epsilon$, $\epsilon \sim$ i.i.d.

Funktionale Spezifikation 3: Homogenität der Grundgesamtheit

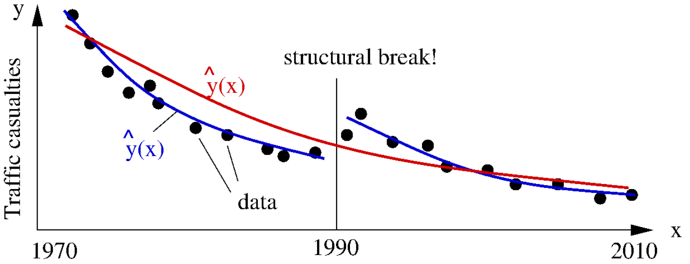


- **Folgen einer Verletzung:** ein nichtberücksichtigter **Strukturbruch** im Raum der exogenen Variablen führt zu einer *Verzerrung*: **junk in, junk out**
- **Lösung:** *Dummyvariable* mit Werten 0 und 1 vor/nach dem Strukturbruch

? Möglicher Strukturbruch in obigem Plot?

! 1. neue Datenbasis (DDR+Westdeutschland $\rightarrow$ Deutschland); 2. Umdefinitionen, z.B. "ernsthaft verletzt" bedeutete vor dem Strukturbruch stationärer, danach ambulanter+stationärer Krankenhausaufenthalt

Funktionale Spezifikation 3: Homogenität der Grundgesamtheit

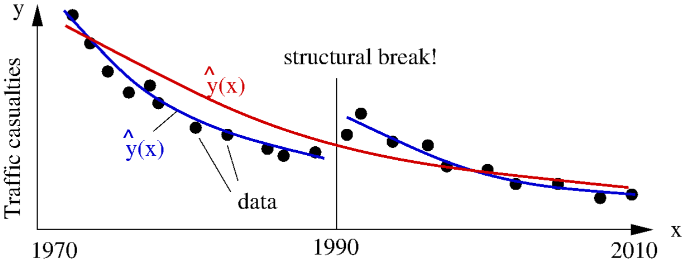


- **Folgen einer Verletzung:** ein nichtberücksichtigter **Strukturbruch** im Raum der exogenen Variablen führt zu einer *Verzerrung*: **junk in, junk out**
- **Lösung:** *Dummyvariable* mit Werten 0 und 1 vor/nach dem Strukturbruch

? Möglicher Strukturbruch in obigem Plot?

! 1. neue Datenbasis (DDR+Westdeutschland → Deutschland); 2. Umdefinitionen, z.B. "ernsthaft verletzt" bedeutete vor dem Strukturbruch stationärer, danach ambulanter+stationärer Krankenhausaufenthalt

Funktionale Spezifikation 3: Homogenität der Grundgesamtheit

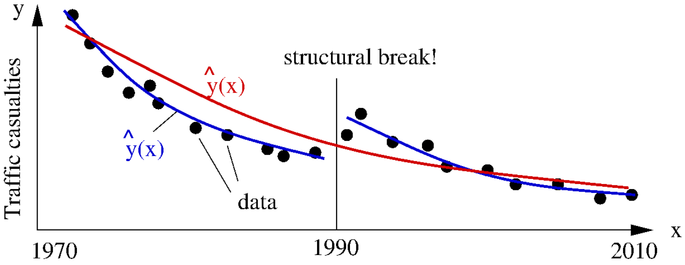


- **Folgen einer Verletzung:** ein nichtberücksichtigter **Strukturbruch** im Raum der exogenen Variablen führt zu einer *Verzerrung*: **junk in, junk out**
- **Lösung:** *Dummyvariable* mit Werten 0 und 1 vor/nach dem Strukturbruch

? Möglicher Strukturbruch in obigem Plot?

! 1. neue Datenbasis (DDR+Westdeutschland → Deutschland); 2. Umdefinitionen, z.B. "ernsthaft verletzt" bedeutete vor dem Strukturbruch stationärer, danach ambulanter+stationärer Krankenhausaufenthalt

Funktionale Spezifikation 3: Homogenität der Grundgesamtheit

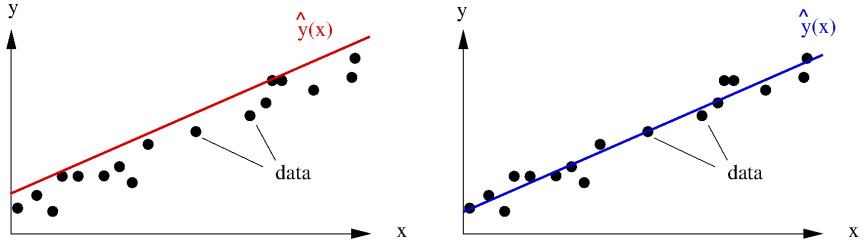


- **Folgen einer Verletzung:** ein nichtberücksichtigter **Strukturbruch** im Raum der exogenen Variablen führt zu einer *Verzerrung*: **junk in, junk out**
- **Lösung:** *Dummyvariable* mit Werten 0 und 1 vor/nach dem Strukturbruch

? Möglicher Strukturbruch in obigem Plot?

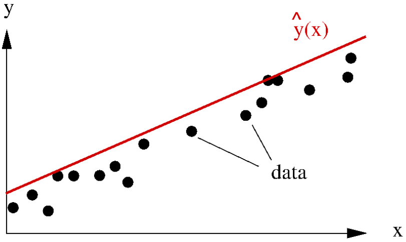
! 1. neue Datenbasis (DDR+Westdeutschland → Deutschland); 2. Umdefinitionen, z.B. "ernsthaft verletzt" bedeutete vor dem Strukturbruch stationärer, danach ambulanter+stationärer Krankenhausaufenthalt

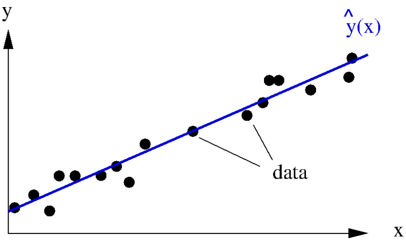
11.2.2 Statistical Spezifikation
1. Residuum ϵ hat Erwartungswert null



- **Bedingung:** $E(\epsilon) = 0$.
- **Konsequenz:** **keine**: Die OLS- (KQ-) Kalibrierung/Schätzmethode berücksichtigt dies automatisch. Bei der diskreten Wahltheorie ist es sogar überhaupt nicht relevant (**Warum?**)

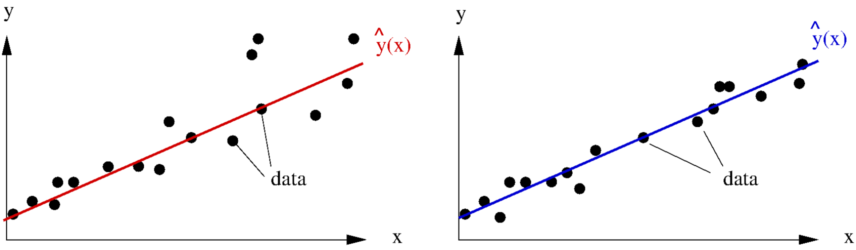
11.2.2 Statistical Spezifikation
1. Residuum ϵ hat Erwartungswert null





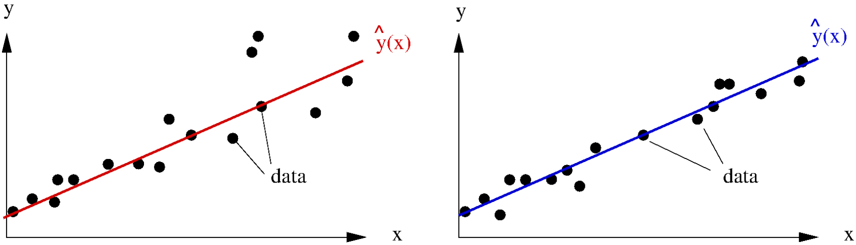
- **Bedingung:** $E(\epsilon) = 0.$
- **Konsequenz:** **keine**: Die OLS- (KQ-) Kalibrierung/Schätzmethode berücksichtigt dies automatisch. Bei der diskreten Wahltheorie ist es sogar überhaupt nicht relevant (**Warum?**)

Statistische Spezifikation 2: Homoskedastizität



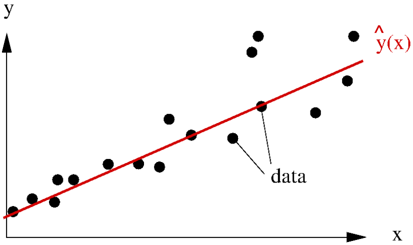
- **Bedingung:** ϵ soll konstante Varianz haben (Homoskedastizität, rechts), keine variable (Heteroskedastizität, links)
- **Konsequenz:** bei Verletzung bleibt der KQ-Schätzer **unverzerrt, ist aber nicht mehr effizient** (ein relativ "milder" Fehler).
- **Lösung:** Fortgeschrittene Methoden wie gewichteter KQ-Schätzer. Ggf Transformation der anhängigen Variable ($\rightarrow$ Dow-Jones-Beispiel

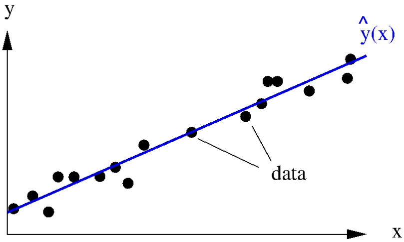
Statistische Spezifikation 2: Homoskedastizität



- **Bedingung:** ϵ soll konstante Varianz haben (Homoskedastizität, rechts), keine variable (Heteroskedastizität, links)
- **Konsequenz:** bei Verletzung bleibt der KQ-Schätzer **unverzerrt, ist aber nicht mehr effizient** (ein relativ "milder" Fehler).
- **Lösung:** Fortgeschrittene Methoden wie gewichteter KQ-Schätzer. Ggf Transformation der anhängigen Variable ($\rightarrow$ Dow-Jones-Beispiel

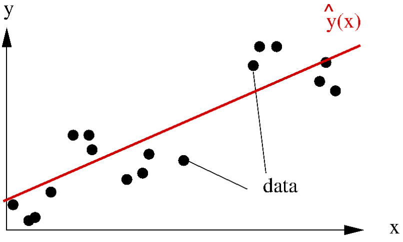
Statistische Spezifikation 2: Homoskedastizität

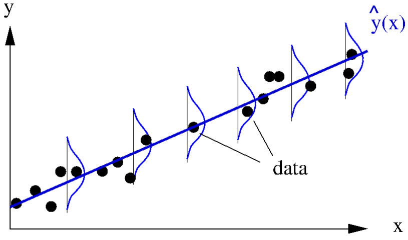




- **Bedingung:** ϵ soll konstante Varianz haben (Homoskedastizität, rechts), keine variable (Heteroskedastizität, links)
- **Konsequenz:** bei Verletzung bleibt der KQ-Schätzer **unverzerrt, ist aber nicht mehr effizient** (ein relativ "milder" Fehler).
- **Lösung:** Fortgeschrittene Methoden wie gewichteter KQ-Schätzer. Ggf Transformation der anhängigen Variable ($\rightarrow$ Dow-Jones-Beispiel

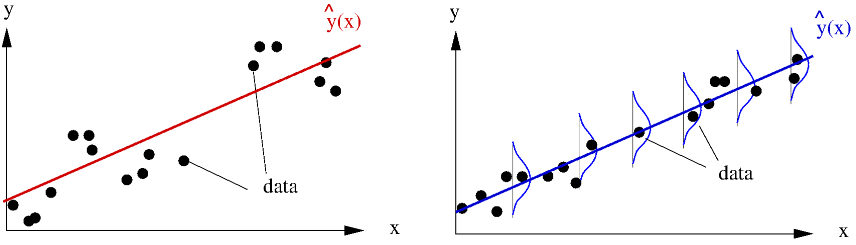
Statistische Spezifikation 3: Korrelationsfreiheit der Residuen





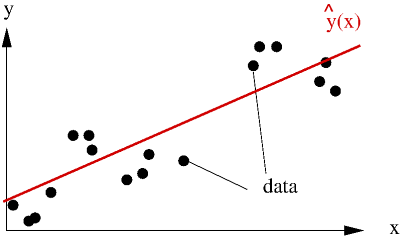
- **Bedingung:** ϵ ist nicht bezüglich der x_i und/oder y korreliert (rechts). Die linke Modell-Daten-Kombination ist fehlspezifiziert
- **Konsequenz:** relativ **mild:** (KQ-Schätzer nicht efficient; Unterschätzung der Schätzfehler, evtl unbedeutende Verzerrung).
- **Lösung:** Identifiziere aus dem Sachverhalt einen nichtberücksichtigten systematischen Einfluss, z.B. Periodizität

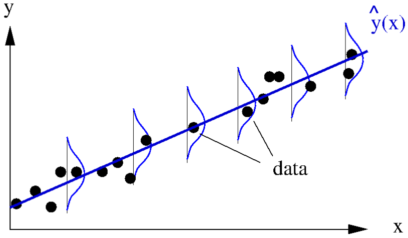
Statistische Spezifikation 3: Korrelationsfreiheit der Residuen



- **Bedingung:** ϵ ist nicht bezüglich der x_i und/oder y korreliert (rechts). Die linke Modell-Daten-Kombination ist fehlspezifiziert
- **Konsequenz:** relativ **mild:** (KQ-Schätzer nicht efficient; Unterschätzung der Schätzfehler, evtl unbedeutende Verzerrung).
- **Lösung:** Identifiziere aus dem Sachverhalt einen nichtberücksichtigten systematischen Einfluss, z.B. Periodizität

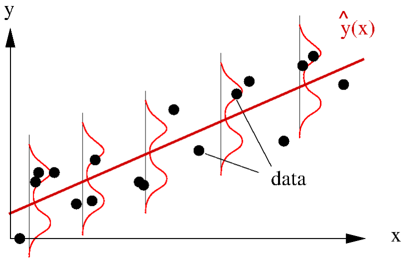
Statistische Spezifikation 3: Korrelationsfreiheit der Residuen

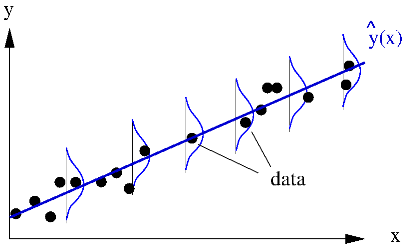




- **Bedingung:** ϵ ist nicht bezüglich der x_i und/oder y korreliert (rechts). Die linke Modell-Daten-Kombination ist fehlspezifiziert
- **Konsequenz:** relativ **mild:** (KQ-Schätzer nicht efficient; Unterschätzung der Schätzfehler, evtl unbedeutende Verzerrung).
- **Lösung:** Identifiziere aus dem Sachverhalt einen nichtberücksichtigten systematischen Einfluss, z.B. Periodizität

Statistische Spezifikation 4: Gaußverteilte Residuen

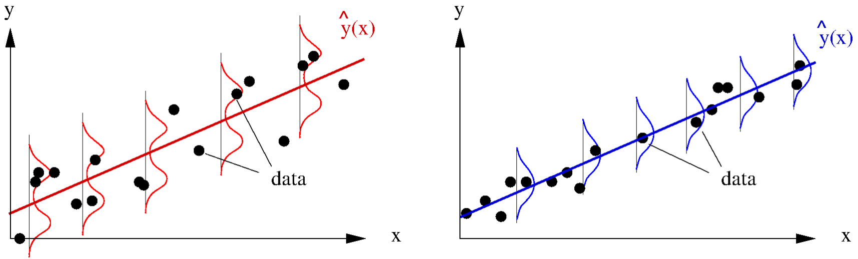




- **Bedingung:** $\epsilon \sim N(0, \sigma^2)$ (rechts), nicht anders verteilt (links).
- **Konsequenz:** eine Verletzung hat **sehr milde** Konsequenzen. Der KQ-Schätzer bleibt unverzerrt *und* effizient, nur die üblichen Fehlerabschätzungen und Tests sind falsch, vor allem bei kleinen Fehlerwahrscheinlichkeiten α
- Alle vier statistischen Spezifikationen zusammen:

$\epsilon \sim$ i.i.d.$N(0, \sigma^2)$ i.i.d.: **i**dentical **i**ndependent **d**istributions

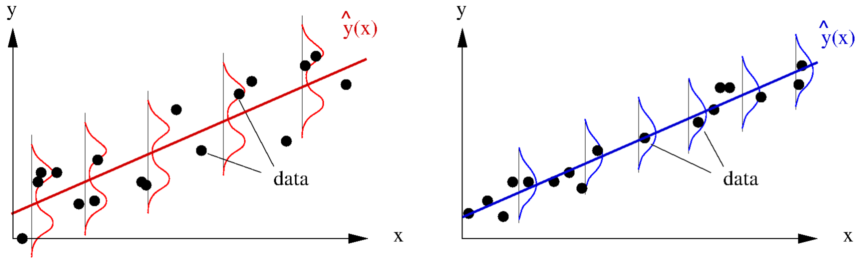
Statistische Spezifikation 4: Gaußverteilte Residuen



- **Bedingung:** $\epsilon \sim N(0, \sigma^2)$ (rechts), nicht anders verteilt (links).
- **Konsequenz:** eine Verletzung hat **sehr milde** Konsequenzen. Der KQ-Schätzer bleibt unverzerrt *und* effizient, nur die üblichen Fehlerabschätzungen und Tests sind falsch, vor allem bei kleinen Fehlerwahrscheinlichkeiten α
- Alle vier statistischen Spezifikationen zusammen:

$\epsilon \sim$ i.i.d.$N(0, \sigma^2)$ i.i.d.: **i**dentical **i**ndependent **d**istributions

Statistische Spezifikation 4: Gaußverteilte Residuen



- **Bedingung:** $\epsilon \sim N(0, \sigma^2)$ (rechts), nicht anders verteilt (links).
- **Konsequenz:** eine Verletzung hat **sehr milde** Konsequenzen. Der KQ-Schätzer bleibt unverzerrt *und* effizient, nur die üblichen Fehlerabschätzungen und Tests sind falsch, vor allem bei kleinen Fehlerwahrscheinlichkeiten α
- Alle vier statistischen Spezifikationen zusammen:

$\epsilon \sim$ i.i.d.$N(0, \sigma^2)$ i.i.d.: **i**dentical **i**ndependent **d**istributions

Datenspezifikation 1: genug Daten

- Jeder Datensatz muss *alle* endogenen und exogenen Variablen enthalten und es muss mehr Datensätze als Parameter geben: $n > J + 1$. Die Daten müssen das Modell also *überbestimmen*

- **Konsequenz einer Verletzung:** Falls $n = J + 1$, kann das Modell *exakt* ohne Fehler an die Daten angepasst werden: $\epsilon_i = 0$ bzw. *overfitting*. Dennoch **harmlos**, da der KQ-Schätzer und andere Tests einen "0/0-Fehler" auswerfen. Bei $n < J + 1$ wird darüberhinaus eine Inversion einer singulären Matrix versucht

- **Konsequenz einer Erfüllung "gerade so":** Falls es nur wenig mehr Datensätze als Parameter gibt, also wenig **Freiheitsgrade** verbleiben, ist der Schätzer unverzerrt und effizient, aber hat **großen Schätzfehler**

- **Lösung:** Mehr Daten ...

Datenspezifikation 1: genug Daten

- Jeder Datensatz muss *alle* endogenen und exogenen Variablen enthalten und es muss mehr Datensätze als Parameter geben: $n > J + 1$. Die Daten müssen das Modell also *überbestimmen*
- **Konsequenz einer Verletzung:** Falls $n = J + 1$, kann das Modell *exakt* ohne Fehler an die Daten angepasst werden: $\epsilon_i = 0$ bzw. *overfitting*. Dennoch **harmlos**, da der KQ-Schätzer und andere Tests einen "0/0-Fehler" auswerfen. Bei $n < J + 1$ wird darüberhinaus eine Inversion einer singulären Matrix versucht
- **Konsequenz einer Erfüllung "gerade so":** Falls es nur wenig mehr Datensätze als Parameter gibt, also wenig **Freiheitsgrade** verbleiben, ist der Schätzer unverzerrt und effizient, aber hat **großen Schätzfehler**
- **Lösung:** Mehr Daten ...

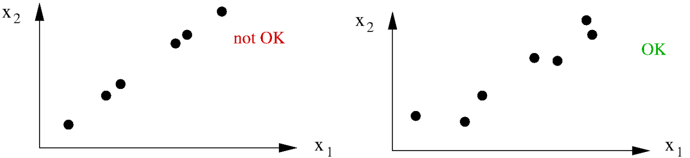
Datenspezifikation 1: genug Daten

- Jeder Datensatz muss *alle* endogenen und exogenen Variablen enthalten und es muss mehr Datensätze als Parameter geben: $n > J+1$. Die Daten müssen das Modell also *überbestimmen*
- **Konsequenz einer Verletzung:** Falls $n = J+1$, kann das Modell *exakt* ohne Fehler an die Daten angepasst werden: $\epsilon_i = 0$ bzw. *overfitting*. Dennoch **harmlos**, da der KQ-Schätzer und andere Tests einen "0/0-Fehler" auswerfen. Bei $n < J+1$ wird darüberhinaus eine Inversion einer singulären Matrix versucht
- **Konsequenz einer Erfüllung "gerade so":** Falls es nur wenig mehr Datensätze als Parameter gibt, also wenig **Freiheitsgrade** verbleiben, ist der Schätzer unverzerrt und effizient, aber hat **großen Schätzfehler**
- **Lösung:** Mehr Daten ...

Datenspezifikation 1: genug Daten

- Jeder Datensatz muss *alle* endogenen und exogenen Variablen enthalten und es muss mehr Datensätze als Parameter geben: $n > J + 1$. Die Daten müssen das Modell also *überbestimmen*
- **Konsequenz einer Verletzung:** Falls $n = J + 1$, kann das Modell *exakt* ohne Fehler an die Daten angepasst werden: $\epsilon_i = 0$ bzw. *overfitting*. Dennoch **harmlos**, da der KQ-Schätzer und andere Tests einen "0/0-Fehler" auswerfen. Bei $n < J + 1$ wird darüberhinaus eine Inversion einer singulären Matrix versucht
- **Konsequenz einer Erfüllung "gerade so":** Falls es nur wenig mehr Datensätze als Parameter gibt, also wenig **Freiheitsgrade** verbleiben, ist der Schätzer unverzerrt und effizient, aber hat **großen Schätzfehler**
- **Lösung:** Mehr Daten ...

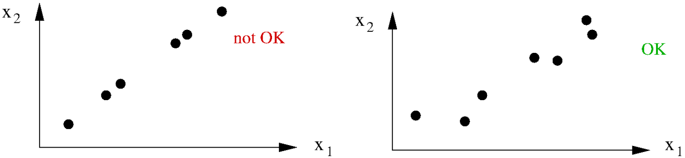
Datenspezifikation 2: Keine Multi-Kolinearität



- Multikolinearität $\to$ Mindestens eine exogene Variable kann *in allen Datensätzen* als Linearkombination anderer Variablen angegeben werden. Die Datenmatrix ist damit *singulär*.
- Nichtperfekte Korrelationen hingegen sind ausdrücklich erlaubt
- Nichtperfekte Korrelationen sind häufig, z.B. Preis vs. Qualität (**Vorzeichen?**)
- **Konsequenz einer verletzung:** der KQ-Schätzer **"entdeckt" für Sie eine Verletzung automatisch**: "0/0"-Fehler. Nahezu perfekte Kolinearität $\Rightarrow$ **große Schätzfehler**

Sind alle Kriterien aus allen drei Spezifikationen erfüllt, sagt man, dass das ökonometrische Problem die **Gauß-Markov Annahmen** erfüllt

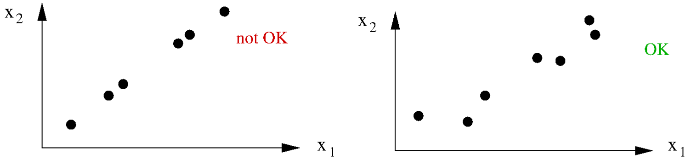
Datenspezifikation 2: Keine Multi-Kolinearität



- Multikolinearität $\to$ Mindestens eine exogene Variable kann *in allen Datensätzen* als Linearkombination anderer Variablen angegeben werden. Die Datenmatrix ist damit *singulär*.
- Nichtperfekte Korrelationen hingegen sind ausdrücklich erlaubt
- Nichtperfekte Korrelationen sind häufig, z.B. Preis vs. Qualität (**Vorzeichen?**)
- **Konsequenz einer verletzung:** der KQ-Schätzer **"entdeckt" für Sie eine Verletzung automatisch**: "0/0"-Fehler. Nahezu perfekte Kolinearität $\Rightarrow$ **große Schätzfehler**

Sind alle Kriterien aus allen drei Spezifikationen erfüllt, sagt man, dass das ökonometrische Problem die **Gauß-Markov Annahmen** erfüllt

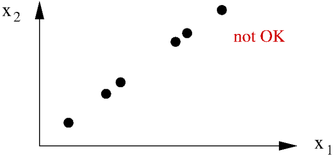
Datenspezifikation 2: Keine Multi-Kolinearität

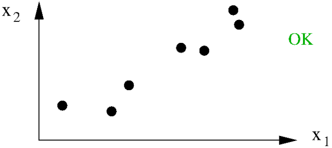


- Multikolinearität → Mindestens eine exogene Variable kann *in allen Datensätzen* als Linearkombination anderer Variablen angegeben werden. Die Datenmatrix ist damit *singulär*.
- Nichtperfekte Korrelationen hingegen sind ausdrücklich erlaubt
- Nichtperfekte Korrelationen sind häufig, z.B. Preis vs. Qualität (**Vorzeichen?**)
- **Konsequenz einer verletzung:** der KQ-Schätzer **"entdeckt" für Sie eine Verletzung automatisch**: "0/0"-Fehler. Nahezu perfekte Kolinearität ⇒ **große Schätzfehler**

Sind alle Kriterien aus allen drei Spezifikationen erfüllt, sagt man, dass das ökonometrische Problem die **Gauß-Markov Annahmen** erfüllt

Datenspezifikation 2: Keine Multi-Kolinearität

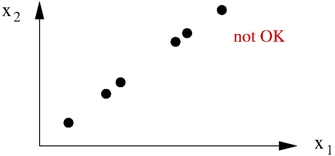


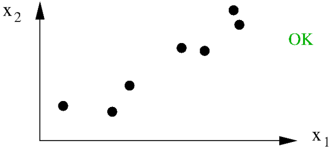


- Multikolinearität → Mindestens eine exogene Variable kann *in allen Datensätzen* als Linearkombination anderer Variablen angegeben werden. Die Datenmatrix ist damit *singulär*.
- Nichtperfekte Korrelationen hingegen sind ausdrücklich erlaubt
- Nichtperfekte Korrelationen sind häufig, z.B. Preis vs. Qualität (**Vorzeichen?**)
- **Konsequenz einer verletzung:** der KQ-Schätzer **"entdeckt" für Sie eine Verletzung automatisch**: "0/0"-Fehler. Nahezu perfekte Kolinearität ⇒ **große Schätzfehler**

Sind alle Kriterien aus allen drei Spezifikationen erfüllt, sagt man, dass das ökonometrische Problem die **Gauß-Markov Annahmen** erfüllt

Datenspezifikation 2: Keine Multi-Kolinearität





- Multikolinearität → Mindestens eine exogene Variable kann *in allen Datensätzen* als Linearkombination anderer Variablen angegeben werden. Die Datenmatrix ist damit *singulär*.
- Nichtperfekte Korrelationen hingegen sind ausdrücklich erlaubt
- Nichtperfekte Korrelationen sind häufig, z.B. Preis vs. Qualität (**Vorzeichen?**)
- **Konsequenz einer verletzung:** der KQ-Schätzer **"entdeckt" für Sie eine Verletzung automatisch**: "0/0"-Fehler. Nahezu perfekte Kolinearität ⇒ **große Schätzfehler**

Sind alle Kriterien aus allen drei Spezifikationen erfüllt, sagt man, dass das ökonometrische Problem die **Gauß-Markov Annahmen** erfüllt

Wie entdeckt man Multi-Kolinearität?

- Gegeben: n Datensätze $\{x_{i0}, ..., x_{ij}, ..., x_{iJ}\}$, $i = 1, ..., n$ (die Datensätze enthalten auch y_i, aber das ist hier nicht relevant)
- x_{ij} ist der j^{th} Faktor im Datensatz i
- Multikollinearität ist gegeben, wenn man einen Faktor x_k als Linearkombination aus allen anderen Faktoren *für alle Datensätze* ausdrücken kann:

$$x_{ik} = \sum_{j \neq k} c_j x_{ij} \quad \forall i = 1, ..., n, \quad \text{konstante Koeffizienten } c_j$$

- Bei zwei Variablen ist dies leicht: $x_2 = c_0 x_1$ bzw Korrelationskoeffizient ist =1 oer $= -1$. Bei mehreren Variablen gilt das nicht.
- Lösung: Check, ob die **deskriptive Varianz-Kovarianz-Matrix**

$$S_{jk} = \frac{1}{n} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)$$

den vollen Rang $J + 1$ hat bzw. $\det \mathbf{S} \neq 0$

- Für $n < J + 1$ ist das trivialerweise nicht erfüllt

Wie entdeckt man Multi-Kolinearität?

- Gegeben: n Datensätze $\{x_{i0}, ..., x_{ij}, ..., x_{iJ}\}$, $i = 1, ..., n$ (die Datensätze enthalten auch y_i, aber das ist hier nicht relevant)
- x_{ij} ist der j^{th} Faktor im Datensatz i
- Multikollinearität ist gegeben, wenn man einen Faktor x_k als Linearkombination aus allen anderen Faktoren *für alle Datensätze* ausdrücken kann:

$$x_{ik} = \sum_{j \neq k} c_j x_{ij} \quad \forall i = 1, ..., n, \quad \text{konstante Koeffizienten } c_j$$

- Bei zwei Variablen ist dies leicht: $x_2 = c_0 x_1$ bzw Korrelationskoeffizient ist $=1$ oer $= -1$. Bei mehreren Variablen gilt das nicht.
- Lösung: Check, ob die **deskriptive Varianz-Kovarianz-Matrix**

$$S_{jk} = \frac{1}{n} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)$$

den vollen Rang $J + 1$ hat bzw. $\det \mathbf{S} \neq 0$

- Für $n < J + 1$ ist das trivialerweise nicht erfüllt

Wie entdeckt man Multi-Kolinearität?

- Gegeben: n Datensätze $\{x_{i0}, ..., x_{ij}, ..., x_{iJ}\}$, $i = 1, ..., n$ (die Datensätze enthalten auch y_i, aber das ist hier nicht relevant)
- x_{ij} ist der j^{th} Faktor im Datensatz i
- Multikollinearität ist gegeben, wenn man einen Faktor x_k als Linearkombination aus allen anderen Faktoren *für alle Datensätze* ausdrücken kann:

$$x_{ik} = \sum_{j \neq k} c_j x_{ij} \quad \forall i = 1, ..., n, \quad \text{konstante Koeffizienten } c_j$$

- Bei zwei Variablen ist dies leicht: $x_2 = c_0 x_1$ bzw Korrelationskoeffizient ist =1 oer $= -1$. Bei mehreren Variablen gilt das nicht.
- Lösung: Check, ob die **deskriptive Varianz-Kovarianz-Matrix**

$$S_{jk} = \frac{1}{n} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)$$

den vollen Rang $J + 1$ hat bzw. $\det \mathbf{S} \neq 0$

- Für $n < J + 1$ ist das trivialerweise nicht erfüllt

Wie entdeckt man Multi-Kolinearität?

- Gegeben: n Datensätze $\{x_{i0}, ..., x_{ij}, ..., x_{iJ}\}$, $i = 1, ..., n$ (die Datensätze enthalten auch y_i, aber das ist hier nicht relevant)
- x_{ij} ist der j^{th} Faktor im Datensatz i
- Multikollinearität ist gegeben, wenn man einen Faktor x_k als Linearkombination aus allen anderen Faktoren *für alle Datensätze* ausdrücken kann:

$$x_{ik} = \sum_{j \neq k} c_j x_{ij} \quad \forall i = 1, ..., n, \quad \text{konstante Koeffizienten } c_j$$

- Bei zwei Variablen ist dies leicht: $x_2 = c_0 x_1$ bzw Korrelationskoeffizient ist =1 oer $= -1$. Bei mehreren Variablen gilt das nicht.
- Lösung: Check, ob die **deskriptive Varianz-Kovarianz-Matrix**

$$S_{jk} = \frac{1}{n} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)$$

den vollen Rang $J + 1$ hat bzw. $\det \mathbf{S} \neq 0$

- Für $n < J + 1$ ist das trivialerweise nicht erfüllt

Wie entdeckt man Multi-Kolinearität?

- Gegeben: n Datensätze $\{x_{i0}, ..., x_{ij}, ..., x_{iJ}\}$, $i = 1, ..., n$ (die Datensätze enthalten auch y_i, aber das ist hier nicht relevant)
- x_{ij} ist der j^{th} Faktor im Datensatz i
- Multikollinearität ist gegeben, wenn man einen Faktor x_k als Linearkombination aus allen anderen Faktoren *für alle Datensätze* ausdrücken kann:

$$x_{ik} = \sum_{j \neq k} c_j x_{ij} \quad \forall i = 1, ..., n, \quad \text{konstante Koeffizienten } c_j$$

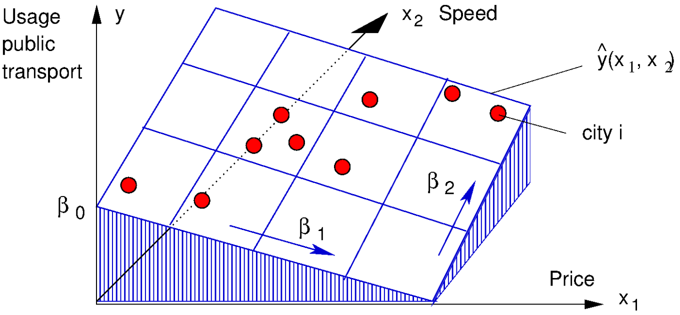
- Bei zwei Variablen ist dies leicht: $x_2 = c_0 x_1$ bzw Korrelationskoeffizient ist =1 oer $= -1$. Bei mehreren Variablen gilt das nicht.
- Lösung: Check, ob die **deskriptive Varianz-Kovarianz-Matrix**

$$S_{jk} = \frac{1}{n} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)$$

den vollen Rang $J + 1$ hat bzw. $\det \mathbf{S} \neq 0$

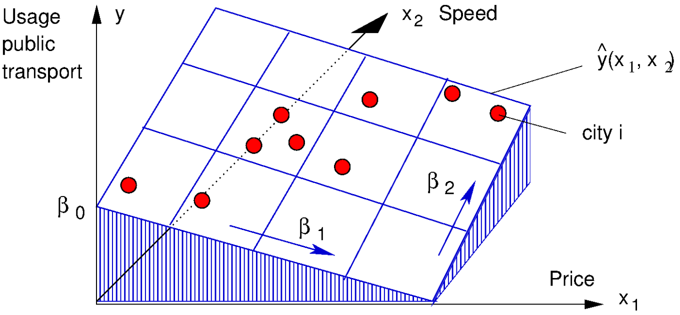
- Für $n < J + 1$ ist das trivialerweise nicht erfüllt

Datenspezifikation 2: Beispiel



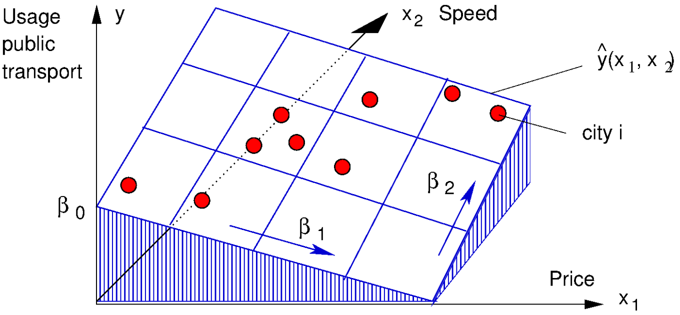
- Die Nachfrage y_i für den ÖPNV in den Städten i hängt vom Preis x_{i1} und der Qualität x_{i2} (Proxy: Haustür-zu-Haustür-Geschwindigkeit) ab.
- Parameter: Achsabschnitt (*intercept*) β_0, Preissensitivität β_1, Qualitätsbewusstsein β_2
- Preis und Qualität sind (*wie bei fast allen Produkten und Dienstleistungen*) korreliert, aber unvollständig: OK

Datenspezifikation 2: Beispiel



- Die Nachfrage y_i für den ÖPNV in den Städten i hängt vom Preis x_{i1} und der Qualität x_{i2} (Proxy: Haustür-zu-Haustür-Geschwindigkeit) ab.
- Parameter: Achsabschnitt (*intercept*) β_0, Preissensitivität β_1, Qualitätsbewusstsein β_2
- Preis und Qualität sind (*wie bei fast allen Produkten und Dienstleistungen*) korreliert, aber unvollständig: OK

Datenspezifikation 2: Beispiel



- Die Nachfrage y_i für den ÖPNV in den Städten i hängt vom Preis x_{i1} und der Qualität x_{i2} (Proxy: Haustür-zu-Haustür-Geschwindigkeit) ab.
- Parameter: Achsabschnitt (*intercept*) β_0, Preissensitivität β_1, Qualitätsbewusstsein β_2
- Preis und Qualität sind (*wie bei fast allen Produkten und Dienstleistungen*) korreliert, aber unvollständig: OK

11.3 Kleinste-Fehlerquadrate (KQ) Schätzung

- Gegeben: lineares Modell der Form

$$y(\boldsymbol{x}) = \boldsymbol{\beta}'\boldsymbol{x} + \epsilon = \hat{y}(\boldsymbol{x}) + \epsilon, \quad \epsilon \sim i.i.d.$$

  welches *alle Gauß-Markow-Annahmen mit Ausnahme der gaußverteilten Residualterme* erfüllt

- Gegeben sind ebenfalls n vollständige Datensätze aller exogenen Faktoren und der endogenen Variablen, welche die Datenspezifikation erfüllen:

$$\{\boldsymbol{p}_i = (x_{i0}, ..., x_{iJ}, y_i)', \quad i = 1, ..., n\}$$

- Gesucht ist ein Parameterschätzer $\hat{\beta}$, welcher die Fehlerquadratsumme zwischen den deterministischen Modellvoraussagen $x_i'\beta$ und der beobachteten endogenen Variable y_i bezüglich der Parameter minimiert:

$$\boxed{\hat{\beta} = \text{arg min}_{\beta}\, S(\beta)}$$

  wobei

$$S(\beta) = \boldsymbol{\epsilon}'\boldsymbol{\epsilon} = (\boldsymbol{y} - \mathbf{X}\,\beta)'(\boldsymbol{y} - \mathbf{X}\,\beta).$$

11.3 Kleinste-Fehlerquadrate (KQ) Schätzung

- Gegeben: lineares Modell der Form

$$y(\boldsymbol{x}) = \boldsymbol{\beta}'\boldsymbol{x} + \epsilon = \hat{y}(\boldsymbol{x}) + \epsilon, \quad \epsilon \sim i.i.d.$$

  welches *alle Gauß-Markow-Annahmen mit Ausnahme der gaußverteilten Residualterme* erfüllt

- Gegeben sind ebenfalls n vollständige Datensätze aller exogenen Faktoren und der endogenen Variablen, welche die Datenspezifikation erfüllen:

$$\{\boldsymbol{p}_i = (x_{i0}, ..., x_{iJ}, y_i)', \quad i = 1, ..., n\}$$

- Gesucht ist ein Parameterschätzer $\hat{\beta}$, welcher die Fehlerquadratsumme zwischen den deterministischen Modellvoraussagen $x_i'\beta$ und der beobachteten endogenen Variable y_i bezüglich der Parameter minimiert:

$$\hat{\beta} = \text{arg min}_\beta \, S(\beta)$$

  wobei

$$S(\beta) = \boldsymbol{\epsilon}'\boldsymbol{\epsilon} = (\boldsymbol{y} - \mathbf{X}\,\beta)'(\boldsymbol{y} - \mathbf{X}\,\beta).$$

11.3 Kleinste-Fehlerquadrate (KQ) Schätzung

- Gegeben: lineares Modell der Form

$$y(\boldsymbol{x}) = \boldsymbol{\beta}'\boldsymbol{x} + \epsilon = \hat{y}(\boldsymbol{x}) + \epsilon, \quad \epsilon \sim i.i.d.$$

  welches *alle Gauß-Markow-Annahmen mit Ausnahme der gaußverteilten Residualterme* erfüllt

- Gegeben sind ebenfalls n vollständige Datensätze aller exogenen Faktoren und der endogenen Variablen, welche die Datenspezifikation erfüllen:

$$\{\boldsymbol{p}_i = (x_{i0}, ..., x_{iJ}, y_i)', \quad i = 1, ..., n\}$$

- Gesucht ist ein Parameterschätzer $\hat{\beta}$, welcher die Fehlerquadratsumme zwischen den deterministischen Modellvoraussagen $x_i'\beta$ und der beobachteten endogenen Variable y_i bezüglich der Parameter minimiert:

$$\boxed{\hat{\beta} = \text{arg min}_{\beta}\, S(\beta)}$$

  wobei

$$S(\beta) = \boldsymbol{\epsilon}'\boldsymbol{\epsilon} = (\boldsymbol{y} - \mathbf{X}\,\beta)'(\boldsymbol{y} - \mathbf{X}\,\beta).$$

11.3 Kleinste-Fehlerquadrate (KQ) Schätzung

- Gegeben: lineares Modell der Form

$$y(\boldsymbol{x}) = \boldsymbol{\beta}'\boldsymbol{x} + \epsilon = \hat{y}(\boldsymbol{x}) + \epsilon, \quad \epsilon \sim i.i.d.$$

  welches *alle Gauß-Markow-Annahmen mit Ausnahme der gaußverteilten Residualterme* erfüllt

- Gegeben sind ebenfalls n vollständige Datensätze aller exogenen Faktoren und der endogenen Variablen, welche die Datenspezifikation erfüllen:

$$\{\boldsymbol{p}_i = (x_{i0}, ..., x_{iJ}, y_i)', \quad i = 1, ..., n\}$$

- Gesucht ist ein Parameterschätzer $\hat{\boldsymbol{\beta}}$, welcher die Fehlerquadratsumme zwischen den deterministischen Modellvoraussagen $\boldsymbol{x}_i'\boldsymbol{\beta}$ und der beobachteten endogenen Variable y_i bezüglich der Parameter minimiert:

$$\boxed{\hat{\boldsymbol{\beta}} = \text{arg min}_{\boldsymbol{\beta}}\, S(\boldsymbol{\beta})}$$

  wobei

$$S(\boldsymbol{\beta}) = \boldsymbol{\epsilon}'\boldsymbol{\epsilon} = (\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta})'(\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta}).$$

Bestimmung des KQ-Schätzers

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

Bestimmung des KQ-Schätzers

$$\begin{aligned} S &= (\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta})'\,(\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta}) \\ \text{[Distributivität } \rightarrow\text{]} \quad &= \boldsymbol{y}'\boldsymbol{y} - (\mathbf{X}\,\boldsymbol{\beta})'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + (\mathbf{X}\,\boldsymbol{\beta})'\mathbf{X}\,\boldsymbol{\beta} \end{aligned}$$

Bestimmung des KQ-Schätzers

$$
\begin{aligned}
S &= (\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta})'\,(\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta}) \\
\text{[Distributivität} \rightarrow] &= \boldsymbol{y}'\boldsymbol{y} - (\mathbf{X}\,\boldsymbol{\beta})'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + (\mathbf{X}\,\boldsymbol{\beta})'\mathbf{X}\,\boldsymbol{\beta} \\
\text{[Transpositionsregeln} \rightarrow] &= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\beta}'\mathbf{X}'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\,\boldsymbol{\beta}
\end{aligned}
$$

Bestimmung des KQ-Schätzers

$$
\begin{aligned}
S &= (\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta})'\,(\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta}) \\
\text{[Distributivität} \rightarrow] &= \boldsymbol{y}'\boldsymbol{y} - (\mathbf{X}\,\boldsymbol{\beta})'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + (\mathbf{X}\,\boldsymbol{\beta})'\mathbf{X}\,\boldsymbol{\beta} \\
\text{[Transpositionsregeln} \rightarrow] &= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\beta}'\mathbf{X}'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\,\boldsymbol{\beta} \\
\text{[Transpositionsregeln} \rightarrow] &= \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{\beta}'(\mathbf{X}'\boldsymbol{y}) + \boldsymbol{\beta}'\left(\mathbf{X}'\mathbf{X}\right)\boldsymbol{\beta}
\end{aligned}
$$

Bestimmung des KQ-Schätzers

$$\begin{aligned}
S &= (\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta})'\,(\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta}) \\
\text{[Distributivität } \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - (\mathbf{X}\,\boldsymbol{\beta})'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + (\mathbf{X}\,\boldsymbol{\beta})'\mathbf{X}\,\boldsymbol{\beta} \\
\text{[Transpositionsregeln } \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\beta}'\mathbf{X}\,'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}\,'\mathbf{X}\,\boldsymbol{\beta} \\
\text{[Transpositionsregeln } \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{\beta}'(\mathbf{X}\,'\boldsymbol{y}) + \boldsymbol{\beta}'\left(\mathbf{X}\,'\mathbf{X}\,\right)\boldsymbol{\beta}
\end{aligned}$$

Ableiten nach $\boldsymbol{\beta}$ und Nullsetzen mit den Regeln $\frac{\partial}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}'\boldsymbol{a}) = \boldsymbol{a}$ und $\frac{\partial}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}'\mathbf{A}\,\boldsymbol{\beta}) = (\mathbf{A}\, + \mathbf{A}\,')\boldsymbol{\beta}$, wobei $\mathbf{A}\, = \mathbf{X}\,'\mathbf{X}$:

Bestimmung des KQ-Schätzers

$$
\begin{aligned}
S &= (\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta})'\,(\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta}) \\
\text{[Distributivität} \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - (\mathbf{X}\,\boldsymbol{\beta})'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + (\mathbf{X}\,\boldsymbol{\beta})'\mathbf{X}\,\boldsymbol{\beta} \\
\text{[Transpositionsregeln} \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\beta}'\mathbf{X}'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\,\boldsymbol{\beta} \\
\text{[Transpositionsregeln} \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{\beta}'(\mathbf{X}'\boldsymbol{y}) + \boldsymbol{\beta}'\left(\mathbf{X}'\mathbf{X}\right)\boldsymbol{\beta}
\end{aligned}
$$

Ableiten nach $\boldsymbol{\beta}$ und Nullsetzen mit den Regeln $\frac{\partial}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}'\boldsymbol{a}) = \boldsymbol{a}$ und $\frac{\partial}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}'\mathbf{A}\,\boldsymbol{\beta}) = (\mathbf{A} + \mathbf{A}')\boldsymbol{\beta}$, wobei $\mathbf{A} = \mathbf{X}'\mathbf{X}$:

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

Bestimmung des KQ-Schätzers

$$\begin{aligned}
S &= (\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta})'\,(\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta}) \\
\text{[Distributivität} \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - (\mathbf{X}\,\boldsymbol{\beta})'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + (\mathbf{X}\,\boldsymbol{\beta})'\mathbf{X}\,\boldsymbol{\beta} \\
\text{[Transpositionsregeln} \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\beta}'\mathbf{X}'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\,\boldsymbol{\beta} \\
\text{[Transpositionsregeln} \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{\beta}'(\mathbf{X}'\boldsymbol{y}) + \boldsymbol{\beta}'\left(\mathbf{X}'\mathbf{X}\right)\boldsymbol{\beta}
\end{aligned}$$

Ableiten nach $\boldsymbol{\beta}$ und Nullsetzen mit den Regeln $\frac{\partial}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}'\boldsymbol{a}) = \boldsymbol{a}$ und $\frac{\partial}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}'\mathbf{A}\,\boldsymbol{\beta}) = (\mathbf{A} + \mathbf{A}')\boldsymbol{\beta}$, wobei $\mathbf{A} = \mathbf{X}'\mathbf{X}$:

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

$$\mathbf{X}'\mathbf{X}\,\boldsymbol{\beta} = \mathbf{X}'\boldsymbol{y}$$

Bestimmung des KQ-Schätzers

$$
\begin{aligned}
S &= (\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta})'\,(\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta}) \\
\text{[Distributivität} \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - (\mathbf{X}\,\boldsymbol{\beta})'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + (\mathbf{X}\,\boldsymbol{\beta})'\mathbf{X}\,\boldsymbol{\beta} \\
\text{[Transpositionsregeln} \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\beta}'\mathbf{X}'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\,\boldsymbol{\beta} \\
\text{[Transpositionsregeln} \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{\beta}'(\mathbf{X}'\boldsymbol{y}) + \boldsymbol{\beta}'\left(\mathbf{X}'\mathbf{X}\right)\boldsymbol{\beta}
\end{aligned}
$$

Ableiten nach $\boldsymbol{\beta}$ und Nullsetzen mit den Regeln $\frac{\partial}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}'\boldsymbol{a}) = \boldsymbol{a}$ und $\frac{\partial}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}'\mathbf{A}\,\boldsymbol{\beta}) = (\mathbf{A} + \mathbf{A}')\boldsymbol{\beta}$, wobei $\mathbf{A} = \mathbf{X}'\mathbf{X}$:

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

$$
\mathbf{X}'\mathbf{X}\,\boldsymbol{\beta} \;=\; \mathbf{X}'\boldsymbol{y} \quad \mid \cdot\, (\mathbf{X}'\mathbf{X})^{-1}
$$

Bestimmung des KQ-Schätzers

$$\begin{aligned} S &= (\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta})'\,(\boldsymbol{y} - \mathbf{X}\,\boldsymbol{\beta}) \\ \text{[Distributivität } \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - (\mathbf{X}\,\boldsymbol{\beta})'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + (\mathbf{X}\,\boldsymbol{\beta})'\mathbf{X}\,\boldsymbol{\beta} \\ \text{[Transpositionsregeln } \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\beta}'\mathbf{X}'\boldsymbol{y} - \boldsymbol{y}'\mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\,\boldsymbol{\beta} \\ \text{[Transpositionsregeln } \rightarrow\text{]} &= \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{\beta}'(\mathbf{X}'\boldsymbol{y}) + \boldsymbol{\beta}'\left(\mathbf{X}'\mathbf{X}\right)\boldsymbol{\beta} \end{aligned}$$

Ableiten nach $\boldsymbol{\beta}$ und Nullsetzen mit den Regeln $\frac{\partial}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}'\boldsymbol{a}) = \boldsymbol{a}$ und $\frac{\partial}{\partial \boldsymbol{\beta}}(\boldsymbol{\beta}'\mathbf{A}\,\boldsymbol{\beta}) = (\mathbf{A} + \mathbf{A}')\boldsymbol{\beta}$, wobei $\mathbf{A} = \mathbf{X}'\mathbf{X}$:

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

$$\mathbf{X}'\mathbf{X}\,\boldsymbol{\beta} \quad = \quad \mathbf{X}'\boldsymbol{y} \quad \Big| \cdot \left(\mathbf{X}'\mathbf{X}\right)^{-1}$$

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