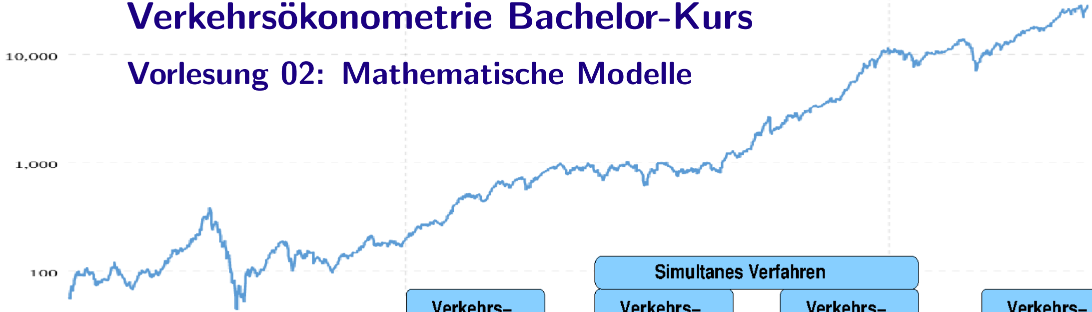
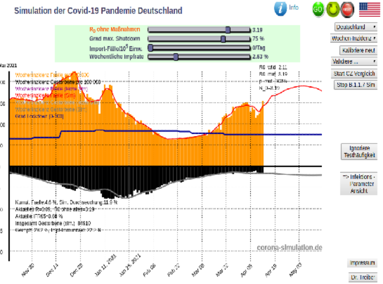
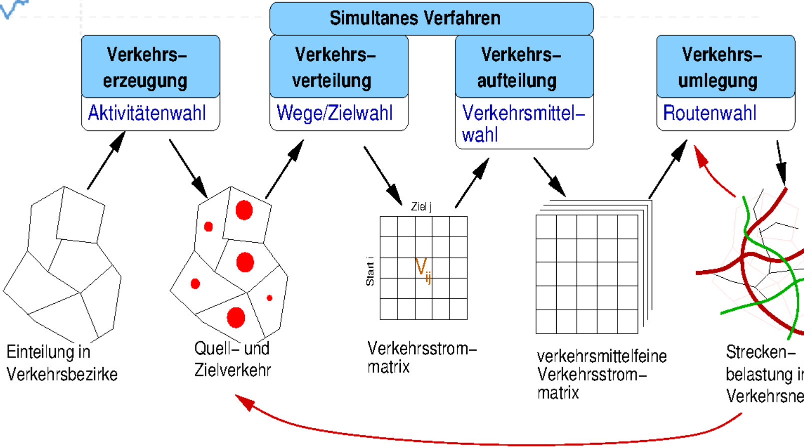
Verkehrsökonometrie Bachelor-Kurs

Vorlesung 02: Mathematische Modelle







2. Mathematische Modelle

Zur Erinnerung: Mathematische Modell $\Rightarrow$ **Modellfunktion**

$$\boldsymbol{f}(\tilde{\boldsymbol{x}}, \boldsymbol{\beta}) + \boldsymbol{\epsilon}$$

- Die Modellfunktion selbst ($\boldsymbol{f}$, ggf $\boldsymbol{\epsilon}$) bestimmt die qualitativen Aspekte
- $\tilde{\boldsymbol{x}}$ sind die exogenen Variablen, die meist in Faktoren $\boldsymbol{x}$ zusammengefasst werden (s.u.)
- $\boldsymbol{\beta}$ sind die Modellparameter, die das Modell quantitativ einstellen
- $\boldsymbol{\epsilon}$ die unbestimmten bzw. Zufallsanteile

Hier geht es nun um die Modellfunktion und mögliche Arten:

- Ein- oder Mehrgleichungsmodelle (Komponentenzahl von $\boldsymbol{f}$ gleich #endogener Variablen)
- linear *vs.* nichtlinear
- deterministisch ($\epsilon = 0$) *vs.* stochastisch
- interne Abhängigkeitsstrukturen: Modell besteht aus mehreren verketteten, gekoppelten oder rückgekoppelten Teilmodellen

> *Mach es so einfach wie möglich, aber nicht einfacher* (Einstein)

2.1 Ein- vs. Mehrgleichungsmodelle

Die Zahl der Gleichungen ist gleich der Zahl der (wesentlichen) endogenen Variablen

- Hat man nur eine endogene Variable Y wie bei Regressionsmodellen, reduziert sich im resultierenden **Eingleichungsmodell** die Funktion $\boldsymbol{f}(.)$ auf eine gewöhnliche skalare Funktion $f(.)$
- **Lineares** Modell oder Modell ohne **Kopplung** der endogenen Variablen $\Rightarrow$ ein Modell für $K > 1$ endogene Variable zerfällt in K Eingleichungsmodelle
- Beispiel Eingleichungsmodell: Verkehrsstärke als Funktion der Tageszeit, Verbrauch als Funktion der Geschwindigkeit, Beschleunigung und Steigung
- Beispiel **Mehrgleichungsmodell**: diskrete Wahlentscheidungen.
  - Die mikroskopischen endogenen Variablen $Y_k \in \{0, 1\}$ sind über die Bedingung $\sum_k Y_k = 1$ gekoppelt bzw. aggregiert
  - Die makroskopischen Wahrscheinlichkeiten $P_k(\boldsymbol{x}, \boldsymbol{\beta})$ sind ebenfalls durch $\sum_k P_k = 1$ gekoppelt

2.2 Linear vs. nichtlinear

Es gibt vier Stufen von Linearität zu Nichtlinearität:

1. **Wahrhaft lineare Modelle:**

$$Y = \hat{y}(\tilde{\boldsymbol{x}}, \boldsymbol{\beta}) + \epsilon = \sum_{j=0}^{J} \beta_j \tilde{x}_j + \epsilon = \boldsymbol{\beta}' \tilde{\boldsymbol{x}} + \epsilon$$

   Linearität $\Rightarrow$ keine Kopplung der endogenen Variablen $\Rightarrow$ man kann jede Komponente separat betrachten

2. **Parameterlineare (quasilineare) Modelle**

$$Y = \hat{y}(\tilde{\boldsymbol{x}}, \boldsymbol{\beta}) + \epsilon = \sum_{j=0}^{J} \beta_j g_j(\tilde{\boldsymbol{x}}) + \epsilon = \sum_{j=0}^{J} \beta_j x_j + \epsilon = \boldsymbol{\beta}' \boldsymbol{x} + \epsilon$$

   Linearisierung mit *aufgrund des Sachverhalts* festen Funktionen $x_j = g_j(\tilde{\boldsymbol{x}})$

   - Die **linearen Faktoren** x_j werden die neuen exogenen Variablen des nun lineraen Modells
   - I.A. ist die Zahl der Faktoren $\neq$ der Zahl der ursprünglichen exogenen Variablen $\Rightarrow$ Beispiele

2.2.2 Beispiel 1: Jährliche Kilometerleistung mit dem Kfz

- y: Kilometerleistung (deterministisch, da "mittlere"), $\tilde{x}_1$: Einkommen [€/Jahr], $\tilde{x}_2$: Spritkosten [€/Liter]
- Zwei exogene Variable $\Rightarrow$ vier Faktoren:

$$\begin{aligned} x_0 &= g_0(\tilde{\boldsymbol{x}}) = 1, \\ x_1 &= g_1(\tilde{\boldsymbol{x}}) = \tilde{x}_1, \\ x_2 &= g_2(\tilde{\boldsymbol{x}}) = \tilde{x}_2, \\ x_3 &= g_3(\tilde{\boldsymbol{x}}) = \tilde{x}_1\tilde{x}_2 \end{aligned}$$

$\Rightarrow$ **Modell**: $$y = \beta_0 + \beta_1\tilde{x}_1 + \beta_2\tilde{x}_2 + \beta_3\tilde{x}_1\tilde{x}_2 = \sum_{j=0}^{3} \beta_j x_j$$

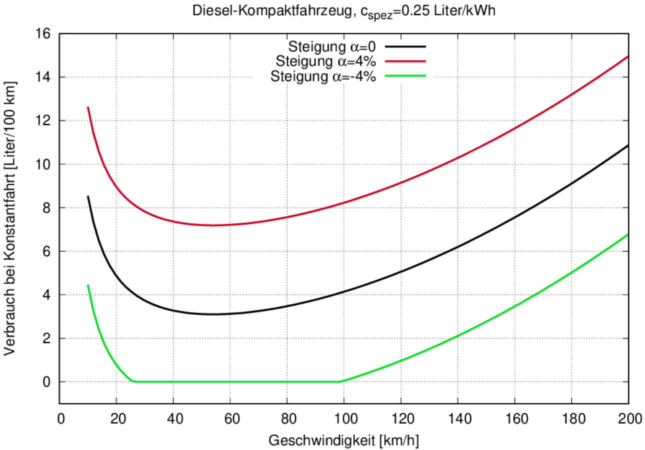
? Diskutiere die Elastizität $\epsilon_2 = \frac{\bar{x}_2}{\bar{y}}\frac{\mathrm{d}y}{\mathrm{d}x_2} = \frac{\beta_2\bar{x}_2}{y} = -0.15$

! 0.15 % Reduzierung der Kilometerleistung um 0.15 % pro Preisanstieg um 1 %

? Diskutiere die Bedeutung der Faktoren, insbesondere das Produkt $\tilde{x}_1\tilde{x}_2$

! $x_0 = 1$: Konstante (Achsabschnitt); $x_1 = \tilde{x}_1$: Erhöhung der Kilometerleistung mit dem Einkommen ($\beta_1 > 0$ erwartet); $x_2 = \tilde{x}_2$: Preissensitivität ($\beta_2 < 0$); $x_3 = \tilde{x}_1\tilde{x}_2$: Erhöhung der Preissensitivität (im Sinne weniger negativer Werte) mit dem Einkommen: Sensitivität $\beta_2 + \beta_3\tilde{x}_1$ ($\beta_3 > 0$ erwartet)

2.2.2 Beispiel 2: Treibstoffverbrauch



- y: Verbrauch [l/100 km]
- $\tilde{x}_1$: Geschwindigkeit v [m/s]
- $\tilde{x}_2$: Steigung α [rad] (1% entspricht $\tilde{x}_2 = 0.01$)
- Modell aus der Physik: Verbrauch $y = cP(v, \alpha)/v$ mit c fester Konstante und Leistung
$$P(v, \alpha) = P_0 + mg(\mu + \alpha)v + \frac{1}{2} c_w \rho A v^3$$
- Damit Modell mit physikbestimmten nichtlin. Funktionen $g_j(\tilde{\boldsymbol{x}})$:
$$y(\tilde{\boldsymbol{x}}) = \beta_0 + \beta_1 \frac{1}{\tilde{x}_1} + \beta_2 \tilde{x}_1^2 + \beta_3 \tilde{x}_2$$

? Wie hängen die Modellparameter mit den physikalischen Parametern zusammen?

! $\beta_0 = cmg\mu$, $\beta_1 = cP_0$, $\beta_2 = c/2 c_w \rho A$, $\beta_3 = cmg$

? Was ist mit der grünen Kurve in der Mitte los?

2.2.3 Linear vs. nichtlinear: reduzibel nichtlinear

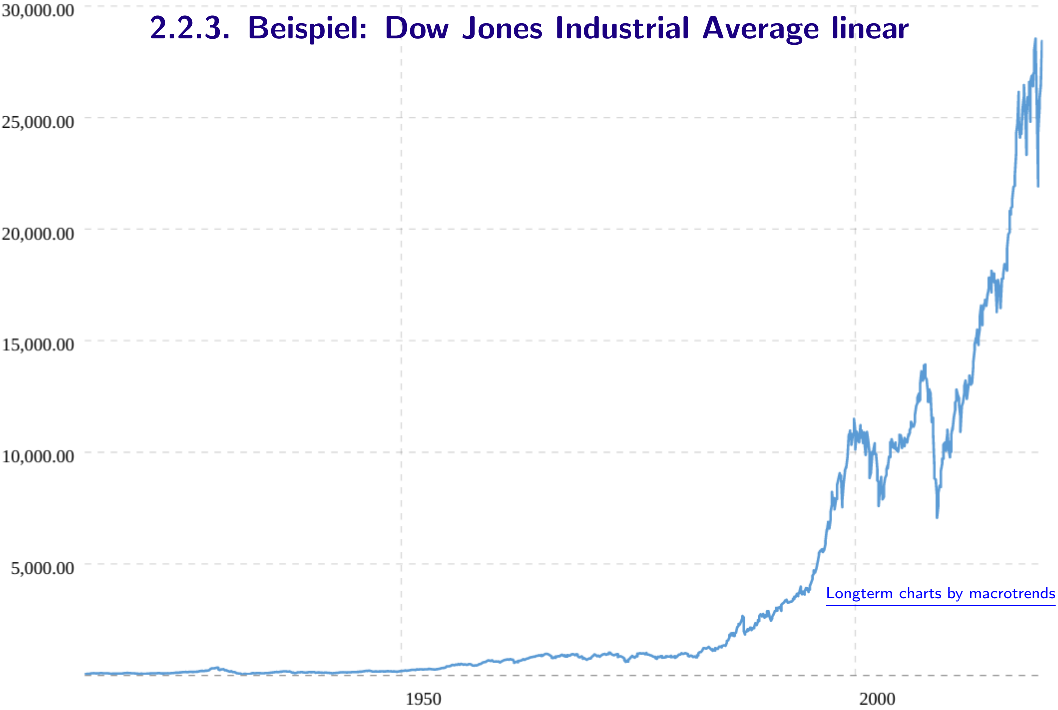
Die Zusammenfassung in Faktoren transformiert die *exogenen* Variablen. Analog kann man manchmal auch die *endogenen* Variablen transformieren, um Linearität zu erreichen. Klassiker ist das

Modell für unbeschränkes Wachstum $G(t) = G_0 e^{rt+\epsilon}$

- Urprüngliche endogene Variable G: Wachstumsmerkmal, z.B. Umsatz oder Gewinn von Firmen, Aktienkurse, BIP, Infektionszahlen ohne jede Herdenimmunität
- Exogene Variable: Zeit t
- Parameter: Anfangswert G_0 und Wachstumsrate r: $1/r$ gibt das Zeitintervall für einen Wachstum um den Faktor $e = 2.71...$ an
- Multiplikativer unbestimmter Faktor e^ϵ
- **Linearisierung**: $Y(t) = \ln G(t) = \ln G_0 + \frac{t}{\tau} + \epsilon$
- **Standardform**: $Y(\boldsymbol{x}) = \boldsymbol{\beta}'\boldsymbol{x} + \epsilon = \beta_0 x_0 + \beta_1 x_1 + \epsilon$
  mit $x_0 = 1$, $x_1 = t$, $\beta_0 = \ln G_0$, $\beta_1 = r$

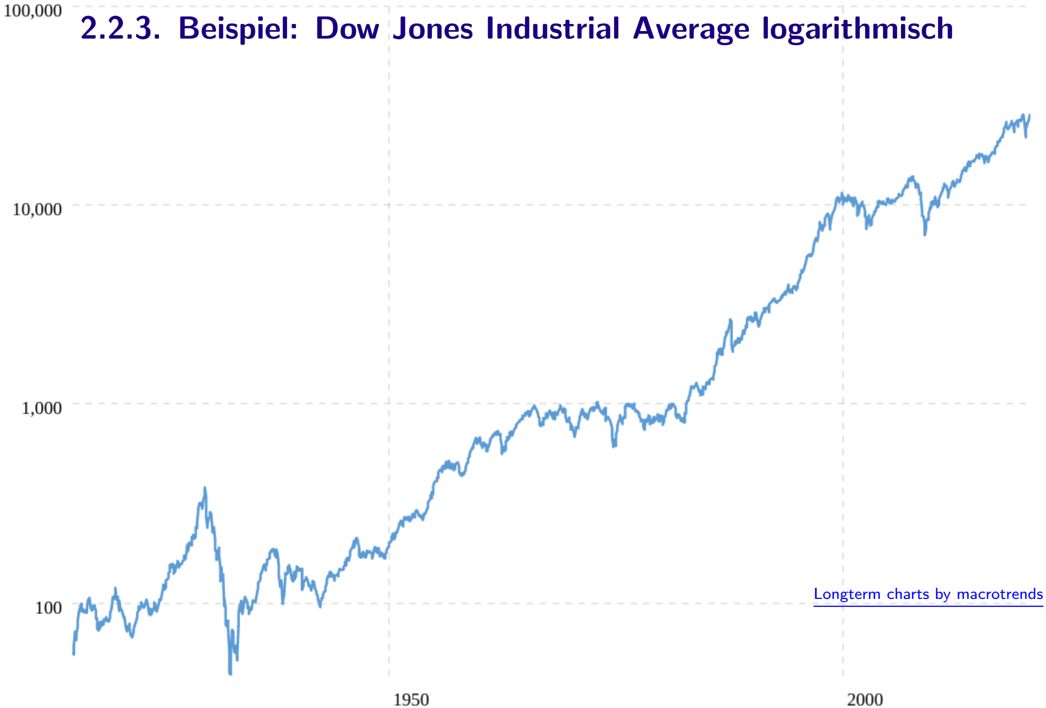
Merke: Anders als die Transformation der exogenen Variablen beeinflusst die der endogenen Variablen auch den Zufallsanteil, macht hier beispielsweise aus einem multiplikativen Zufallsfaktor einen additiven Anteil

2.2.3. Beispiel: Dow Jones Industrial Average linear



Longterm charts by macrotrends

2.2.3. Beispiel: Dow Jones Industrial Average logarithmisch



Longterm charts by macrotrends

2.2.4 Linear vs. nichtlinear: irreduzibel nichtlinear

Klassiker: beschränktes Wachstum

$$y(t) = \frac{y_s}{1 + (y_s/y_0 - 1)e^{-rt}}$$

- Lösung der ODE

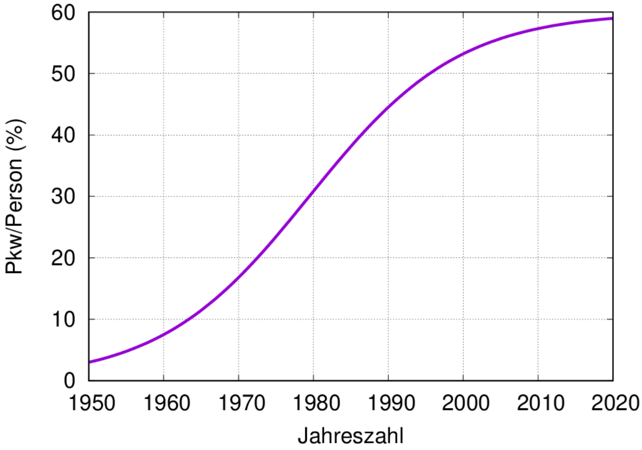
$$\frac{\mathrm{d}y}{\mathrm{d}t} = ry(t)\left(1 - \frac{y(t)}{y_s}\right)$$

für den Anfangswert $y(t_0) = y_0$

- Plot für $t_0 = 1950$, $y_0 = 3\,\%$, $y_s = 60\,\%$, und $r = 1/10\,\mathrm{Jahre}$

? Was könnte dies darstellen?

! z.B. Markdurchdringung eines neuen Produktes (Smartphone, Autos ...)



2.3 Weitere Kriterien für die Modellfunktion

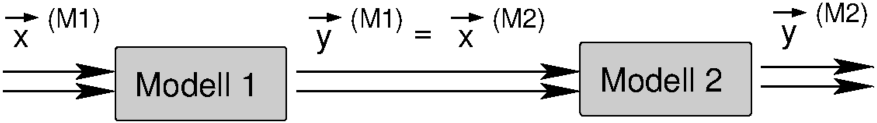
- **Deterministisch** *vs.* **stochastisch**
- Natur des Modells: **mikroskopisch** *vs.* **makroskopisch**
- Zahl der endogenen Variablen: Ein- *vs.* Mehrgleichungsmodelle ($\Rightarrow$ 2.1)
- Skalierung der endogenen Variablen: reellwertig (metrisch) $\rightarrow$ z.B. **Regressionsmodelle** *vs.* diskret (ordinal- oder nominalskaliert) $\rightarrow$ z.B. **diskrete Wahltheorie**
- Abhängigkeitsstrukturen: Verkettung, Kopplung, Rückkopplung

Beispiele:

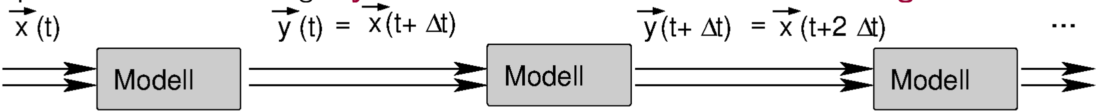
- Ein Modell für die zu erwartenden Wahrscheinlichkeiten beim Modal-Split ist ein *deterministisches*, *makroskopisches*, *kontinuierliches Mehrgleichungsmodell* mit *Verkettung* zweier Teilmodelle
- Ein Modell für die Wahl durch eine einzelne Person ist ein *stochastisches*, *mikroskopisches*, *diskretes Mehrgleichungsmodell* mit *Verkettung* zweier Teilmodelle

2.4 Abhängigkeitsstrukturen 1: Verkettung

Die endogenen Variablen des Modell 1 dienen als exogene Variable für Modell 2



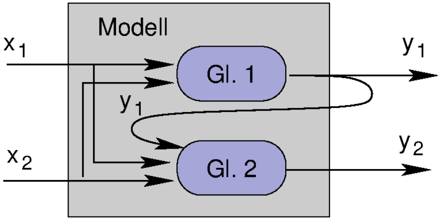
Spezialfall der Verkettung: **dynamische Modelle** für die **Zeitentwicklung**:



- Die endogenen Variablen zur Zeit t sind die exogenen zur Zeit $t+\Delta t$
- Das Modell für jeden Schritt ist oft dasselbe (**autonomes Modell**)
- *Beispiel*: Umwandlung der Differenzialgleichung $\frac{\mathrm{d}y}{\mathrm{d}t} = ry(t)\left(1 - y(t)/y_s\right)$ des Modells für beschränkten Wachstums in eine verkette Differenzengleichung

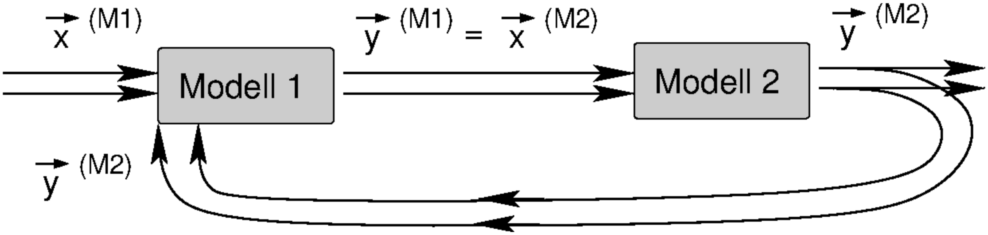
$$y(t+\Delta t) = y(t) + r\ \Delta t\ y(t)\left(1 - \frac{y(t)}{y_s}\right)$$

2.4 Abhängigkeitsstrukturen 2: Kopplung



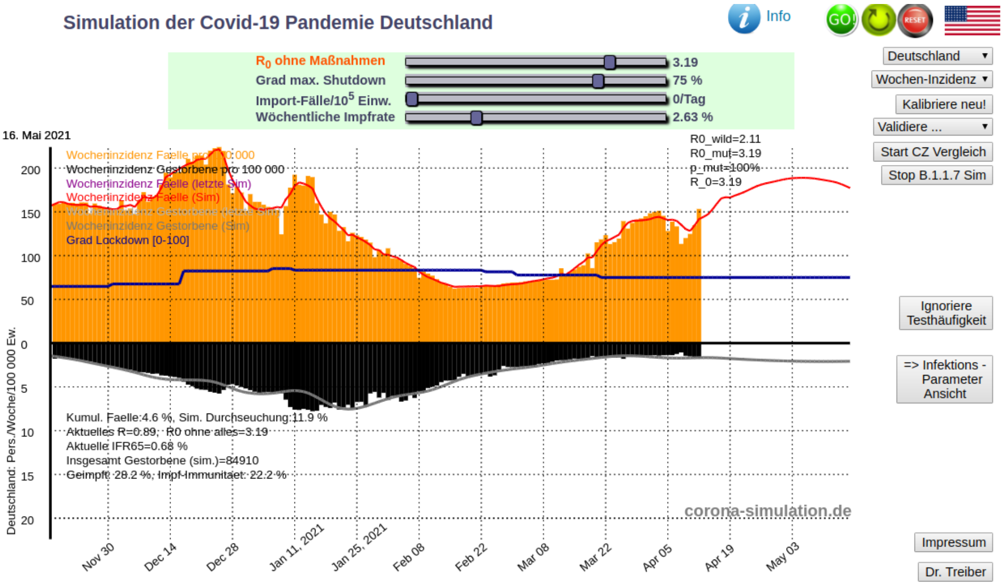
Eine oder mehrere der endogenen Variablen der Modellgleichung 1 werden als exogene Variable an Gleichung 2 gekoppelt

2.4 Abhängigkeitsstrukturen 3: Rückkopplung

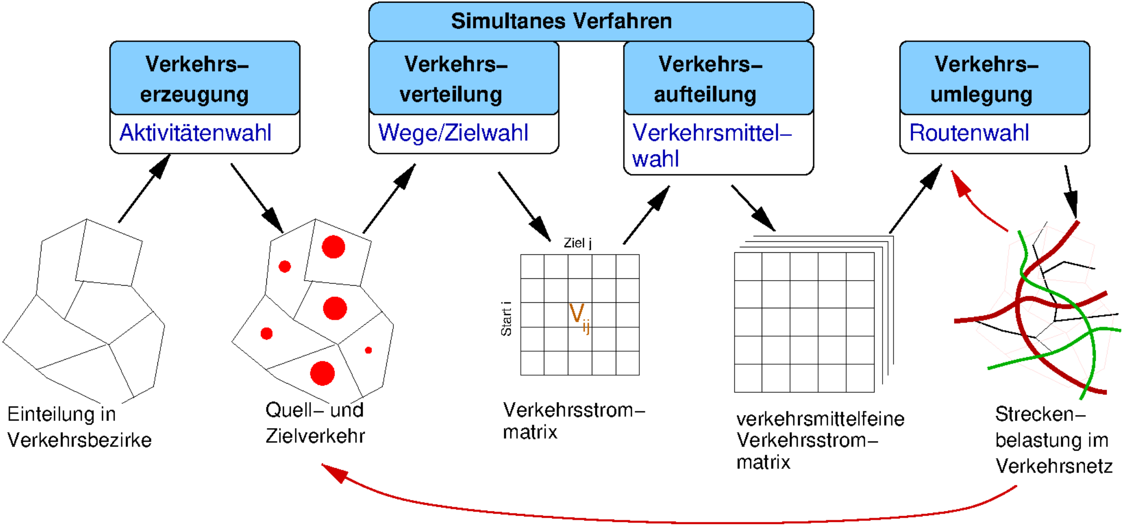


Kombination von Verkettung und Kopplung

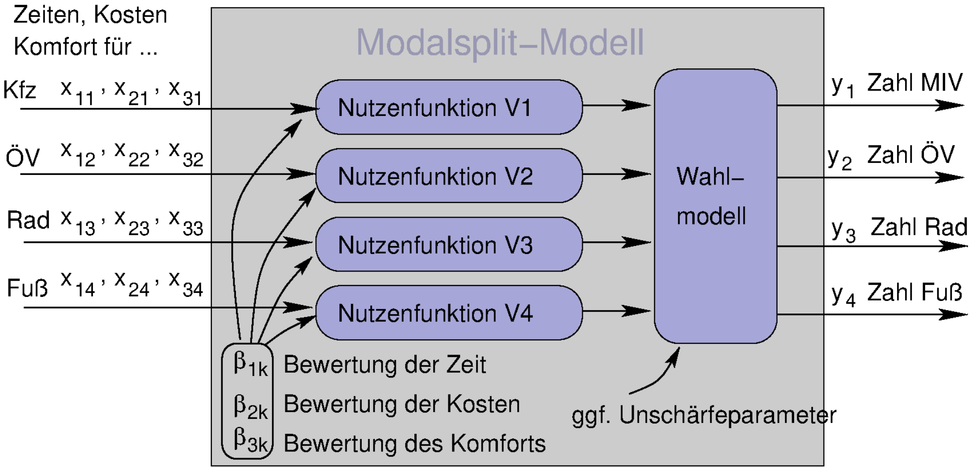
2.5 Modellbeispiel 1: Makromodell der Infektionsausbreitung $\Rightarrow$ VL 3



2.5 Modellbeispiel 2: Vierstufenmodell der Verkehrsplanung ⇒ VL 4ff



2.5 Modellbeispiel 3: Modal-Split Modell $\Rightarrow$ VL 6ff



- Exogene Variablen x_{jk}: Einflussfaktor j für Modus k
- Endogenous variables y_k: Auswahldummy $\in \{0, 1\}$ oder Auswahlwahrscheinlichkeit für Verkehrsmodus k

2.5 zu Beispiel 3: Inputdaten bei zwei Alternativen

	Alter	Ge-schlecht	Zeit TT Rad	Kosten Rad	Zeit TT ÖV	Kosten ÖV	Wahl Rad	Wahl PT
Variable	x_1	x_2	x_3	x_4	x_5	x_6	y_{1i}	y_{2i}
Person 1	30	w	20 min	0 €	30 min	1.00 €	0	1
Person 2	24	m	11 min	0 €	20 min	2.00 €	1	0
Person 3	27	m	34 min	0 €	15 min	2.00 €	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Daten z.B. von Interviews/Interneterhebungen $\Rightarrow$ VL 8