

ao.Univ.Prof. Dr. Edwin Deutsch

Institut für Wirtschaftsmathematik E105

Arbeitsgruppe EOS: Ökonometrie und Systemtheorie

Homepage: [www.eos.tuwien.ac.at/Oeko/EDeutsch](http://www.eos.tuwien.ac.at/Oeko/EDeutsch)

Einführung in die Vorlesungsreihe

Mikroökometrie (105.126 VO 3.0 WS, 105.128 UE 1.5 WS)

AKOEK Methoden der Mikroökometrie

(109.005 VO 2.0 WS, 109.004 UE 1.0 WS)

AKOEK angewandte Mikroökometrie in Wirtschafts- und

Marktforschung (119.005 VO 2.0 SS, 119.002 UE 1.0 SS)

Was ist Mikroökonomie ?

Sie werden als MathematikerIn mit der Aufbereitung von Entscheidungsdaten in Wirtschaft, Finanz oder Politik beauftragt.

In einem Umfrageinstitut sollen Sie anhand von Haushaltsumfragen für ein Firmenmanagement ermitteln,

- ob Leute zur gegebenen Marktlage lieber einen PKW der Standardklasse oder einer gehobenen Klasse nachfragen;

oder Sie prognostizieren im Auftrag einer Zeitung,

- wie die nächste Nationalratswahl ausgeht;

Sie untersuchen in einem Forschungsinstitut mit Mikrozensusdaten,

- welche Arbeitskräfte und Ausbildungswege zur Arbeitslosigkeit neigen und welche Maßnahmen die Dauer der Arbeitslosigkeit verkürzen,
- wieviel bestimmte sozialpolitischen Maßnahmen (zB für Wohnen, Familie oder Pension) dem Staat auf längere Frist kosten werden;

Sie arbeiten in einem Rating-Institut und evaluieren,

- welche Firmen auf Grund ihrer Bilanzen in die Kategorien AAA bis C einzuordnen sind;

- Sie arbeiten im Kreditscoring einer Bank und evaluieren,
- welche Darlehensanträge zwar ertragsversprechend, aber riskant sind und eine Absicherung gegen eine Zahlungsunfähigkeit empfehlen.

All das ist Gegenstand der Mikroökonomie.

Entscheidendes Kennzeichen:

Untersuchung von Individualdaten anhand von Methoden der deskriptiven und schließenden Statistik sowie zugrundeliegenden sozioökonomischen Modellen.

Anwendungsbereiche:

wirtschaftliche und soziale Problemstellungen bis hin zu Gebieten der Verkehrstechnik, der Umwelttechnik, der Pharmakologie und Medizin.

Zur Illustration:

Ein typischer (anonymisierter) Datensatz, bestehend aus einer simulierten Stichprobe von 100 Individuen (Personen). In der Praxis werden derartige Datensätze in einer Umfrage erhoben, oder sie entstammen den records eines Bank- bzw. Fondsinstituts.

Gegenstand: Wahl einer Eigentumswohnung bzw. eines dafür aufgenommenen Kredits, durch Personen gegebenen Alters und liquiden Vermögensbestandes (ihrer ersparten Eigenmittel).

Datensatz hergestellt aus Zufallszahlen

Alter: Gleichverteilung

Vermögen: lognormal-Verteilung mit Alterstrend, dh.  $\log(\text{Eigspar ohne Trend})$  ist normalverteilt

Explorative Analyse: Ausgangspunkt für jeden Modellansatz.

Die Histogramme zeigen ungleiche Ersparnisse mit beträchtlicher Schiefe.

Scatter-Diagramm informativ !

Streuung der Vermögen steigt mit fortschreitendem Alter (bis auf Zufallsausreißer)

Basisformeln (für  $N=100$  Beobachtungen)

Es seien  $X_n$  und  $Y_n$  Querschnittsreihen von Beobachtungen.

Mittelwert

$$\bar{X} = \frac{1}{N} \sum_{n=1}^N X_n$$

korrigierte (Stichproben-)Varianz samt Standardabweichung  $\bar{S}$

$$\text{var}(X) = \bar{S}^2 = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2$$

Kovarianz zweier Reihen

$$s_{XY} = \text{cov}(X, Y) = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})(Y_n - \bar{Y})$$

Daraus Kovarianzmatrix und Korrelationsmatrix

$$\Sigma = \text{cov}(X, Y) = \begin{bmatrix} s_X^2 & s_{XY} \\ s_{XY} & s_Y^2 \end{bmatrix}, \quad R = \text{corr}(X, Y) = \begin{bmatrix} 1 & \frac{s_{XY}}{s_X s_Y} \\ \frac{s_{XY}}{s_X s_Y} & 1 \end{bmatrix}$$

im Beispiel haben wir samt der Inversen  $\Sigma^{-1}$

$$\Sigma = \begin{bmatrix} 60.8 & 137.4 \\ 137.4 & 5659.8 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0.23 \\ 0.23 & 1 \end{bmatrix}, \quad \Sigma^{-1} = \frac{1}{1000} \begin{bmatrix} 17.41 & -0.42 \\ -0.42 & 0.19 \end{bmatrix}$$

Die Variablen ALTER und EIGSPAR sind wie zu erwarten positiv korreliert.

**Datenblatt Wohnungswahl**  
**100 Personen von 24 bis 50 Jahren**

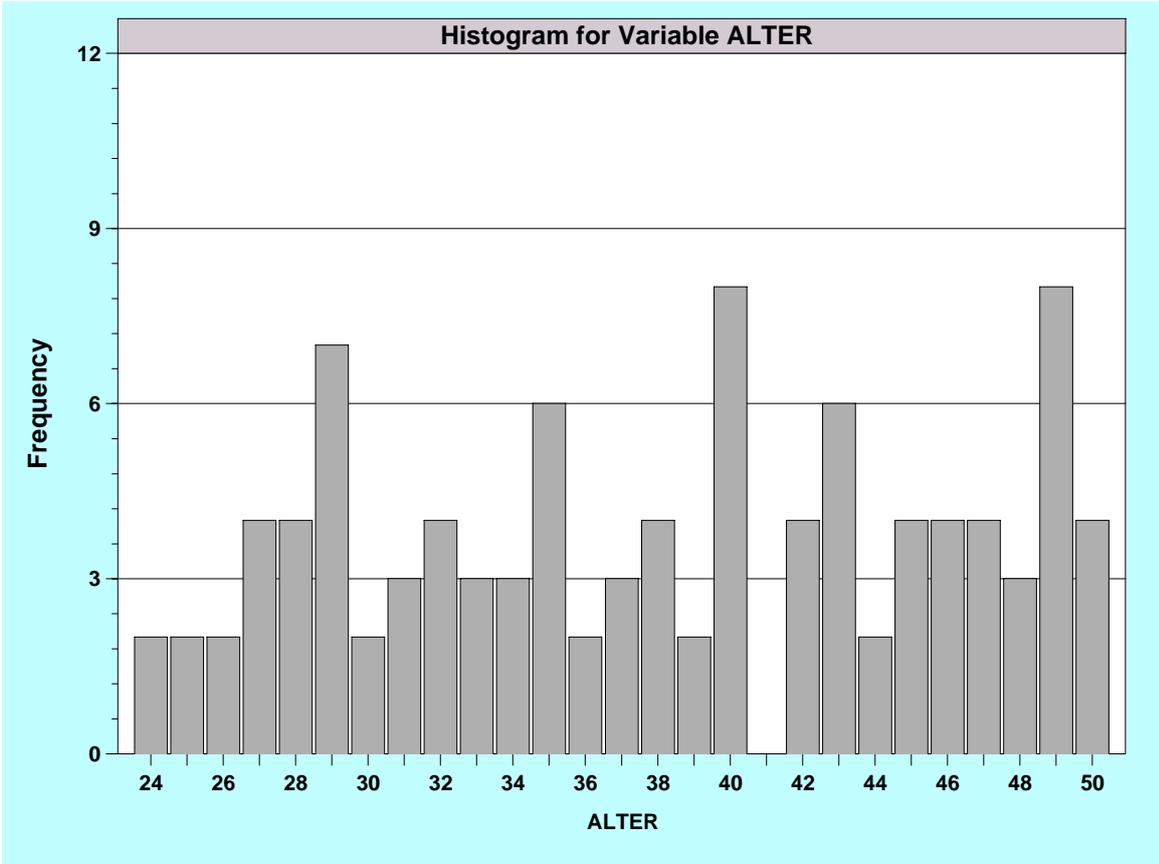
LNR	WAHL	ALTER	EIGSPAR
1	1	24	40,9
2	0	24	16,9
3	0	25	26,7
4	1	25	39,6
5	1	26	40,3
6	0	26	34,3
7	0	27	13,9
8	0	27	26,1
9	0	27	8,1
10	0	27	44,9
11	0	28	14,5
12	1	28	31,1
13	1	28	87,9
14	1	28	73,3
15	0	29	9,9
.	.	.	.
.	.	.	.
.	.	.	.
92	0	49	77,3
93	1	49	128,3
94	1	49	79,0
95	1	49	71,2
96	0	49	48,5
97	0	50	13,5
98	1	50	60,7
99	1	50	131,9
100	1	50	144,3

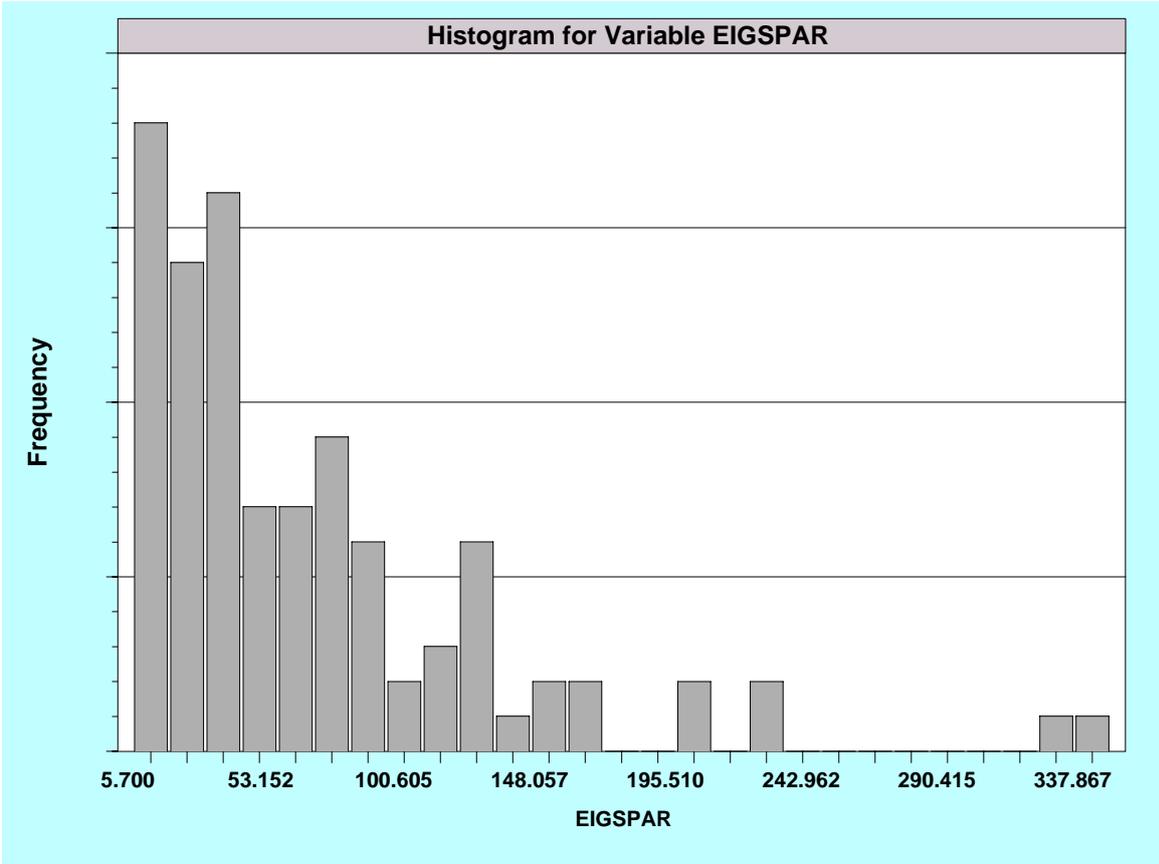
WAHL= 1: Wohnung (Haus) wird gekauft

mittleres Alter: 38 Jahre

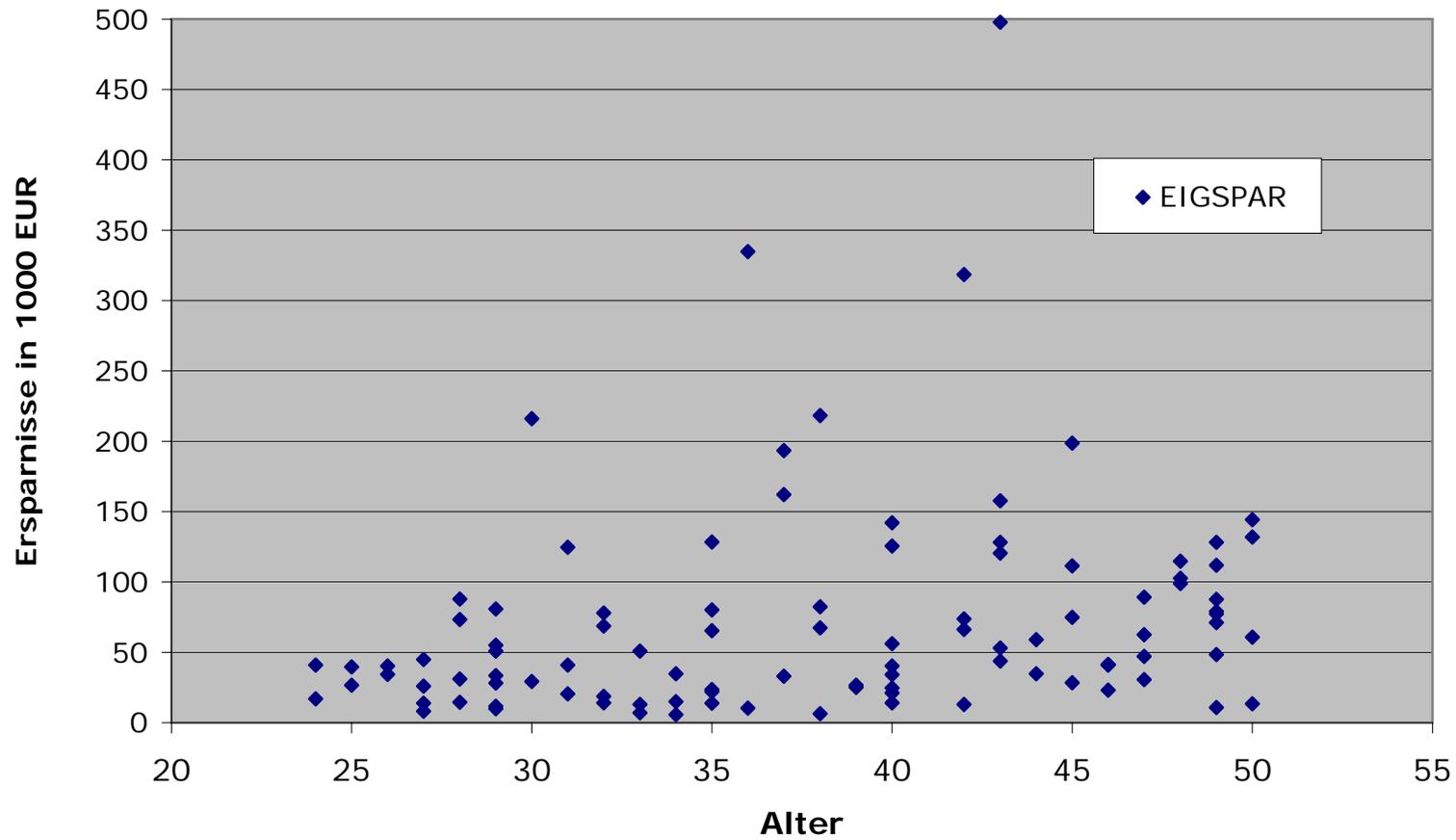
mittlere Ersparnisse: 71 Tsd. Euro

Median der Ersparnisse: 54 Tsd. Euro

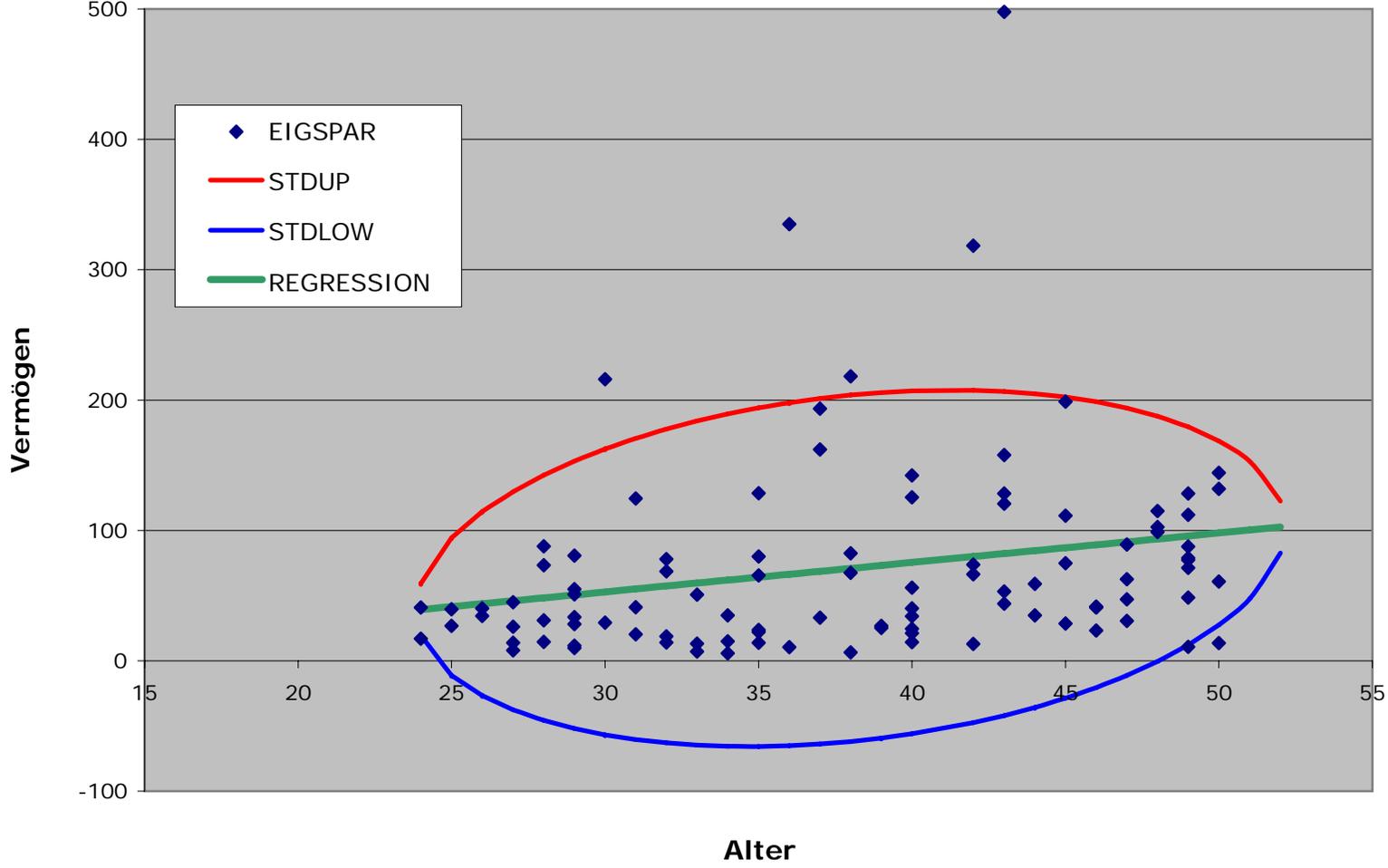




**Scatter Diagramm Ersparnisse**  
**Altersgruppen 24-50**  
**Mittelwert: Alter=38, Eigspar=71**  
**(simulierte Daten)**



# Streuungsellipse und Regressionsgerade



Wir bezeichnen bei gegebenem  $X, Y$  die Abweichung vom jeweiligen Mittelwert mit den Variablen  $x = X - \bar{X}$  und  $y = Y - \bar{Y}$ . Die Elemente der Inversen  $\Sigma^{-1}$  benennen wir mit  $v$ . Dann liefert die quadratische Form

$$\begin{bmatrix} x & y \end{bmatrix} \Sigma^{-1} \begin{bmatrix} x \\ y \end{bmatrix} = v_{XX} \cdot x^2 + 2v_{XY} \cdot xy + v_{YY} \cdot y^2 = \text{const}$$

eine Ellipse. In der Abbildung ist const. so gewählt, daß die Ellipse die Daten etwa überdeckt.

An der Lage dieser "Streuungsellipse" ist nun gut zu erkennen, wie die Variablen ALTER und EIGSPAR streuen. Zudem ist die Lage der Ellipse leicht schräg nach oben rechts. Höheres Alter erlaubt mehr Ersparnis. Dieses simple Faktum wird durch eine nachträgliche Regression bestätigt. Die grün eingezeichnete Gerade folgt aus der Kleinstquadrate-Schätzung OLS. Besser noch, aber nicht leicht visuell erkennbar, ist die OLS-Gleichung in Logarithmen:

$$EIGSPAR = -14.95 + 2.26 * ALTER \quad (\text{t-Wert } 2.4, R^2 = 0,05)$$

$$\log(EIGSPAR) = 2.33 + 0.04 * ALTER \quad (\text{t-Wert } 3.4, R^2 = 0.11)$$

Gemäß t-Werten nehmen die Ersparnisse mit dem Alter "signifikant" zu. Im Durchschnitt spart die Population 2260 Euro pro Jahr, bzw. vermehrt sie ihre Ersparnisse um 4% p.a.. Trotzdem ist der am  $R^2$  gemessene Fit gering. Das ist eine für mikroökonomische Untersuchungen typische Situation  $\implies$  die Individuen unterscheiden sich stark in ihrem Sparverhalten.

Bemerkung: die lange Ellipsenachse ist mit der Regressionsgeraden nicht deckungsgleich.

Zum Vorlesungsplan:

## Mikroökonomie (WS)

Einführungsvorlesung für Studenten des Bakkalaureats Mathematik, und für nachfolgendes Diplomstudium. Die Themen lauten:

(1) deskriptive Statistiken zur Durchforstung eines „Datenkaleidoskops“, Kontingenztafeln, Principal Component Analysis, Count data Modelle; wie behandelt man Stichprobenmerkmale, die qualitativer Natur sind oder in Intervallen gegeben sind;

(2) Methoden der qualitativen Wahl zwischen diskreten Alternativen, „Qualitative Response“: mit welcher Wahrscheinlichkeit Individuen eine bestimmte Wahl treffen und welche Faktoren diese Wahl begünstigen oder beeinträchtigen;

(3) Einführung in die „Survival Analyse“: wie lange ein Individuum in einem bestimmten Zustand verharrt, und welche Faktoren das Umschwenken in einen anderen Zustand beschleunigen oder verzögern.

## Methoden der Mikroökonomie (WS)

Vertiefende Vorlesung für Diplomstudenten.

Datengenerierender Prozesse: Man stelle sich das Entstehen von Daten als "Mechanismus" vor, der bestimmten Zufallsgesetzen gehorcht und in den Beobachtungen einer Stichprobe sichtbar wird.

Es werden stochastische Prozesse diskutiert, welche die Daten erzeugen, und Methoden, die bei der Analyse solcher Daten eingesetzt werden

- Random Walks
- Brown'sche Prozesse
- Ornstein-Uhlenbeck Prozesse (mean reverting processes)
- Erneuerungsprozesse
- Poisson Prozesse

## Angewandte Mikroökonomie (SS)

Anwendungsorientierte Vorlesung für Diplomstudenten (in Zukunft eventuell Praktikum).

Diskussion repräsentativer Modelle aus den Gebieten der Wirtschafts- und Marktforschung.

In Übungen werden Fragestellungen anhand von konkreten Umfragedaten erörtert.

Aufbauend auf Grundvorlesung gilt besonders Augenmerk den folgenden Themen:

- (1) Choice based sampling und Self-Selection: lassen sich Aussagen aus einer Teilstichprobe auf die Grundgesamtheit ausdehnen, für Marktanalysen besonders wichtig
- (2) Switching Regression samt Tobit (Teilstichproben weisen unterschiedliche lineare Modelle auf, samt Schätzung unbekannter Schwellenwerte, ab denen eine Aktion beobachtbar, zB Wohnungskauf ab Mindesteinkommen).
- (3) Grundzüge der Panelschätzung (betrifft Individualdaten bestehend aus einer zeitlichen Serie von Beobachtungen pro Individuum)

Literatur: findet sich in den Vorlesungsskripten des Vortragenden.

Kurzer Blick auf Modellprobleme (Gegenstand der Vorlesung Mikroökometrie)

Discrete Choice:

Konkrete Problemstellung mit Blick auf simulierte Daten von vorhin: sind die Individuen beim Hauserwerb erfolgreich?

Modellstruktur: Akteure (Personen)  $i = 1, 2, \dots, N$ ,

Wahlalternativen (binomial)  $Y_i = 0, Y_i = 1$

beobachtbare Einflußfaktoren  $X_i = \{x_{i1}, x_{i2}, \dots, x_{iK}\}$

Gesucht ist die **bedingte** Wahrscheinlichkeit  $P\{Y = 1|X\}$

Dann lautet Modell

$$P\{Y_i = 1\} = F(\mathbf{x}'_i\boldsymbol{\beta})$$

mit  $F$  als Verteilungsfunktion der Normalverteilung, und

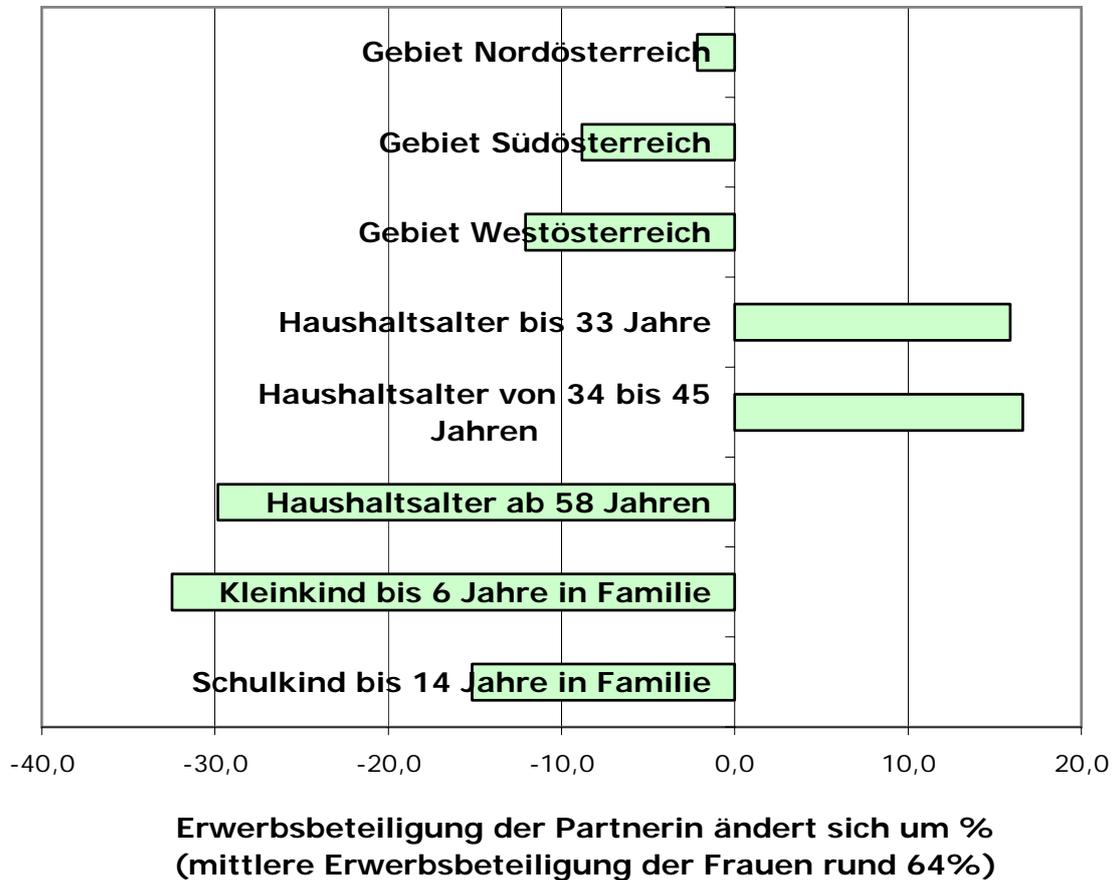
$$\mathbf{x}'_i\boldsymbol{\beta} = \sum_k x_{ik}\beta_k$$

Modellidee: systematischer Einfluß via  $\mathbf{x}'_i\boldsymbol{\beta}$  wird durch individuelle Präferenzen (normalverteilte Abweichungen) übersteuert, daher bei gegebenen Charakteristiken nur Wahrscheinlichkeit der Erwerbsbeteiligung berechenbar.

Abbildung zeigt empirisches Modell der Erwerbsbeteiligung von Frauen in Partnerschaft auf Basis von Mikrozensusdaten

## Simulation Erwerbsbeteiligung, Demographie

Referenzhaushalt: Partnerschaft, mittlerer Beruf, Alter 45-56, Inländerin, ohne Kinder, Dienstleistungssektor. Quelle: österreichische Mikrozensus



Wenn Charakteristiken miteinander korrelieren, dann ist es möglich, daß die Wahrscheinlichkeit (Charakteristik trifft zu) und deren Alternative (Charakteristik trifft nicht zu) in bestimmten Fällen die Randwahrscheinlichkeit nicht einschließt → typisches Paradox bei Umfragedaten.

## Duration analysis (Verweildauern)

Zwei Zustände der Welt, Ausgangszustand  $E_0$ , Alternativzustand (Endzustand)  $E_1$ .

Wie lange dauert es, bis ein Akteur vom Ausgangszustand in den Alternativzustand übergeht ?

(Problemstellung mit Blick auf simuliertes Beispiel: wie lange sparen die Individuen?)

Akteure (Personen)  $i = 1, 2, \dots, N$

Zufallsvariable  $T$ : Übergang des Akteurs  $i$  zum Zeitpunkt  $T_i$ .

beobachtbare Einflußfaktoren  $X_i = \{x_{i1}, x_{i2}, \dots, x_{iK}\}$

Modellstruktur:

Gesucht ist die bedingte Übergangswahrscheinlichkeit  $F(t) = P\{T < t|X\}$

instruktiver formuliert als Survival-Funktion

$$S(t) = P\{T \geq t|X\} = 1 - F(t) = 1 - P\{T < t|X\}$$

Kandidaten der Wahrscheinlichkeitsverteilung sind Exponential, Weibull oder andere.

Abbildung zeigt empirisches Modell des Firmenüberlebens mit Schätzung der Dauer bis zur eventuellen Insolvenz. Mehr als zwei Drittel der Firmen überleben nachhaltig, Insolvenzen sind im 4. Jahr am häufigsten.

### Stylised survivals and hazards of loan guarantees

Source: FGG-records 1982-2000, observed over guarantee term

