

# Kapitel 1: Ökonometrie und Ökonomische Daten



Moodle



Lehrbuch

# Übersicht:

Was ist Ökonometrie?

Schritte einer empirischen ökonomischen Analyse

Die Struktur ökonomischer Daten

Kausalität und Ceteris Paribus

# Was ist Ökonometrie?

# Was ist Ökonometrie?

Ökonometrie ist der Gebrauch statistischer Methoden zur Analyse ökonomischer Daten.

Ökonometriker:innen analysieren typischerweise nicht-experimentelle Daten.

## **Typische Ziele**

- ▶ Schätzung ökonomischer Zusammenhänge
- ▶ Testung ökonomischer Theorien
- ▶ Bewertung wirtschaftspolitischer Maßnahmen

## **Häufige Anwendungen**

- ▶ Prognose wichtiger makroökonomischer Variablen
- ▶ Prognose von Zeitreihen

# Schritte einer empirischen ökonomischen Analyse

# Schritte einer empirischen ökonomischen Analyse

Eine empirische Analyse benutzt Daten um eine Theorie zu testen oder um einen Zusammenhang zu schätzen.

## 1. Fragestellung formulieren!

- ▶ Welches sind die Einflussfaktoren für kriminelle Aktivitäten?
- ▶ Was sind die Effekte von Ausbildung auf Löhne?

## 2. Ökonomisches Modell aufstellen!

- ▶ Mikro- / Makromodell
- ▶ Optimierungs- und Gleichgewichtsannahmen
- ▶ Zusammenhänge zwischen ökonomischen Variablen
- ▶ Beispiele: Nachfragefunktionen, Preis-Schemata

# Ein ökonomisches Modell der Kriminalität (Becker 1968)

Gleichung, welche die Zeit, die mit kriminellen Aktivitäten verbracht wird als Funktion verschiedener Faktoren darstellt

$$y = f(x_1, x_2, x_3, x_4, x_5, x_6, x_7) ,$$

wobei

- $y$  : Zeit, die mit kriminellen Aktivitäten verbracht wird
- $x_1$  : Stundenlohn für kriminelle Aktivität
- $x_2$  : Stundenlohn für legale Arbeit
- $x_3$  : Einkommen, welches unabhängig von Arbeit und Kriminalität ist
- $x_4$  : Wahrscheinlichkeit erwischt zu werden
- $x_5$  : Wahrscheinlichkeit verurteilt zu werden, falls erwischt
- $x_6$  : Erwartete Strafe, falls verurteilt
- $x_7$  : Alter

# Ein Modell zu Ausbildung und Entlohnung

Modell:

$$wage = f(educ, exper, training) ,$$

wobei

<i>wage</i>	:	Stundenlohn
<i>educ</i>	:	Jahre Bildung
<i>exper</i>	:	Jahre Arbeitserfahrung
<i>training</i>	:	Wochen Fortbildungen

## 3. Ökonometrisches Modell aufstellen!

- ▶ Spezifizierung funktionaler Zusammenhänge durch parametrische Darstellung

→ unbekannte Parameter  $\beta_0, \beta_1, \beta_2, \dots$

Beschreiben Richtung und Intensität des Zusammenhangs ökonomischer Variablen

- ▶ Bereitstellung von Indikatoren für unbeobachtbare Variablen

# Ein ökonometrisches Modell der Kriminalität

$$\begin{aligned} crime &= \beta_0 + \beta_1 wage + \beta_2 othinc + \beta_3 freqarr \\ &\quad + \beta_4 freqconv + \beta_5 avgsen + \beta_6 age + u, \end{aligned}$$

wobei

- crime* : Maß für die Häufigkeit krimineller Aktivität
- wage* : Stundenlohn für legale Arbeit
- othinc* : Erwerbsunabhängiges Einkommen
- freqarr* : Häufigkeit Inhaftierungen für Straftaten
- freqconv* : Häufigkeit Verurteilungen für Straftaten
- avgsen* : durchschnittliche Haftlänge nach Verurteilungen
- u* : unbeobachtbare Faktoren

# Ökonometrisches Modell: Ausbildung und Entlohnung

$$wage = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 training + u ,$$

wobei

*wage* : Stundenlohn

*educ* : Jahre Bildung

*exper* : Jahre Arbeitserfahrung

*training* : Wochen Fortbildungen

*u* : unbeobachtbare Faktoren

# Schritte einer empirischen ökonomischen Analyse

## 4. Hypothesen aufstellen!

Hypothesen beinhalten die Vorzeichen der unbekannt Parameter.

Beispiel:

„Der Stundenlohn für legale Arbeit hat keinen Einfluss auf die Häufigkeit krimineller Aktivitäten.“

$$\Rightarrow H_0 : \beta_1 = 0.$$

**5. Daten bereitstellen!** → nächster Abschnitt

# Die Struktur ökonomischer Daten

- ▶ Querschnittsdaten
- ▶ Zeitreihendaten
- ▶ (Gepoolte) Querschnittsreihen
- ▶ Panel oder Längsschnittdaten

## Querschnittsdaten / Cross-Sectional Data

Eine Stichprobe von **Querschnittsdaten** besteht aus zeitgleichen Beobachtungen von verschiedenen Einheiten (Individuen, Firmen, Länder, ...).

- ▶ Zeitgleich kann auch zulassen, dass die Daten aus dem selben Jahr stammen.
- ▶ Querschnittsdaten sind (mehr oder weniger) unabhängig.
- ▶ Uniform-zufällige Auswahl von individuellen Einheiten.
- ▶ Probleme durch Selbstselektion
- ▶ Anwendungsschwerpunkt in mikroökonomischen Fragestellungen

## Querschnittsdaten: Löhne & individuelle Kenngrößen

obsno	wage	educ	exper	female	married
1	3.10	11	2	1	0
2	3.24	12	22	1	1
3	3.00	11	2	0	0
4	6.00	8	44	0	1
5	5.30	12	7	0	1
⋮	⋮	⋮	⋮	⋮	⋮
525	11.56	16	5	0	1
526	3.50	14	5	1	0

Die Daten stehen als Ressource zum Lehrbuch [hier](#) bereit.

Im Archiv: wage1.xls

## Querschnittsdaten: Wachstumsraten 1960-1985

De Long & Summers "Equipment investment and economic growth" *The Quarterly Journal of Economics* 106.2 (1991): 445-502.

obsno	country	gpcrgdp	govcons60	second60
1	Argentina	0.89	9	32
2	Austria	3.32	16	50
3	Belgium	2.56	13	69
4	Bolivia	1.24	18	12
⋮	⋮	⋮	⋮	⋮
61	Zimbabwe	2.30	17	6

*gpcrgdp* : average growth in real per capita GDP '60-'85

*govcon60* : government consumption as a percentage of GDP

*second60* : percentage of adult population with a secondary education

Hier stammen die Daten *govcon60* und *second60* aus dem Jahr 1960 und *gpcrgdp* aus den Jahren 1960 bis 1985.

# Zeitreihendaten / Time Series Data

Eine Stichprobe von **Zeitreihendaten** besteht aus Beobachtungen von einer Größe (Aktienkurs, Inflation, BIP, ...) zu verschiedenen Zeitpunkten.

- ▶ Typischerweise: korrelierte Daten
- ▶ Wichtig: Reihenfolge der Daten
- ▶ Tägliche, wöchentliche, monatliche, jährliche Daten möglich
- ▶ Trend- und saisonale Eigenschaften
- ▶ Anwendungen in der Makrökonomie und Finanzwissenschaften

# Zeitreihendaten: Mindestlohn und Arbeitslosigkeit

Castillo-Freeman & Freeman "When the Minimum Wage Really Bites: The Effect of the U.S.-Level Minimum on Puerto Rico" in G. Borjas and R. Freeman, eds., *Immigration and the Work Force*, Chicago: University of Chicago Press, 1992, pp. 177-212.

obsno	year	avgmin	avgcov	prunemp	prgnp
1	1950	0.20	20.1	15.4	878.7
2	1951	0.21	20.7	16.0	925.0
3	1952	0.23	22.6	14.8	1015.9
⋮	⋮	⋮	⋮	⋮	⋮
37	1986	3.35	58.1	18.9	4281.6
38	1987	3.35	58.2	16.8	4496.7

*avgmin* : Mindestlohn des Jahres

*avgcov* : Anteil der Arbeitskräfte mit Mindestlohn

*prunemp* : Arbeitslosenrate

*prgnp* : Bruttoinlandsprodukt (in Mio 1954\$)

## (Gepoolte) Querschnittsreihen / Pooled Cross Sections

**(Gepoolte) Querschnittsreihen** umfassen zwei oder mehr zufällig ausgewählte Querschnitts-Stichproben zu verschiedenen Zeitpunkten.

Diese Datensätze vereinen also die Eigenschaften von Querschnittsdaten und Zeitreihendaten.

- ▶ Die Querschnitts-Stichproben müssen unabhängig gezogen werden!
- ▶ Effektiv um die Wirkungen politischer Maßnahmen zu messen

# (Gepoolte) Querschnittsreihen / Pooled Cross Sections

Beispiel: Immobilienpreise vor und nach der Reduktion der Grundsteuer 1994

obsno	year	hprice	proptax	sqrft	bdrms	bthrms
1	1993	85,500	42	1600	3	2.0
2	1993	67,300	36	1440	3	2.5
3	1993	134,000	38	2000	4	2.5
⋮	⋮	⋮	⋮	⋮	⋮	⋮
250	1993	243,600	41	2600	4	3.0
251	1995	65,000	16	1250	2	1.0
252	1995	182,400	20	2200	4	2.0
253	1995	97,500	15	1540	3	2.0
⋮	⋮	⋮	⋮	⋮	⋮	⋮

# Panel oder Längsschnittdaten / Panel or Longitudinal Data

**Paneldaten** bestehen aus einer Zeitreihe für mehrere, aber gleichbleibende Querschnittsdaten.

- ▶ Paneldaten haben eine Querschnitts- und eine Zeitreihen-Dimension.
- ▶ Hauptunterschied zu gepoolten Querschnittsreihen: die einzelnen Einheiten der Querschnitte werden nicht zufällig gezogen, sondern im Laufe der Zeit „verfolgt“.
- ▶ Erfassung zeitunabhängiger, aber unbeobachtbarer Größen möglich.
- ▶ Erfassung zeitversetzter Zusammenhänge möglich.

# Panel oder Längsschnittdaten

Zweijahres Panel für urbane Kriminalität

obsno	city	year	murders	population	unem	police
1	1	1986	5	350,000	8.7	440
2	1	1990	8	359,200	7.2	471
3	2	1986	2	64,300	5.4	75
4	2	1990	1	65,100	5.5	75
⋮	⋮	⋮	⋮	⋮	⋮	⋮
297	149	1986	10	260,700	9.6	286
298	149	1990	6	245,000	9.8	334
299	150	1986	25	543,000	4.3	520
300	150	1990	32	546,200	5.2	493

# Datenstrukturen in dieser Vorlesung

Im ersten Teil der Vorlesung behandeln wir Querschnittsdaten, da hier die wichtige Annahme der Unabhängigkeit eher gewährleistet ist.

Erst im zweiten Teil der Vorlesung besprechen wir Zeitreihendaten, da diese den Gebrauch anspruchsvollerer Methoden voraussetzen.

# Kausalität und Ceteris Paribus

# Kausalität und Ceteris Paribus

Ziel der Ökonometrie:

Schlussfolgern, dass eine Variable (wie z.B. Bildung) eine **kausale Wirkung** auf eine andere Variable (wie z.B. Gehalt) hat.

Die kausale Wirkung ist vom Begriff der statistischen Korrelation abzugrenzen!

Ceteris (lat.): „der Rest“

Paribus (lat.): „gleich“

Die Annahme **Ceteris Paribus** bedeutet, dass man die Veränderung einer einzelnen Variable betrachtet und alle anderen Größen konstant bleiben.

Ohne Ceteris Paribus ist die Schlussfolgerung einer kausalen Wirkung sehr diffizil.

# Experimente

In wissenschaftlichen (Labor-)Experimenten kann die Ceteris Paribus-Annahme leicht aufrechterhalten werden.

In realen ökonomischen Situationen sind Ökonometriker:innen lediglich Beobachter:innen. Einzelne Faktoren können nur schwer isoliert werden.

Dennoch ist die Fragestellung hilfreich:

Wie sollte ein Experiment gestaltet sein, um einen einzelnen kausalen Effekt zu isolieren?

Hierzu einige Beispiele:

- ▶ Die Wirkung von Dünger auf die Ernte
- ▶ Die Wirkung von Bildung auf das Gehalt
- ▶ Die Wirkung von Polizei auf urbane Kriminalität
- ▶ Die Wirkung von Mindestlöhnen auf die Arbeitslosenquote

# Die Wirkung von Dünger auf die Ernte

„Wie stark wird die Produktion von Sojabohnen ansteigen, wenn die aufgetragene Düngermenge angehoben wird?“

Auch andere Faktoren wie Niederschlag, Bodenqualität und Parasiten beeinflussen den Ernteerfolg.

Experiment:

Trage auf verschieden Äckern gleicher Größe unterschiedliche Mengen von Dünger auf. Messe die erzielten Ernteerträge.

→ Querschnittsdaten!

Problem: Manche Einflussfaktoren sind nicht beobachtbar.

Lösung:

Variiere die Düngemengen **unabhängig von anderen Einflussfaktoren** auf den verschiedenen Feldern!

# Die Wirkung von Bildung auf das Gehalt

„Wenn eine zufällig ausgewählte Person ein zusätzliches Jahr Bildung absolviert, um wieviel steigt ihr Gehalt?“

Gedankenexperiment: Was wäre, wenn ein:e soziale Planer:in zufällig ausgewählten Personen ein Bildungsniveau zuweisen könnte?

In diesem Fall könnte das Dünger-Experiment zu einem Bildungs-Experiment transformiert werden.

Aufgrund von psychischen, ethischen und monetären Kosten ist dieses Gedankenexperiment nicht real durchführbar.

Problem ohne zufällige Zuweisung von Bildung: Die Menschen wählen Bildungsniveau **in Abhängigkeit** von anderen wichtigen aber potentiell unbeobachtbaren Einflussfaktoren wie intellektueller Fähigkeit, Motivation, Elternhaus, gesellschaftlichen Einflüssen.

# Die Wirkung von Polizei auf urbane Kriminalität

„Bewirkt die Präsenz von mehr Polizist:innen eine Reduktion von Kriminalität?“

Ceteris Paribus-Frage:

Wie stark würde die Kriminalitätsrate sinken, wenn eine zufällig ausgewählte Stadt 10 zusätzliche Polizist:innen erhielte?

Kontrafaktische Frage:

Wie hätte sich die Kriminalitätsrate in dieser Stadt ohne die zusätzlichen Polizist:innen entwickelt?

Probleme:

Es gibt keine zwei identischen Städte, die sich ausschließlich durch die Anzahl von Polizist:innen unterscheiden.

Ein:e soziale Planer:in kann den Städten die Anzahl der Polizist:innen nicht vorschreiben.

Die Entscheidung der einzelnen Stadt ist **abhängig** von unbeobachtbaren Einflussfaktoren der jeweiligen Kriminalitätsrate.

# Die Wirkung von Mindestlöhnen auf die Arbeitslosenquote

„Wie stark (wenn überhaupt) wird die Arbeitslosenquote ansteigen, wenn der Mindestlohn erhöht wird?“

Ein Beispiel für Zeitreihen zu Mindestlöhnen und Arbeitslosenquote ist auf Folie 19.

Die Gleichgewichts-Theorie prognostiziert einen Anstieg der Arbeitslosenquote, falls der Mindestlohn höher ist als der Gleichgewichtslohn. Das Ausmaß des Anstiegs hängt von der Steigung der Nachfragekurve (für Arbeit) ab.

Würde der Mindestlohn zufällig (und unabhängig) festgelegt, wäre der Effekt des Mindestlohnes messbar.

Wenn die Mindestlohnentscheidung jedoch mit unbeobachtbaren Größen **zusammenhängt**, die ebenfalls die Arbeitslosenquote beeinflussen, ist die Messbarkeit komplizierter.

# Begriffe aus Kapitel 1

- ▶ Empirische Analyse
- ▶ Ökonomisches & ökonometrisches Modell
- ▶ Querschnittsdaten
- ▶ Stichprobe / Zufallsauswahl
- ▶ Zeitreihen
- ▶ Gepoolte Querschnittsreihen
- ▶ Längsschnittdaten
- ▶ Kausaler Effekt, Ceteris Paribus
- ▶ Experimentelle und beobachtete Daten