

# Statistische Ansätze in der Wirtschafts- und Sozialgeschichte : eine Übersicht

Autor(en): **Baten, Jörg**

Objektyp: **Article**

Zeitschrift: **Geschichte und Informatik = Histoire et informatique**

Band (Jahr): **9 (1998)**

PDF erstellt am: **24.06.2024**

Persistenter Link: <https://doi.org/10.5169/seals-7227>

## **Nutzungsbedingungen**

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern. Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden. Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

## **Haftungsausschluss**

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

# Statistische Ansätze in der Wirtschafts- und Sozialgeschichte: Eine Übersicht\*

---

Jörg Baten

## Zusammenfassung

Wünschen «Geschichtskonsumenten» Statistiken in den Aufsätzen und Büchern, die wir schreiben? Welche Leistungen können statistische Methoden im geschichtswissenschaftlichen Bereich überhaupt erbringen, und zu welchem Preis? Dieser Beitrag gibt einen Überblick zum Statistikeinsatz in der Wirtschafts- und Sozialgeschichte in verschiedenen Ländern und fragt nach den Ursachen für die enormen Unterschiede in diesem Bereich.

Ein zweiter Teil formuliert Strategien, wie die «Kosten» in Form von Arbeitszeit reduziert werden können. Es werden einige Verfahren vorgestellt, die für historische Datenqualitäten besonders geeignet erscheinen, jedoch bisher in der Geschichtswissenschaft selten eingesetzt wurden (z.B. OLS-Regression, ordered probit regression, count data analysis).

## 1. Erkenntnismöglichkeiten der schließenden Statistik im historischen Bereich und ihre Probleme

Meine anfängliche Hypothese bei dieser kleinen Untersuchung zum länderspezifischen Statistikeinsatz in der Wirtschafts- und Sozialgeschichte bestand darin, daß Mitte der 1990er Jahren die Anwendung von Methoden der schließenden Statistik<sup>1</sup> in Europa eher die Ausnahme als die Regel ist, und daß dies ganz besonders auch für Deutschland gilt. Diese Anfangshypothese wird nachfolgend zu prüfen sein.<sup>2</sup>

Zwar gibt es auch in Europa an einigen Orten Wirtschafts- und Sozialhistoriker, die mit quantitativen Methoden arbeiten. Aber die «kritische Masse» scheint zu fehlen. Ein Indiz dafür besteht in der Tatsache, daß eine

---

\* Ich danke Peter Bär, Christophe Koller, John Komlos, Christian Pfister und den Teilnehmern der Tagung für wichtige Hinweise.

1 Schließende Statistik: Im Gegensatz zur deskriptiven Statistik, die zum Ziel der Beschreibung u.a. Mittelwerte und Variationsmaße benutzt, untersucht die schließende Statistik Einflüsse und Wechselwirkungen von mehreren Variablen (Beispiele finden sich im Text).

2 Eine Auswertung der Artikel im *Journal of Economic History* 1941-1990 über die «klimetrische Revolution» in den USA findet sich in Whaples, Robert: «A Quantitative History of the Journal of Economic History and the Cliometric Revolution». In: *Journal of Economic History* 51-2 (1991), S. 289-301.

intensive Diskussion z.B. auf den Historikerkongressen über die richtige Anwendung dieser Methoden selten oder nie geführt wird. Richard Tilly merkte kürzlich an, daß im deutschsprachigen Bereich für einen Nachwuchswissenschaftler die Anreize aus seiner Umwelt nicht groß seien, statistische Methoden anzuwenden.<sup>3</sup> Die Motivation für diese Art der Arbeit müsse vom Erkenntnisinteresse des Forschers selbst kommen. Zudem denke ich aber auch, daß diese Techniken die internationale Kommunikation erleichtern und dabei helfen, konkurrierende Hypothese zu überprüfen.

Im angloamerikanischen Raum hingegen hat sich die Kliometrie<sup>4</sup> vor allem im Teilbereich Wirtschaftsgeschichte, aber auch in der Sozialgeschichte rasant entwickelt. Scott Eddie argumentiert, daß dieser Aufschwung ganz wesentlich von der rapiden Expansion der amerikanischen Universitäten in den 1970er Jahren getragen wurde, die überproportional viele Wirtschaftswissenschaftler einstellten und unter diesen wiederum viele Wirtschaftshistoriker. Das Ziel war damals, den Studenten ein breiteres Angebot an Lehrveranstaltungen bieten und ökonomische Theorien verstärkt empirisch überprüfen zu können. In den 70er Jahren ging allerdings auch eine Welle der «Quantifizierung» durch die europäische Wirtschafts- und Sozialgeschichte, die mit den damals verfügbaren statistischen und informationstechnischen Möglichkeiten bereits interessante Ergebnisse erbrachte.<sup>5</sup> Diese Möglichkeiten haben sich inzwischen noch bedeutend erweitert, vor allem durch den Preisverfall leistungsfähiger Computer und der viel angenehmer gewordenen Software. Angesichts der Tendenzen zur Globalisierung und Ökonomisierung der Weltgesellschaft ist deutlich abzusehen, daß sich in den nächsten Jahren auch die Geschichtswissenschaft wieder sehr viel stärker auf quantitative, wirtschaftshistorische Methoden hin orientieren wird.

Natürlich gibt es auch an diesem Wissenschaftsstil ernsthafte Kritikpunkte. Neben den üblichen Problemen statistischer Methoden (Meßfehler

---

3 Tilly, Richard: «Cliometrics in Germany». In: Komlos, John; Eddie, Scott (Hgg.): *Selected Cliometric Studies on German Economic History*. Stuttgart 1997, S. 17-33; Eddie, Scott: «Cliometrics: What is it, whither came it forth?». In: ebda., S. 12-16; Komlos, John: «Die Zukunft der Wirtschaftsgeschichte». In: *VSWG* 82 (1995), S. 404-410.

4 Kliometrie: In den 1970er Jahren löste die Kliometrie in den USA die eher beschreibende, traditionelle Wirtschaftsgeschichte ab. Sie orientierte sich zunächst an ökonomischen, später auch an anderen sozialwissenschaftlichen Theorien und benutzt statistische Techniken zur Hypothesenprüfung.

5 Schon innerhalb der deutschen Wirtschafts- und Sozialgeschichte gäbe es hier so viele Namen zu nennen, daß ich auf eine ausführliche Bibliographie an diesem Punkt verzichte. Ohne Anspruch auf Vollständigkeit seien die Namen Werner Abelshäuser, Knut Borchardt, Peter Borscheid, Wilfried Feldenkirchen, Rainer Fremdling, Rainer Gömmel, Carl-L. Holtfrerich, Hartmut Kaelble, Hubert Kiesewetter, Jürgen Kocka, Rainer Metz, Dietmar Petzina, Toni Pierenkemper, Hans Pohl, Jürgen Schneider, Eckart Schremmer, Reinhard Spree, Richard Tilly, Rolf Walter, Hans-Ulrich Wehler und Harald Winkel exemplarisch genannt.

etc.) besteht das wichtigste Problem darin, daß sich die kometrische Wirtschaftsgeschichte von den nicht statistisch geschulten Lesern meist weit entfernt. Die narrativen Elemente, die einen historischen Text angenehm lesbar machen, treten in der Regel in den Hintergrund. Es bedeutet doppelten Aufwand, eine historische Darstellung zum einen «statistikorientiert» und zum andern narrativ zu gestalten.

Andererseits bemerkt man sogar in populärwissenschaftlichen Fernsehsendungen immer häufiger einfache Streudiagramme, die Zusammenhänge aufzeigen. Ich denke, daß auf lange Sicht die Möglichkeiten, Zusammenhänge zu erforschen und mit Hilfe statistischer Methoden zu überprüfen, so attraktiv sind, daß einige dieser Methoden zu Standardinstrumenten werden sollten. Zudem erscheint die Prognose wahrscheinlich, daß der Durchschnittsbürger in naher Zukunft schon in der Schule in einfachen Formen des «statistischen Denkens» ausgebildet werden wird, weil er in einer sonst immer unüberschaubarer werdenden, globalisierten Welt den Überblick verliert. Schon heute kann es sich durchaus sogar im privaten Bereich lohnen, einfache statistische Techniken z.B. beim Anmieten einer Wohnung zu benutzen. Das Angebot ist umfangreich und differenziert, und die Mietsumme ist über die Jahre hinweg hoch, so daß sich eine Stunde statistische Analyse hier leicht lohnen kann. Die unten näher besprochene OLS-Regression macht deutlich, wie unterschiedlich die Preise nach Stadtteilen, Wohnungsgröße etc. sind, und man kann die Resultate mit der eigenen Zahlungsbereitschaft für diese Kriterien vergleichen.

Unter dem oben angesprochenen «statistischen Denken» verstehe ich, daß beim Nachdenken über Zusammenhänge zufällige Fehler als solche gewertet werden. Ein Beispiel aus der Geschichtswissenschaft mag dies illustrieren: Ein Forscher denkt über den Zusammenhang von Proteinproduktion und Körpergröße nach. Die Hypothese sei, daß niedrige Werte lokaler Milchproduktion (gemessen an der Rinderzahl pro Kopf) hohe Untauglichenanteile wegen Körpergröße bei der Rekrutierung in den preussischen Regierungsbezirken verursacht haben (Abb. 1).<sup>6</sup> Er läßt seinen Computer ein Streudiagramm mit einer Regressionslinie ausdrucken. Ein zweiter Historiker tritt hinzu und zeigt auf die Punkte, die besonders weit von der Regressionslinie entfernt liegen. Er glaubt nicht an den Zusammenhang, weil es ja Gegenbeispiele gibt, z.B. hat der Regierungsbezirk Stralsund eine niedrigere Milchproduktion pro Kopf als der Bezirk Trier, wo die Untauglichenquote etwas höher liegt. Natürlich läßt sich ein solcher

---

6 Hierbei wurde angenommen, daß die regionalen Unterschiede der Milchproduktion sich innerhalb eines Jahrhunderts nur wenig ändern.

Zusammenhang nie im strikten Sinn <beweisen>, weil es immer Scheinkorrelationen geben kann. Ist jedoch ein Zusammenhang theoretisch sinnvoll zu begründen und wird er anhand einer Mindestanzahl von Beobachtungen bestätigt, so ist die Wahrscheinlichkeit recht hoch, daß dieser Zusammenhang existiert.

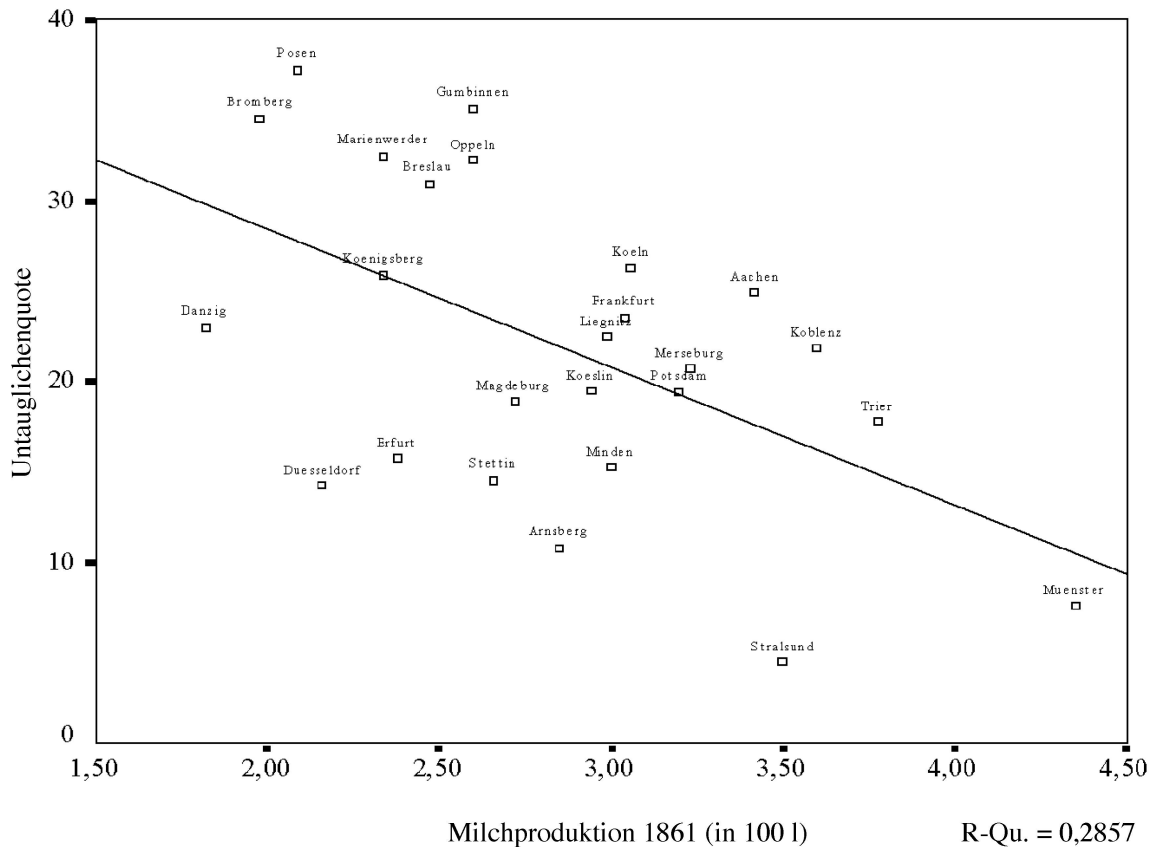


Abb. 1: Der Zusammenhang von Pro-Kopf-Milchproduktion (1861) und Untauglichenquote wegen ungenügender Körpergröße 1854

Ich denke, daß in der universitären Ausbildung die angewandte Statistik sehr wichtig ist, gerade auch weil viele Universitätsabsolventen später nicht an der Universität arbeiten. In vielen Bereichen (Unternehmensberatungen, Marktforschungsinstitute, staatliche Stellen, Raumplanung etc.) zeichnet sich m.E. eine Tendenz zur «Verwissenschaftlichung» und zum Einsatz quantitativer, statistischer Instrumente zur Entscheidungsfindung ab; wengleich es freilich retardierende Elemente geben mag, wie z.B. ältere Vorgesetzte, die solche Methoden nicht lieben.

Um zur Verbreitung statistischer Methoden beizutragen, kommt es ganz wesentlich darauf an, die <Kosten> in Form von Arbeitszeit zu verringern, etwa durch Kurse, die mit einem Minimum an mathematischen Vorkennt-

nissen auskommen.<sup>7</sup> Auch die Orientierung an Beispielen und der Einsatz neuer Lernmedien (Videos, computergestützte Lernprogramme etc.) birgt noch ein großes Potential, um den Aufwand weiter zu verringern. So könnte sich eine Geschichtswissenschaft herausbilden, die narrative und statistische Elemente kombiniert, eine Synthese der Stile von Theodor Mommsen, Douglass North und Robert Fogel sozusagen, des ersten und der beiden letzten Nobelpreisträger aus den historischen Wissenschaften.<sup>8</sup>

## 2. Ländervergleich beim Statistikeinsatz

Im folgenden will ich kurz einige Länder unter der Fragestellung vergleichen, inwieweit ihre Wirtschafts- und Sozialhistoriker schließende Statistik und Simulation einsetzen.

Für diesen Zweck habe ich eine kleine Datenbank aufgebaut, die aus 623 Zeitschriftenartikeln aus den Jahren 1994 bis 1996 besteht, soweit diese Jahrgänge bei uns der Bibliothek verfügbar waren. Ich habe diese Artikel auf ihren Statistikeinsatz hin untersucht, wobei einschränkend anzumerken ist, daß ich die einzelnen Artikel und insbesondere die darin enthaltenen Tabellen nur kurz angeschauen konnte. Denn bei einem Gesamtumfang von ca. 18'000 Seiten war eine allzu stark in Detail gehende Durchsicht der Artikel unmöglich. Ebenfalls mit der Begrenzung des zeitlichen Aufwandes ist zu erklären, daß die Mehrzahl der Zeitschriften aus dem wirtschaftshistorischen und weniger aus dem sozialhistorischen Bereich stammt.<sup>9</sup> Zudem war ich gezwungen, die Zahl der betrachteten Länder auf sechs zu beschränken. Weiterhin ist anzumerken, daß etwas <versteckt> angewendete statistische Verfahren unentdeckt geblieben sein könnten.

«Statistikeinsatz» ist natürlich ein umfassender Begriff, auch der Vergleich von zwei Mittelwerten ist ja bereits Statistik. Um die Fragestellung etwas einzuengen, habe ich mich in diesem Beitrag darauf konzentriert, den Einsatz von schließender Statistik und Simulation zu betrachten (Tab. 1), also insbesondere die Verfahren der Regressionsanalyse und Faktorenana-

---

7 Bisher gibt es noch wenig Einführungen, die auf das intuitive Verständnis abzielen. Die entsprechenden Programmhandbücher der populären Statistikprogramme, z.B. von SPSS für Windows oder Eviews, haben hier überraschende Leistungen vollbracht, wobei es diesen natürlich darauf ankommt, alle Programmfunktionen zu erwähnen, auch die weniger wichtigen. Peter Kennedy hat ein gutes Einführungsbuch geschrieben, das allerdings schon sehr umfassend ist, vgl. Kennedy, Peter: *A Guide to Econometrics*. Malden (MA) 1998.

8 Die ihre Auszeichnungen für ihre Leistungen in den Bereichen Literatur (Mommsen) und Wirtschaftswissenschaften erhielten.

9 Neben den unten ausgewerteten Zeitschriften hätten weitere Periodica mit Artikeln, die schließende Statistiken enthalten, ausgewertet werden können, z.B. das *Journal of Interdisciplinary History*, *Social Science History*, *Scripta Mercaturae* oder *Historical Methods*.

lyse, wobei erstere in der wirtschaftshistorischen Literatur klar dominiert. Hinzugezählt habe ich auch einfache Streudiagramme, soweit mit ihnen der Zusammenhang zwischen zwei Variablen untersucht wird, von den nicht eine die Zeitachse darstellt.

Insgesamt ist zu beachten, daß sich der Einsatz von schließender Statistik natürlich nicht bei jeder Fragestellung und bei jeder Art von Artikel eignet.<sup>10</sup> Aber der Prozentsatz ist doch ein Indikator, inwieweit diese Methodik ein normaler Bestandteil des wissenschaftlichen Diskurses ist.

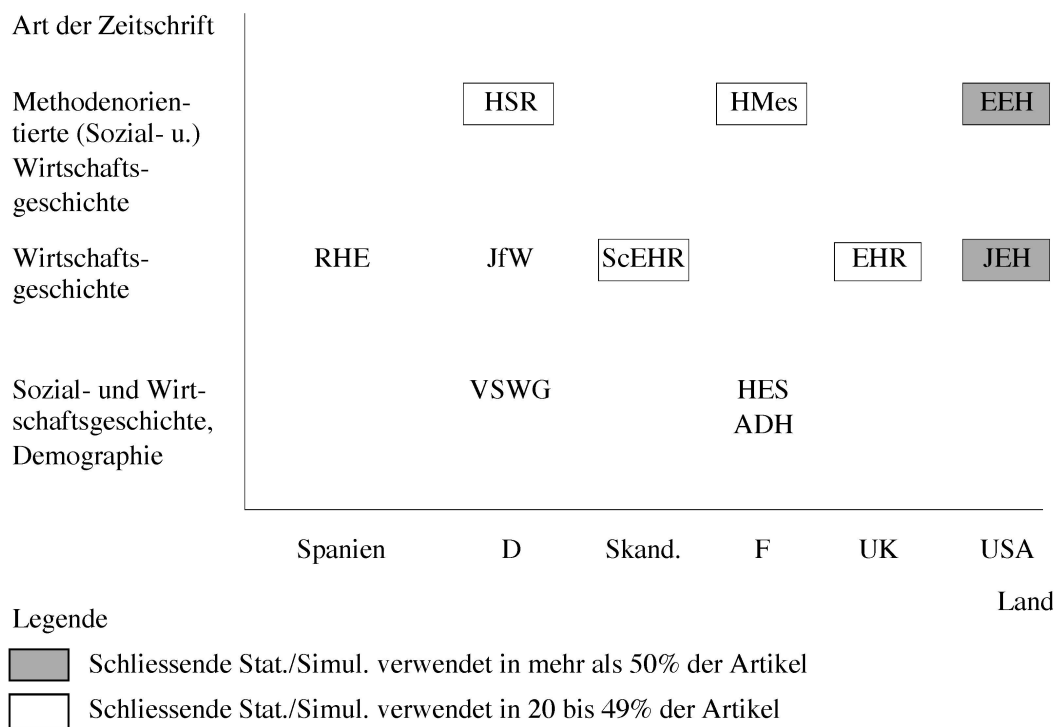


Abb. 2: Einige Charakteristika der betrachteten Zeitschriften

Einige der Zeitschriften sind explizit oder implizit (wie die «Explorations in Economic History», EEH) auf den Einsatz von statistischen Methoden hin orientiert (Abb. 2). Dies ist sicherlich auch bei der deutschen HSR der Fall, die thematisch stärker sozial- als wirtschaftshistorisch ausgerichtet ist, ebenso bei der französischen Zeitschrift «Histoire et Mesure». Andere Zeitschriften sind sozial- *und* wirtschaftsgeschichtlich oder auch demographiehistorisch ausgerichtet (wie die anderen beiden französischen Zeitschriften und die deutsche VSWG). Meine anfängliche Erwartung war, daß in wirtschaftshistorischen Zeitschriften häufiger statistische Methoden anzutreffen sind, weil Statistik meist Teil der ökonomischen, nicht aber der

<sup>10</sup> Ausgesprochene Literaturüberblicke wurden aus der Tabelle ausgeschlossen.

historischen Grundausbildung ist. Im Durchschnitt mag dies auch stimmen, es gibt jedoch Ausnahmen (z.B. das JfW und RHE, die seltener schließende Statistiken aufwiesen).

Land und Zeitschrift	Alle Jahre	1994	1995	1996	Absolute Anzahl
<b>USA</b>					
EEH	81%	59%	87%	95%	67
JEH	52%	35%	63%	58%	94
<b>UK</b>					
EHR	22%	19%	30%	18%	130
<b>Frkr.</b>					
HES	1%	-	-	-	111
ADH	19%	-	-	-	26
HMes	36%	-	-	-	11
<b>Skand.</b>					
ScEHR	29%	38%	18%	-	24
<b>Dtld.</b>					
JfW	11%	0%	12%	23%	45
VSWG	10%	10%	0%	20%	29
HSR	48%	17%	50%	56%	40
<b>Span.</b>					
RHE	11%	-	-	-	46
<b>Durchschnitt</b>					
USA-Zt.	67%	47%	75%	77%	
UK-Zt.	22%	19%	30%	18%	
Sonst. Europ. Zt.	21%	-	-	-	

Tab. 1: Der Prozentanteil der Zeitschriftenartikel, die schließende Statistik oder Simulation verwenden. Abkürzungen und überwiegende Sprache:

EEH	Explorations in Economic History (engl.)
JEH	Journal of Economic History (engl.)
EHR	Economic History Review (engl.)
HES	Histoire, économie, société (frz.), nicht jahrgangweise gezählt
ADH	Annales Demographie Historique (frz.), nicht jahrgangweise gezählt
HMes	Histoire et Mesure, nur 1996 berücksichtigt (frz./engl.)
ScEHR	Scandinavian Economic History Review, nur 1994 u. 1995-1 u. 3 berücksichtigt (engl.)
JfW	Jahrbuch für Wirtschaftsgeschichte (dt./engl.)
VSWG	Vierteljahrschrift für Sozial- und Wirtschaftsgeschichte (dt./engl.)
HSR	Historical Social Research (engl./dt.)
RHE	Revista de Historia Económica (span./engl.)

Die beiden Zeitschriften mit der deutlichsten Statistikorientierung sind also das JEH und EEH, die beide in den USA erscheinen. Gleichzeitig befinden sich darin einige Artikel, die nicht von Nordamerikanern stammten (Tab. 2). Die Anzahl dieser Artikel gibt einen Hinweis darauf, wie groß die Bereitschaft von Wirtschaftshistorikern der verschiedenen Länder ist, dort



zu publizieren (und wie erfolgreich sie dabei sind). Dabei ist anzumerken, daß der Refereeprozess die <Mühen> einer Publikation in einer anderen als der Muttersprache noch wesentlich erhöht. Nehmen wir einmal an, daß die Zahl der Wirtschaftshistoriker mit der Bevölkerungszahl korreliert, so deutet die Anzahl von 15 Artikel aus Israel (pro 10 Mio. d. Bev.) darauf hin, daß die dortigen Wirtschaftshistoriker besonders interessiert sind, in diesen beiden Zeitschriften zu veröffentlichen. Großbritannien und Australien folgen auf den nächsten Plätzen, was – ebenso wie bei Irland auf Platz 5 – wegen des <Native-Speaker>-Effekts nicht so überraschend ist. Bemerkenswert sind hingegen die prominenten Plätze von Norwegen, den Niederlanden und Schweden. Auch Frankreich und Spanien haben eine beträchtliche absolute Anzahl von Artikeln, während z.B. Italien und Deutschland eher weiter unten auf der Skala angesiedelt sind. Beiträge aus der Schweiz wurden in den untersuchten Jahrgängen des JEH und EEH keine gefunden.

Nationalität	Artikelzahl relativ zur Bevölkerung*	Artikelzahl absolut
Israel	15,0	6
Grossbritannien	6,3	36
Australien	6,3	10
Norwegen	5,0	2
Irland	5,0	2
Niederlande	3,3	5
Schweden	2,9	2
Frankreich	1,4	8
Spanien	1,3	5
Portugal	1,0	1
Griechenland	1,0	1
Italien	0,5	3
Deutschland	0,5	4
Korea	0,3	2
Argentinien	0,3	1
Türkei	0,2	1
Japan	0,2	2
<b>Summe</b>		<b>91</b>
<b>%-Anteil an allen Artikeln**</b>		<b>20.7%</b>

Tab. 2: Artikel im JEH und EEH, die nicht von Nordamerikanern stammen (1989-1996)

\* Pro 10 Mio. der Bevölkerung

\*\* Wobei auch Artikel den «Ausländern» zugerechnet wurden, wenn nur ein/e Autor/in Nicht-Nordamerikaner/in war

Quellen: Explorations in Economic History, Oktober 1989 - Oktober 1996 und Journal of Economic History, September 1989 - September 1996.

Eine hier interessante weitere «Stichprobe» besteht in den Teilnehmern von zwei internationalen wirtschaftsgeschichtlichen Konferenzen, die in den letzten beiden Jahren stattfanden und deren Tagungssprache nur Englisch war (Tab. 3).<sup>11</sup> Wiederum wurde die Zahl der Teilnehmer auf die Bevölkerung des jeweiligen Landes bezogen, dabei zeigen Dänemark und Israel die höchsten Teilnahmeziffern, ebenfalls Finnland, Irland und Norwegen. Nun ist natürlich zu beachten, daß die beiden Konferenzen in Deutschland und Italien stattfanden, daher differierten die Reisekosten der Wirtschaftshistoriker aus den unterschiedlichen Ländern sehr stark. In Spalte 4 wurde die Teilnehmerzahl (pro Bev.) unter Berücksichtigung der geschätzten Anreisekosten berechnet. Dies fängt natürlich nicht die gesamten Kosten und den erwarteten Nutzen ein, daher landen die Teilnehmer aus den USA trotz der höchsten absoluten Teilnehmerzahl im unteren Mittelfeld. Die Zahl der deutschen Teilnehmer war relativ gering, obwohl die eine der beiden etwa gleich großen Konferenzen in München stattfand. Insgesamt schneiden wiederum die nichtenglischsprachigen, skandinavischen Länder und die Niederlande recht gut ab. Zumal diese Beobachtung auch bei ähnlichen Untersuchungen gemacht wurde – z.B. wurde kürzlich die Anzahl der Veröffentlichungen in volkswirtschaftlichen Top-Journals gezählt – liegt der Schluß nahe, daß Niederländer und Skandinavier wesentlich «internationaler» orientiert sind als z.B. Deutsche oder Franzosen. Ein recht großer wissenschaftlicher «Binnenmarkt» für deutsch- bzw. französischsprachige Studien mag hier dämpfend wirken.<sup>12</sup>

Die Anfangshypothese, dass in Europa der Einsatz statistischer Methoden eher die Ausnahme als die Regel sei, hat sich daher teilweise am empirischen Material bestätigt, konnte jedoch für einige Länder modifiziert werden. Nachfolgend sollen einige Strategien genannt werden, die die «Kosten» des Statistikeinsatzes senken könnten.

---

11 Man hätte auch die *European Social Science History Conference* betrachten können, um stärker die Sozial- und Wirtschaftsgeschichte einzubeziehen, aber hierfür stand kein Teilnehmerverzeichnis zur Verfügung.

12 Zum Teil können auch institutionelle Beschränkungen eine Rolle spielen. Z.B. ist es bei einer wichtigen deutschen sozial- und wirtschaftsgeschichtlichen Zeitschrift untersagt, als Deutscher englischsprachige Aufsätze zu veröffentlichen.

Land	Teilnehmer/Bev.*	Anzahl Teilnehmer, absolut	Kontrolliert für Reisekosten**
Israel	75	3	100
Dänemark	80	4	98
Finnland	60	3	74
Irland	50	2	60
Norwegen	50	2	60
Großbritannien	47	27	56
Australien	25	4	51
Schweden	43	3	51
Italien	43	25	48
Niederlande	33	5	38
Spanien	32	12	37
Belgien	30	3	35
Argentinien	3	1	34
Japan	2	2	33
Korea	2	1	33
Brasilien	1	1	32
USA	16	39	25
Portugal	20	2	23
Kanada	12	3	20
Deutschland	18	14	18
Frankreich	13	7	13
Österreich	13	1	9
Russland	2	3	6
Schweiz	8	1	3
Türkei	2	1	0

Tab. 3: Teilnehmer an den beiden wirtschaftsgeschichtlichen Konferenzen «European Historical Economics Society» (Venedig 1996) und «World Clio Conference» (München 1997)

\* Pro 100 Mio. Einwohner eines Landes

\*\* Die Residuen einer Regression von Teilnehmeranzahl pro Bevölkerung auf die geschätzten Reisekosten, standardisiert auf Werte zwischen 0 und 100.

### 3. Wie können die «Kosten» des Statistikeinsatzes minimiert werden?

Versetzen wir uns in die hypothetische Situation eines Historikers, der quantitative Zusammenhänge untersuchen möchte. Nehmen wir an, er hat nicht die Zeit, mehrere Monate Statistik zu studieren. Zunächst: Was für Hilfsmittel kann man empfehlen?

## **Erforderliche <Ressourcen> für den historischen Statistikeinsatz:**

### *a) Technische Ausstattung*

Sie benötigen einen 486er oder Pentium Computer, eine SPSS für Windows-Lizenz, zusätzlich vielleicht eine Software, um die Daten bequem eingeben und Kategorien<sup>13</sup> zuordnen zu können (z.B. Excel oder dBase).

### *b) Geduld*

Viele Untersuchungen enden mit zweifelhaften Ergebnissen, weil

- bestimmte Tests auf problematische Schätzungen hinweisen,
- Zweifel an der Meßgenauigkeit einer Variable aufkommen,
- ein anderer, viel interessanterer und besser geeigneter Datensatz dem Forscher im Archiv begegnet.

### *c) Risikobereitschaft*

Es wird <Statistikexperten> geben, die sich über scheinbare oder tatsächliche Fehler mokieren; natürlich auch andere, die hilfreiche Kommentare geben.

### *d) Standfestigkeit*

Man sollte sich von fast unaussprechbaren Begriffen wie «Heteroskedastizität» nicht abschrecken lassen. Zudem sollte man darauf vorbereitet sein, daß man sich in einem eher traditionell ausgerichteten geschichtswissenschaftlichen Umfeld oft rechtfertigen muß: Was bringt ein quantitativ belegter Zusammenhang an zusätzlicher Erkenntnis?

### *e) Kritikakzeptanz*

Vielleicht noch wichtiger als in anderen Forschungsbereichen ist der wissenschaftliche Austausch von Kritik, weil viele (aber eben oft identifizierbare!) Fehlerquellen existieren.

Gerade über den letzten Punkt sollten wir uns Gedanken machen. Es ist eine Binsenweisheit, daß der Mensch Fehler macht. In der Wissenschaft spielt jedoch der Geniegedanke des 19. Jahrhunderts in unseren Köpfen noch eine wichtige Rolle: Selbst macht man keine Fehler, alle anderen Studien zum selben Thema sind hingegen unbefriedigend oder weniger wichtig. Unterstützt wird dies natürlich durch die Tatsache, daß man in der Einleitung von Dissertationen und anderen wissenschaftlichen Werken darlegen muß, warum gerade die eigene Arbeit unglaublich wichtig ist und alle anderen Studien überragt. Das ist zwar im Prinzip eine wichtige Betrachtung: Man muß schließlich rechtfertigen, warum man so viel Mühe in die Untersuchung steckte und meistens auch finanzielle Opfer brachte, und

---

13 Z.B. Ja/Nein-Kategorien der erklärenden Variablen («Dummyvariablen»): Jede Dateneinheit bekommt eine 1, wenn ein Kriterium zutrifft und sonst eine Null.

warum der Leser die Zeit investieren sollte, die Untersuchung zu lesen. Nur gibt es eben oft den Nebeneffekt, daß man später geäußerte Kritik persönlich nimmt, z.B. wenn man sehr kritische Gutachten von Zeitschriften bekommt. Kritikakzeptanz ist bei statistischen Untersuchungen besonders entscheidend, weil es viele Möglichkeiten gibt, bei statistischen Untersuchungen Fehler zu machen.

Durch den Einsatz von modernen, Windows-orientierten Statistikprogrammen lassen sich die «Kosten» deutlich verringern. Zwar lassen sich einige speziellere Probleme mit den besonders anwenderfreundlichen Programmen bisher noch nicht lösen, aber dies ist auch meist nicht nötig. Ein wirklicher Fortschritt in quantitativer Wirtschafts- und Sozialgeschichte läßt sich m.E. eher dadurch erreichen, daß viele Anwender die Basistechniken benutzen, als daß wenige Kollegen mit besonders fortschrittlicher Raffinesse glänzen.

## **4. Statistische Verfahren**

### *4.1. OLS-Regression*

Die OLS-Regression (ordinary least squares, «Kleinste Quadrate-Methode») mit Querschnittsdaten<sup>14</sup> ermöglicht es in vielen Situationen, Einflüsse von mehreren erklärenden auf eine zu erklärende Variable zu untersuchen.

Wer dennoch mit besonderen statistischen Verfahren arbeiten möchte, kann sich die Arbeit wesentlich erleichtern, wenn er einen Beispielartikel anschaut, der diese Technik schon verwendet. Hierfür habe ich auf meiner Homepage<sup>15</sup> die oben erwähnte Datenbank mit Zeitschriftenartikeln abgelegt. Erfahrungsgemäß hilft es oft, wenn man beim Einsatz eines neuen Verfahrens Beispiele von anderen Anwendern zur Hand hat. Man kann schauen, wie diese die statistischen Tests und Kennzahlen interpretieren, mit welchen Problemen sie zu kämpfen hatten. Wenn man näher an bestimmten Forschungen interessiert ist, stellen die Autoren gelegentlich auch ihre Datensätze zur Verfügung.

---

14 Querschnittsdaten beinhalten Informationen für verschiedene Einheiten zu einem Zeitpunkt (also z.B. Angaben für alle Verwaltungsdistrikte der Schweiz, alle Unternehmen Norddeutschlands, alle Einwohner Berns etc.), im Gegensatz zu Zeitreihen für eine Einheit oder sogenannte Paneldaten, die eine Kombination aus beiden Dimensionen darstellen und sich auf verschiedene Einheiten über die Zeit hinweg beziehen, z.B. Informationen über alle Verwaltungseinheiten eines Landes über die Zeit hinweg.

15 [http://www.vwl.uni-muenchen.de/ls\\_komlos/joerg.html](http://www.vwl.uni-muenchen.de/ls_komlos/joerg.html)

#### 4.2. Logistische Regression und (bisher) selten eingesetzte Verfahren: Ordered Probit, Count Data Analysis

Es gibt bestimmte Situationen, in denen eine OLS-Regression nicht zu interpretierbaren Ergebnissen führt (Tab. 4). Insbesondere ist dies der Fall, wenn die zu erklärende Variable nur die Werte Ja oder Nein annehmen kann (im Computer 1 oder 0). Gerade in der historischen Analyse treten solche Daten oft auf, zum Beispiel bei politischen Abstimmungen über einen Volksentscheid, bei den Determinanten der Heiratswahrscheinlichkeit von Individuen.

Diese Technik hat in den oben diskutierten Zeitschriften bereits eine breite Anwendung gefunden (Anhang Übersicht 1), daher stehen hier viele Anwendungsbeispiele zur Verfügung.

- **OLS-Regression (Kleinste Quadrate-Methode):** Verwendbar für Daten, bei denen die zu erklärende Variable aus kontinuierlich gemessenen Zahlen besteht (im Gegensatz zu den nachfolgenden beiden Beispielen). Implementiert in SPSS für Windows.<sup>16</sup>
- **Logistische Regression:** Verwendbar für Daten, bei denen die zu erklärende Variable aus «Ja-oder-Nein»-Werten besteht. Implementiert in SPSS für Windows.
- **Ordered Probit:** Für Daten, bei denen die zu erklärende Variable mehr als zwei Kategorien umfaßt, die ordinal zu interpretieren sind (z.B. sehr wenig-wenig-viel-sehr viel). In komplizierteren Programmen wie z.B. Limdep implementiert, für SPSS hat Steffen Kühnel ein Makroprogramm geschrieben.<sup>17</sup>
- **Zählmodellen (Count Data Analysis):** Für den Fall, daß die zu erklärende Variable kleine, ganzzahlige, positive Werte annimmt (oder Null), aber als Zahl zu interpretieren ist (z.B. Schiffsunfälle). In Programmen wie Limdep oder GAUSS implementiert.

Tab. 4: Übersicht über einige statistische Verfahren zur Analyse von Querschnittsdaten

16 In dieser Software sind auch Möglichkeiten vorgesehen, die Annahmen der Regression zu überprüfen, deren Verletzung die Schätzung beeinträchtigen kann. Peter Kennedy (vgl. Anm. 7) hat diese Annahmen sehr anschaulich beschrieben.

17 Kühnel, Steffen: «Programm zur Logitanalyse von kategorialen abhängigen Variablen auf Individualdatenebenen». In: *Historical Social Research* 1995-3, S. 63-87.

Es gibt jedoch Datenmaterial, das nur unter (erheblichem) Informationsverlust in 1/0-Variablen umkodieren lassen. Hier zwei solche Fälle:

1. Die zu erklärende (= abhängige Variable) kann nur wenige Werte annehmen, die zudem nicht als numerische Abstände zu interpretieren sind, sondern als Rangfolge (z.B. die Fragebogenantworten: Wenig/Mittel/Viel).
2. Die zu erklärende Variable nimmt meist nur kleine, ganzzahlige, positive Werte an (oder Null), ist aber als Zahl zu interpretieren (z.B. Schiffsunfälle verschiedener Reedereien oder Hexenverbrennungen in verschiedenen Herrschaften).

Die Antwort auf den erstgenannten Fall heißt «*ordered probit*» und hat m.E. ein großes Potential gerade für historische Anwendungen. Stellen wir uns vor, wir würden mit einem Fragebogen eine Ortsbevölkerung befragen. Sehr oft würden wir Fragen stellen, die mit «sehr wenig/wenig/viel/sehr viel» zu beantworten wären. Um Zusammenhänge zu erkennen, möchten wir die Antworten auf Charakteristika der Befragten zurückführen, deren soziale Schicht, Einkommen, Einstellungen zu anderen Dingen, Alter usw.

Für diese Art von Untersuchung wird bei der Ordered Probit Analyse die Antwort – in diesem Fall – als Werte von 0 bis 3 kodiert, wobei die Unterschiede zwischen 0 und 1, 2 und 3 usw. nicht numerisch zu verstehen sind, sondern ordinal: 1 ist mehr als 0, aber nicht «genauso viel» mehr wie der Unterschied von 2 und 3.

Leider ist diese Schätzmethode zur Zeit noch nicht in den benutzerfreundlichen Standardprogrammen implementiert, man muß sich an spezialisierte Statistikprogramme wie z.B. Limdep<sup>18</sup> wenden. Limdep ist eine Software, bei der man kleine Programmdateien erzeugen muß. Möchte man die Verwendung dieser Programme umgehen, lassen sich die Daten manchmal auch sinnvoll in 1/0 Werte umkodieren, z.B. indem man die Fragebogenantworten «sehr wenig» und «wenig» mit Null bewertet und die anderen mit Eins. Es ist offensichtlich, daß dabei jedoch Information verloren geht.

Dies gilt auch für das zweite Verfahren, daß m.E. ein großes Potential für historische Anwendungen birgt, die «*Count Data Analysis*» (Zähl-datenanalyse).<sup>19</sup> Für Objekte der historischen Analyse lassen sich oft

---

18 Falls diese Software nicht am lokalen Rechenzentrum oder einem sonst zugänglichen Ort verfügbar ist, muß sie für ca. \$ 600 bei Econometric Software (Fax +1-516-286-6848) gekauft werden.

19 Vgl. zur Zähl-datenanalyse Greene, William H.: *Econometric analysis*. 3. Aufl. Upper Saddle River (NJ) 1997.

bestimmte «Ereignisse» zählen, wobei viele Nullen und kleine Werte auftreten. Z.B. wird bei Schiffsunglücken oft der Wert «kein Unglück» vorkommen, sehr viel seltener der Wert «14». Für solche ungleichmäßig verteilten und zudem keine negativen Werte annehmenden Daten eignet sich die Zählratenanalyse wesentlich besser als die OLS-Schätzung. Auch solche Variablen lassen sich notfalls in 1/0-Werte transformieren, wobei ebenfalls Informationen verloren gehen.

Es sei nochmals kurz darauf hingewiesen, daß diese spezielleren Verfahren nur nötig sind, wenn die zu erklärenden Variablen bestimmte Eigenschaften aufweisen. Wenn die abhängige Variable eine kontinuierliche Variable ist, lassen sich mit der einfach anzuwendenden OLS-Regression die relevanten Ergebnisse erzeugen.<sup>20</sup>

---

<sup>20</sup> In einigen Fällen, wenn z.B. die abhängigen Variablen Prozentsätze sind, wären eigentlich anspruchsvollere Verfahren angebracht (Logit-Regression mit Anteilsdaten), aber erfahrungsgemäß ergeben diese Verfahren meist dieselben Ergebnisse.



## Anhang

### Übersicht 1: Beispiele des Einsatzes von Logit und Probit zur Untersuchung von Einflüssen auf qualitative Variablen

<b>Methode</b>	<b>Zusammenhang</b>	<b>Untersuchung: Titel, Autor/-in, Erscheinungsort</b>
<b>probit</b>	<b>participation</b>	Women's Labour Force Participation and the Transition to the Male-Breadwinner Family, 1790-1865. Autoren/innen: Horrell, Sara; Humphries, Jane. In: Economic History Review 48 (1), February 1995, S. 89-117.
<b>logit</b>	<b>vote</b>	Possession Is Nine Points of the Law: The Political Economy of Early Public Land Disposal. Autor/in: Kanazawa, Mark T. In: Explorations in Economic History 33 (2), April 1996, S. 227-249.
<b>probit</b>	<b>opinion of labor quality</b>	Human Capital and Direct Investment in Poor Countries. Autor/in: Hanson, John R., II. In: Explorations in Economic History 33 (1), January 1996, S. 86-106.
<b>probit</b>	<b>children participation</b>	The Exploitation of Little Children: Child Labor and the Family Economy in the Industrial Revolution. Autoren/innen: Horrell, Sara; Humphries, Jane. In: Explorations in Economic History 32 (4), October 1995, S. 485-516.
<b>logit</b>	<b>death from yellow fever</b>	Strangers' Disease: Determinants of Yellow Fever Mortality during the New Orleans Epidemic of 1853. Autoren/innen: Pritchett, Jonathan B.; Tunali, Insan. In: Explorations in Economic History 32 (4), Oct. 1995, S. 517-539.
<b>logit</b>	<b>retirement</b>	Agricultural Decline and the Secular Rise in Male Retirement Rates. Autor/in: Costa, Dora L. In: Explorations in Economic History 32 (4), October 1995, S. 540-552.
<b>logit</b>	<b>legislative intervention</b>	The Political Economy of Coalition-Formation: The Case of Livestock Enclosure in the Postbellum South. Autor/in: Kantor, Shawn Everett. In: Explorations in Economic History 32 (1), January 1995, S. 82-108
<b>logit</b>	<b>lifetime shaker</b>	Human Capital in Religious Communes: Literacy and Selection of Nineteenth Century Shakers. Autor/in: Murray, John E. In: Explorations in Economic History 32 (2), April 1995, S. 217-235.
<b>logit</b>	<b>work in agriculture</b>	A New Look at U.S. Agricultural Productivity Growth, 1800-1910. Autoren/innen: Geib Gundersen, Lisa; Zahrt, Elizabeth. In: Journal of Economic History 56 (3), September 1996, S. 679-686.
<b>probit</b>	<b>self-employment</b>	Myth of the Industrial Scrap Heap: A Revisionist View of Turn-of-the-Century American Retirement. Autoren/innen: Carter, Susan B.; Sutch, Richard. In: Journal of Economic History 56 (1), March 1996, S. 5-38.
<b>probit</b>	<b>participation</b>	Health and Labor Force Participation of Older Men, 1900-1991. Autor/in: Costa, Dora L. In: Journal of Economic History 56 (1), March 1996, S. 62-89.

<b>logit mit Anteilsdaten</b>	<b>infant mortality</b>	Infant Mortality and Living Standards of English Workers during the Industrial Revolution. Autor/in: Huck, Paul. In: Journal of Economic History 55 (3), September 1995, S. 528-550.
<b>logit</b>	<b>voting populist</b>	Supplanting the Roots of Southern Populism: The Contours of Political Protest in the Georgia Hills. Autor/in: Kantor, Shawn Everett. In: Journal of Economic History 55 (3), September 1995, S. 637-46.
<b>logit</b>	<b>participation</b>	Women, Marital Status, and Work Opportunities in 1880 Colorado. Autor/in: Herr, Elizabeth. In: Journal of Economic History 55 (2), June 1995, S. 339-366.
<b>logit (restr.)</b>	<b>mortality</b>	Family Income, Mortality, and Fertility on the Eve of the Demographic Transition: A Case Study of Rosny-sous-Bois. Autor/in: Weir, David R. In: Journal of Economic History 55 (1), March 1995, S. 1-26.
<b>logit</b>	<b>decision (in court)</b>	Property Rights and Patent Litigation in Early Nineteenth-Century America. Autor/in: Khan, B. Zorina. In: Journal of Economic History 55 (1), March 1995, S. 58-97.
<b>probit</b>	<b>banking crisis</b>	The Shoe That Didn't Drop: Explaining Banking Stability during the Great Deparession. Autor/in: Grossman, Richard S. In: Journal of Economic History 54 (3), September 1994, S. 654-682.
<b>probit</b>	<b>vote</b>	Who Voted for Smoot-Hawley?. Autoren/innen: Callahan, Colleen M.; McDonald, Judith A.; O'Brien, Anthony Patrick. In: Journal of Economic History 54 (3), September 1994, S. 683-690.
<b>logit</b>	<b>living alone</b>	At What Cost a Room of Her Own? Factors Contributing to the Feminization of Poverty among Prime-Age Women, 1939-1959. Autoren/innen: Barrington, Linda; Conrad, Cecilia A. In: Journal of Economic History 54 (2), June 1994, S. 342-357.
<b>ordered probit</b>	<b>insurance choice</b>	The Durable Experiment: State Insurance of Workers' Compensation Risk in the Early Twentieth Century. Autoren/innen: Price v. Fishback; Shawn E. Kantor. In: Journal of Economic History 56 (4), 1996.

## Übersicht 2: Einige weitere Beispiele für Statistikeinsatz

<b>Methode</b>	<b>Zusammenhang</b>	<b>Untersuchung: Titel, Autor/-in, Erscheinungsort</b>
<b>ADF, ARMA</b> (augmented dickey fuller test, auto-regression/moving average-time series analysis)	<b>currency</b>	Government Spending, Monetary Policies, and Exchange Rate Regime Switches: The Drachma in the Gold Standard Period. Autor/in: Lazaretou, Sophia. In: Explorations in Economic History 32 (1), January 1995, S. 28-50.
<b>ARCH</b> (auto-regression-conditional heteroscedasticity)	<b>financial history</b>	British Bank Conservatism in the Late 19th Century. Autoren/innen: Capie, Forrest H.; Mills, Terence C. In: Explorations in Economic History 32 (3), July 1995, S. 409-420.
<b>CGE</b> (computable general equilibrium simulation)	<b>depression</b>	The Causes of the Depression in Australia in the 1930s: A General Equilibrium Evaluation. Autor/in: Siriwardana, Mahinda. In: Explorations in Economic History 32 (1), January 1995, S. 51-81.
<b>ECM</b> (error correction model)	<b>demand elasticity</b>	An Econometric Model of World Silk Production, 1870-1914. Autor/in: Federico, Giovanni. In: Explorations in Economic History 33 (2), April 1996, S. 250-274.
<b>SURE</b> (seemingly unrelated regression)	<b>unemployment rate</b>	The Political Economy of Public-Private Compensation Differentials: The Case of Federal Pensions. Autor/in: Craig, Lee A. In: Journal of Economic History 55 (2), June 1995, S. 304-320.
<b>VAR</b> (restr.), instrumental variables (vector autoregressive model)	<b>shocks et al.; output, employment</b>	Wages and Employment in Britain between the Wars: Quarterly Evidence from the Shipbuilding Industry. Autoren/innen: Gazeley, Ian; Rice, Patricia. In: Explorations in Economic History 33 (3), July 1996, S. 296-318.
<b>basic structural model</b>	<b>real wages</b>	Trends in Real Wages in Britain, 1750-1913. Autoren/innen: Crafts, N. F. R.; Mills, Terence C. In: Explorations in Economic History 31(2), April 1994, S. 176-194.
<b>cochrane/orcutt transformed regression</b>	<b>wage at al.; quit rate</b>	Worker Turnover in the 1920s: What Labor-Supply Arguments Don't Tell Us. Autor/in: Owen, Laura J. In: Journal of Economic History 55 (4), December 1995, S. 822-841.
<b>cointegration analysis</b>	<b>currency</b>	The French Crime of 1873: An Essay on the Emergence of the International Gold Standard, 1870-1880. Autor/in: Marc Flandreau. In: Journal of Economic History 56 (4), 1996
<b>compound poisson</b>	<b>claims</b>	Risky Business? Nonactuarial Pricing Practices and the Financial Viability of Fraternal Sickness Insurers. Autor/in: Emery, J. C. Herbert. In: Explorations in Economic History 33 (2), April 1996, S. 195-226.
<b>GLS, random effects</b> (generalized least squares)	<b>% outdoor relief recipient</b>	Institutional Choice Matters: The Poor Law and Implicit Labor Contracts in Victorian Lancashire. Autor/in: Kiesling, L. Lynne. In: Explorations in Economic History 33 (1), January 1996, S. 65-85.

<b>hazard models</b>	<b>quits, layoffs</b>	Quits and Layoffs in Early Twentieth Century Labor Markets. Autoren/innen: Hamilton, Barton; MacKinnon, Mary. In: Explorations in Economic History 33 (3), July 1996, S. 346-366.
<b>hedonic regression</b>	<b>housing rent</b>	The Rental Price of Housing in New York City, 1830-1860. Autor/in: Margo, Robert A. In: Journal of Economic History 56 (3), September 1996, S. 605-625.
<b>proportional hazard</b>	<b>turnover</b>	Making the Effort: The Contours of Racial Discrimination in Detroit's Labor Markets, 1920-1940. Autoren/innen: Maloney, Thomas N.; Whatley, Warren C. In: Journal of Economic History 55 (3), September 1995, S. 465-493.
<b>robust regression</b>	<b>gross domestic product (gdp)</b>	Exogenous or Endogenous Growth? The Industrial Revolution Reconsidered. Autor/in: Crafts, N. F. R. In: Journal of Economic History 55 (4), December 1995, S. 745-772.
<b>spectral analysis, coherency</b>	<b>tithe &amp; prices</b>	Agrarian Cycles in Germany 1339-1670: A Spectral Analysis of Grain Prices and Output in Nuremberg. Autoren/innen: Bauernfeind, Walter; Woitek, Ulrich. In: Explorations in Economic History 33 (4), October 1996, S. 459-478.
<b>tobit, tsls</b> (two-stage-least-squares);	<b>preference for fem/male work</b>	Testing for Occupational Crowding in Eighteenth-Century British Agriculture. Autor/in: Burnette, Joyce. In: Explorations in Economic History 33 (3), July 1996, S. 319-345.
<b>truncated regression, simulation</b>	<b>heights</b>	Height and Health in the United Kingdom 1815-1860: Evidence from the East India Company Army. Autoren/innen: Mokyr, Joel; O Grada, Cormac. In: Explorations in Economic History 33 (2), April 1996, S. 141-168.
<b>wls</b> (weighted least square regression)	<b>production function</b>	Establishment Size and Economies of Scale in 19th-Century France. Autor/in: Siesic, Pierre. In: Explorations in Economic History 31 (4), October 1994, S. 453-478.

Leere Seite  
Blank page  
Page vide