

Was hat uns die "Frankfurter Schule" der Statistik gebracht?

Darstellung einer deutschen Fehlentwicklung am Beispiel der Indextheorie von Paul Flaskämper

von

Peter von der Lippe

Einführung

Die "Frankfurter Schule" in der Statistik (kurz FSS) stellte eine Richtung innerhalb der Statistik dar, die stark die begrifflichen Grundlagen der Datenerhebung, die verbale Interpretation von Statistiken, die Deskriptive Statistik sowie die "Wirtschaftsstatistik" und damit auch die Beziehung zur Praxis der amtlichen Statistik betonte.¹ Sie spielte eine nicht unerhebliche Rolle beim Wiederaufbau des Faches Statistik an den westdeutschen Universitäten nach dem Zweiten Weltkriege und war auch außerhalb der Hochschulen bis in die Zeit etwa der 1980er und 1990er Jahre in Deutschland recht einflussreich. Die FSS ist aber inzwischen von einer mehr formalen Richtung in der Statistik wegehend verdrängt.²

Zu den Hauptvertretern der FSS rechnet man meist Franz Zizek (1876 – 1938), Paul Flaskämper (1886 – 1979), Adolf Blind (1906 – 1996), sowie Heinz Grohmann, und Werner Neubauer, die alle als Professoren der Statistik an der Goethe Universität Frankfurt/Main tätig waren. Kennzeichnend für diese Richtung war, wie gesagt, die Überzeugung, dass sich die Statistik als Wissenschaft nicht nur mit mathematischen, "formalen" Aspekten (der "Zahlenlogik", wie man das nannte), sondern auch verstärkt mit inhaltlichen Aspekten (der sog. "Sachlogik") beschäftigen sollte. In der Zeit, in der sie vor allem von Flaskämper geprägt war, verstand sie sich auch als ein Gegengewicht gegen eine vorrangig formal orientierte angloamerikanische Betrachtungsweise in der Statistik.

Wir versuchen im Folgenden zu zeigen: Man kann die Betonung der Bedeutung der inhaltlichen Aussagefähigkeit und Interpretation von Methoden und statistischen Berechnungen sowie die Beschäftigung mit sehr fundamentalen Fragen, wie z.B. dem "Adäquationsproblem" durchaus für *im Grundsatz* richtig halten und auch die Gegenrichtung einer übertrieben und ausschließlich aufs Formale fokussierten Statistik und das damit verbundene Desinteresse an "Wirtschaftsstatistik" und an der amtlichen Statistik unbefriedigend finden.

Gleichwohl ist aber die FSS eher kritisch zu beurteilen und zwar aus folgenden Gründen,

- ihre Grundüberzeugungen sind wenig operational und sie bauen auf eine nicht mehr zeitgemäße Philosophie auf;

¹ Wirtschaftsstatistik wurde früher oft als gleichwertiger Bestandteil der Statistik neben Statistik als reine Methodenwissenschaft begriffen. Die Anfangszeit der FSS war noch stark vom juristischen Sprachgebrauch geprägt. Es war seinerzeit deshalb auch noch üblich zwischen einem "allgemeinen" und "besonderen" Teil einer Disziplin zu unterscheiden. So ist die "Methodenlehre" eine "allgemeine" oder "formale" Statistik und die sehr viel mehrbeachtete Wirtschaftsstatistik, bzw. ihre Teilgebiete wie Bevölkerungs-, Agrar-, Preisstatistik usw. liefen unter dem Titel "besondere" oder "materielle" Statistik. Die in Deutschland länger als in anderen Ländern anhaltende Beherrschung der Nationalökonomie durch die Juristen, und damit das "legalist thinking" wird von Englisch sprechenden Autoren auch gern als Grund für die "German fights over definitions" und das gerade auch für die FSS typische "concept splitting" genannt.

² Mit dem damit einher gehenden Bedeutungsverlust der "Wirtschaftsstatistik" hat sich wohl auch die Distanz zwischen den Statistikern an den Hochschulen und in den Ämtern vergrößert. Heutzutage ist schon der Begriff "Wirtschaftsstatistik" häufig unbekannt oder er wird verstanden als Methoden, die besonders für die Wirtschaftsforschung von Interesse sind.

- "sachlogische" Überlegungen sind oft inexakt und unwissenschaftlich in der Form, und sie liefern in der Sache auch nicht selten keine über die "Zahlenlogik" hinausgehenden Erkenntnisse,
- sie können vielmehr sogar zu widersprüchlichen und gravierend fehlerhaften Aussagen führen und schließlich
- hat die FSS mit ihrem Selbstverständnis als typisch deutsche Richtung in der Statistik nicht unerheblich dazu beigetragen, dass in Deutschland mathematisch anspruchsvollere Betrachtungen in der Statistik vernachlässigt wurden, und die Folgen eines deutschen Sonderweges noch länger nachwirkten, denn "die an angelsächsischen Vorbildern orientierte methodische Weiterentwicklung kam nur langsam voran" (Rinne 2011, S. 82).

Der zuletzt genannte Punkt dürfte der Hauptgrund dafür sein, dass die FSS im Rückblick eher kritisch zu beurteilen ist aber gleichwohl eine Erfahrung ist, dies zu analysieren lohnt.

Wir demonstrieren die Kritikpunkte am Beispiel der Indextheorie von Paul Flaskämper, einem Statistiker, der über einen auch politisch sehr entscheidenden Zeitraum hinweg die FSS geprägt hatte. Aufbauend auf der Grundüberzeugung dieser "Schule" hielt Flaskämper es nicht nur für möglich, sondern auch für den allein richtigen Weg, die Indexformel allein aus der "Sachlogik" heraus zu deduzieren. Er glaubte so zu der "logisch" einzig richtigen Formel für einen Preisindex zu gelangen und dabei verstand er seine Herangehensweise explizit als Gegenposition zu der von ihm (aber auch anderen deutschsprachigen Autoren) vehement als "formalistisch" abgelehnten Indextheorie des US Ökonomen Irving Fisher.

Wir halten Flaskämpers Überzeugung, man komme nur mit "Logik" zwingend zur *der* einzig richtigen Indexformel und man könne "beweisen", welche Formel die allein richtige ist, für im Ansatz verfehlt. Solche Auffassungen sind die Folge grundlegender methodologischer Überzeugungen der FSS, die nach unserer Auffassung falsch, zumindest sehr bedenklich sind.

Es sollte einschränkend beachtet werden, dass unsere Kritik der FSS vorrangig die Grundauffassungen dieser Schule betrifft, und die Zeit, die man als Blütezeit der FSS bezeichnen könnte (Flaskämper und Blind). Spätere, zeitgenössische Autoren, die noch als Vertreter der FSS gelten, wie Grohmann und Neubauer haben für deutlich andere, positiv zu bewertende Akzente in dieser Richtung der Statistik gesorgt und dabei sehr beachtliche Beiträge geleistet.

Die folgende Betrachtung gliedert sich in drei Teile. Im ersten Teil stellen wir die Grundüberzeugungen der FSS dar, demonstrieren dann im zweiten Teil am konkreten Beispiel von Flaskämpers Indextheorie, wohin der Versuch führt, alles aus einer Theorie des statistischen Vergleichs ableiten zu wollen, und gehen abschließend im dritten Teil noch einmal auf die zu Beginn genannten Grundüberzeugungen ein und fragen uns, inwieweit sie für Leistungen und Fehlleistungen der FSS verantwortlich zu machen sind.

1. Grundlegende Positionen der Frankfurter Schule

In diesem Teil stellen wir die Vorstellungen dar, die nach Auffassung der FSS notwendig (und wohl auch hinreichend)³ sind, um daraus eine wissenschaftliche Statistik zu entwickeln. Wir betrachten auch die für sie zentralen Begriffe Quantifizierbarkeit und Vergleichbarkeit.

a) Drei Dichotomien

Grundlegend für die FSS ist die Vorstellung von zwei sich jeweils *prinzipiell* und nicht nur graduell unterscheidenden Dingen (einem "Dualismus"⁴ oder einer Dichotomie). Wir haben

³ Siehe den im zweiten Teil dargestellten Versuch Flaskämpers, allein hieraus eine Indextheorie abzuleiten.

zunächst als eine erste Dichotomie die Unterscheidung in Natur- und Geisteswissenschaften,⁵ die von der FSS primär an der (schwer operational zu definierenden) "Quantifizierbarkeit" festgemacht wurde,⁶ sich aber auch am Objekt (belebt/unbelebt, vorgefunden/vom Menschen gemacht usw.)⁷ und an der für erforderlich gehaltenen Betrachtungsweise orientiert. Was letztere betrifft, so mag es genügen, einige Schlagworte (in Gestalt von Gegenüberstellungen) zu nennen: nomothetisch (mit Ursachen erklären), generalisieren, nach Gesetzen suchen vs. idiographisch, ("verstehend" beschreiben), individualisieren (historisch begreifen), nach "Sinn" und "Bedeutung" fragen; oder intersubjektive Nachprüfbarkeit und Wahrheitsfindung nur durch Erfahrung einerseits und Erkenntnisgewinnung durch reines Nachdenken oder auch durch Intuition und Einfühlung andererseits.

Auf Seiten der Anhänger eines Dualismus der Wissenschaften, zu dem sich die FSS dezidiert bekennt ist auch oft die Rede von "Sinn"-Interpretation, einem hinter der bloßen "Erscheinung" verborgenen tieferem "Wesen" oder auch von "organischen" und geistigen Entitäten statt mechanischen oder stofflichen Phänomenen.

Die FSS ist (oder eher: war) mit den beiden Vorstellungen der Existenz geistiger, aber prinzipiell unerkennbarer Entitäten, wie das "Wesen" (hinter der "bloßen" Erscheinung), oder das "Ding an sich" (Kant), und dass alles aus *einem* "letzten" Grund heraus deduktiv ableitbar ist, also zwingend zu folgern ist, ein Kind des "Deutschen Idealismus" (Fichte, Hegel, Schelling), und damit auch in diesem Sinne, eine sehr "deutsche" Richtung in der Statistik. Solche heutzutage eigentümlich anmutende, seinerzeit jedoch sehr gebräuchliche Vorstellungen stehen im krassen Widerspruch zu einer uns jetzt, gerade auch als Statistiker sehr viel näher stehenden Philosophie, nämlich dem Positivismus.⁸

Vor diesem philosophischen Hintergrund werden viele nachfolgend zitierte Aussagen von Vertretern der FSS, insbesondere von Flaskämper überhaupt erst verständlich und hierauf wurde auch (implizit) Bezug genommen, um zwei weitere Dichotomien einführen zu können:

- die Existenz von zwei gleichberechtigten Arten von Statistik, der "mathematischen" Richtung (Stochastik⁹) und einer anderen, der Richtung der FSS, bei der aber schon allein die Namensgebung Schwierigkeiten bereitet: "nichtmathematische Statistik" wurde als nicht genügend trennscharf und als diskriminierend empfunden,¹⁰ andere Vorschläge, die sich nicht durchsetzten, waren "soziale Arithmetik" (Flaskämper), "logische Richtung" (W. Winkler) oder "soziale Zähl- und Meßkunst" (Blind 1952) und in späterer Zeit sprach man auch gerne von "Sozialstatistik", und
- das "Postulat" des "Parallelismus von Sach- und Zahlenlogik", wonach eine formale Betrachtung mit einer begrifflichen Klärung beginnen und mit einer inhaltlichen Reflexion und Interpretation abgeschlossen werden muss, um "sinnvoll" zu sein.

⁴ "Polarität" entspricht wohl wegen der fließenden Übergänge und denkbaren Zwischenformen nicht den Vorstellungen der FSS.

⁵ Es gibt auch die Unterscheidung zwischen Naturwissenschaft einerseits und Kultur- oder Sozialwissenschaften andererseits.

⁶ Nach Flaskämper 1933, S. 60 ist "das Organische, der Sinnzusammenhang" typisch für die Sozialwissenschaften, "die Begriffe sind 'ganzheitlicher Art'" und es gilt, "daß der Kern der Erscheinungen hier sich jeder Zähl- und Meßbarkeit entzieht" (S. 62).

⁷ Bei Flaskämper 1950 aber auch Below 1956 findet sich auch die Vorstellung, dass die Biologie eine dritte Kategorie zwischen Natur- und Kulturwissenschaft darstellt, dass es also geboten sei, bei "lebend" noch zu unterscheiden zwischen höheren und niedrigeren Lebensformen.

⁸ Konzepte wie "Wesen" usw. waren metaphysisch und deshalb unhaltbar für den Positivismus, der u.a. besonders prominent vom englischen Statistiker Karl Pearson (1857 – 1936) vertreten wurde und der vom Prinzip der Einheit der Wissenschaft ausging, also den Gedanken prinzipiell verschiedener Wissenschaftstypen ablehnte.

⁹ Der Begriff ist heute sehr üblich. Er scheint in Deutschland von L. von Bortkiewicz eingeführt worden zu sein.

¹⁰ Richtig ist aber sicher, dass sich diese, wie immer genannte Richtung dadurch von der "mathematischen" unterscheidet, dass sie meint, sich mit "elementarmathematischen" Methoden begnügen zu können.

Für die zweite Dichotomie (zwei Arten von Statistik) ist vor allem die Begründung eines eigenständigen Werts der Deskription (Deskriptive Statistik und Wirtschaftsstatistik) gegenüber der auf der Wahrscheinlichkeitsrechnung aufbauenden "mathematischen" Statistik zentral. Nach Vorstellung der FSS kann sie durch die "Logik des Vergleichs" erfolgen, genauso wie das "Gesetz der großen Zahlen" die Basis der mathematischen Statistik sei.¹¹

Der Versuch der "mathematischen" Statistik etwas entgegenzusetzen provoziert natürlich die Frage, wie man es in der FSS mit der Mathematik halten will. Blind 1952 ist bemüht, dem Eindruck entgegenzutreten, es gehe hier nur um "mangelnde Vertrautheit mit der Mathematik" (S. 530).¹² Es sei vielmehr so, dass sich gezeigt habe, dass die größere mathematische Raffinesse in der "Sozialstatistik" nicht anwendbar sei und man *deshalb* der mathematischen Statistik eine andere Art von Statistik gegenüberstellen müsse.

Die dritte Dichotomie ("Parallelismus") dürfte aus heutiger Sicht am ehesten akzeptabel sein (so lange man sich nicht die Mühe macht, die Begriffe Sachlogik und Zahlenlogik exakt definieren zu wollen). Die herausragende Rolle der "Sachlogik" rechtfertigt sich mit dem Gedanken, dass soziale und kulturelle Gegenstände prinzipiell nicht "quantifizierbar" seien. Sie seien nicht "messbar", weil es fließende Übergänge zwischen Erscheinungen und einen Unterschied zwischen dem Beobachtungskonzept und dem ideellen Konzept gibt. Das damit, also mit der fehlenden (oder eingeschränkten) Quantifizierbarkeit angesprochene "Adäquationsproblem" und der Gedanke der "Validität" einer Messung dürften, wie auch die Forderung nach einer sachlichen und anschaulichen Interpretation auf dem ersten Blick überzeugend sein. Bei genauerem Hinsehen zeigt sich jedoch, dass auch hier entscheidend auf unklare Konzepte aufgebaut wird. Das wird deutlich, wenn man die zur Begründung der behaupteten Nicht-Quantifizierbarkeit angebotenen Argumente betrachtet.

b) "Wesen" und "Quantifizierbarkeit"

Nur oberflächlich betrachtet kann man, so heißt es bei Flaskämper, Betriebe zählen (so wie man z.B. Bäume zählen kann). "Das Wesen des Betriebsbegriffes aber können wir niemals in eine mathematische Formel auflösen, es ist ein sinnhafter, organischer Zusammenhang, kein summenhafter Begriff" (Flaskämper 1950, S. 155), was wohl so zu verstehen ist, dass man die "wahre" Anzahl der Betriebe nicht feststellen kann, wie immer man "Betriebe" definiert (und dies wird auch wohl immer so bleiben, weil wenig klar ist, was hier mit dem "Wesen" des Betriebs gemeint ist). Auch bei Bäumen gibt es Definitionsprobleme und fließende Übergänge zu anderen Gewächsorten, aber kein "Wesen des Baumbegriffs". Denn es sind offenbar Qualitäten, wie "sinnhaft" und "organisch", die ein Betrieb hat, aber ein Baum nicht hat. Im gleichen Aufsatz heißt es, in den Sozialwissenschaften "kommen quantifizierbare Begriffe so gut wie gar nicht vor, wohl aber solche ganzheitlicher Art... (Man) hat es nicht mit zähl- oder messbaren Tatbeständen zu tun, sondern mit Sinnzusammenhängen ... Der Kern wirtschaftlichen Geschehens ist nicht quantifizierbar" (S. 154). Neben Sinnhaftigkeit gibt es also noch die "Ganzheitlichkeit" und die unsichtbaren Zielgrößen "Kern" und "Wesen", die eine Quantifizierung unmöglich machen oder sie zumindest erschweren.

¹¹ Es war seinerzeit ein Thema, ob die "Logik des Vergleichs" und das "Gesetz der großen Zahlen" jeweils die Basis der einen, bzw. anderen Richtung (der "logischen" und der mathematischen") sein könne, oder ob es nur *eine* Statistik gibt, die auf diese zwei (für gleich fundamental gehaltenen) Säulen aufbaut. Für uns wirkt es heutzutage eigenartig, wie man von einer "Überschätzung des Gesetz der großen Zahlen" (Flaskämper 1929b, S. 232) als mögliches Fundament der Statistik sprechen kann. Vgl. auch Flaskämper 1927a.

¹² Genau dieser Eindruck drängt sich aber bei vielen Äußerungen der entsprechenden Autoren über Gegenstände der "mathematischen Statistik" auf. Aus heutiger Sicht ist es geradezu peinlich, was von den Vertretern der "nichtmathematischen" Richtung über die mathematischen Richtung und deren angeblich problematischen Konzepte geschrieben wurde. Man könnte hier zahlreiche Zitate aufführen, aus denen deutlich hervorgeht, dass sie nicht wussten, wovon sie sprachen, als sie sich über die von ihnen mehr oder weniger abgelehnten mathematischen Statistik äußerten.

Es stellt sich dann natürlich die Frage, was unter solchen Voraussetzungen überhaupt mit "Statistik" ausgerichtet werden kann, denn ohne Quantifizierung kann es keine Statistik geben. Aus der nicht *von vornherein* gegebenen Quantifizierbarkeit folgt für die FSS nicht, dass Statistik nicht möglich sei, sondern "nur", dass sie überhaupt erst durch Konkretisierung von Begriffen möglich gemacht werden kann.

Das Definieren eines dem ideellen Konzept möglichst nahe kommenden ("adäquaten") Zähl- oder Beobachtungsbegriffs wird meist als "Adäquationsproblem" oder Problem der "Operationalisierung" von Konzepten bezeichnet. Es wird gerne als *das* zentrale Problem der Wirtschaftsstatistik angesehen und hierauf besonders aufmerksam gemacht zu haben gilt allgemein als eine besondere, verdienstvolle Leistung der FSS.

Kennzeichnend für die "Adäquation" (Näherung auf der Ebene der Begriffe), ist dass sie als rein gedankliche Leistung verstanden wird und mit Begriffen oder "Konzepten" operiert statt mit Gleichungen und Formeln, im Unterschied etwa auch zu organisatorischen und erhebungstechnischen Leistungen. Dass "logisch" ein Lieblingwort der FSS war, ist wohl darauf zurückzuführen, dass seinerzeit dieses Wort gerne für eben diese gedankliche Arbeit des Definierens, Subsummierens usw. gebraucht wurde, und weil *diese* Art Tätigkeit zum Angelpunkt der Statistik wurde verstand sich die FSS auch als "logische Richtung".

Neben der Quantifizierbarkeit (und Adäquation verstanden als messbar *machen*) stand auch die Frage der Vergleichbarkeit (oder das Vergleichbar-*machen*) im Fokus der FSS und verwandter Richtungen in der Statistik.¹³ Auch das ist primär ein Thema der "Logik" oder "Sachlogik" in der Terminologie der FSS, oder bescheidener ausgedrückt, eine gedankliche Arbeit.

c) Vergleichbarkeit als zentraler Gegenstand der Statistik

Schon für Zizek, dem Begründer der FSS war *das* zentrale Problem der Statistik der "statistische Vergleich", bzw. es waren die "logischen" Voraussetzungen, die erfüllt sein müssen, damit ein solcher Vergleich "sinnvoll" ist. Eine grundlegende Einsicht dabei war, dass Vergleichbarkeit Ungleichheit in *einem* Aspekt, dem Gegenstand des Vergleichs, aber Gleichheit in allen anderen Aspekten voraussetzt. Nach Zizek, wäre es z.B. sinnlos, die Körperlänge deutscher Männer mit der von englischen Frauen zu vergleichen (Zizek 1931, S. 536), weil hier offenbar Ungleichheit in *zwei* Merkmalen (Nationalität und Geschlecht) und nicht nur in einem besteht.¹⁴ Weil es danach bei zwei zu vergleichenden (Teil-)Gesamtheiten entscheidend darauf ankam, ob diese in allen Aspekten bis auf einen gleich sind bestand die nächste Aufgabe darin, sich Gedanken darüber zu machen, inwiefern (oder in welchem Maße) statistische Massen "homogen" oder "gleich" sind.

Weil in der Praxis Massen stets nur mehr oder weniger "gleich" und damit vergleichbar sind, wurde schnell allgemein akzeptiert, dass es eine absolute Vergleichbarkeit nicht gibt.¹⁵ Es ist auch klar, dass sich Gleichheit auf verschiedene Betrachtungsebenen erstrecken kann. Im einfachsten Fall mögen Massen gleich sein, weil sie unter den gleichen Oberbegriff fallen, wie dies bei den Positionen einer "Klassifikation" oder "Systematik", z.B. von Waren, Wirt-

¹³ Ein anderes Thema, das seinerzeit die Statistiker bewegte, war die "Ursachenforschung" mit Mitteln der Deskriptiven Statistik (also ohne Signifikanztests). Es spielte eine Rolle in der Indextheorie von Laspeyres (Gedanke des "reinen Preisvergleichs"), weniger aber in der von Flaskämper.

¹⁴ Wie bei der Ursachenforschung standen also auch hier bedingte Mittelwerte (bzw. Anteile) im Vordergrund. Man dachte noch nicht an eine Regression mit zwei und mehr Dummy Variablen und auch nicht an zufällige und signifikante Unterschiede. Nur vage Andeutungen in dieser Richtung (3σ Bereich) findet man bei Zizek.

¹⁵ Nach der hier konsultierten Literatur soll L. v. Bortkiewicz an eine Rangskala der Vergleichbarkeit gedacht haben. Das Problem beim Gedanken eines Kontinuums der Vergleichbarkeit ist, dass es auch darauf ankommt, mit wie vielen Merkmalen und in welcher Merkmalsgliederung man die Einheiten einer Masse beschreiben möchte. Bei hinreichend detaillierter Beschreibung ist fast nichts miteinander vergleichbar, und bei einer entsprechend groben Betrachtung sind fast alle Massen "vergleichbar".

schaftszweigen usw. intendiert ist. Wie allgemein bekannt spielen Fragen dieser Art gerade für die statistische Praxis eine große Rolle.

Eine andere Frage ist, ob Überlegungen dieser Art ausreichend Substanz liefern, damit man von einer "Logik des Vergleichs" sprechen kann und ob man aus ihr gar eine Indextheorie oder eine konkrete Preisindexformel entwickeln kann (worin ja Flaskämpers Programm bestand). Die Idee war ja, man könne aus dieser "Logik" alle statistische Methoden "deduktiv" ableiten.

Dass dies aber ganz offensichtlich nicht der Fall sein kann, war schon daran erkennbar, dass man mehr oder weniger strenge Forderungen an die Vergleichbarkeit stellen kann,¹⁶ und dass man sich bei diesem Thema schnell in eine (typisch deutsche?) Sophisterei begab.¹⁷ Das manifestierte sich in einer Vielzahl von Gleichheitsbegriffen. So unterschied man materielle und formale Gleichheit, "Wesensgleichheit", "Gefügegenheit"¹⁸, "Gleichwertigkeit" usw. mit z.T. sehr feinsinnigen Differenzierungen, auf die wir hier nicht eingehen wollen, obgleich sie durchaus ihre Berechtigung haben mögen.

Schon das Subsummieren unter dem gleichen Oberbegriff kann ja aus verschiedenem Grunde erfolgen. Das tertium comparationis kann z.B. darin bestehen, dass sich die Einheiten unterschiedlichen Teilgesamtheiten der gleichen Gesamtheit sind, über ihre Entstehungsgeschichte miteinander (genetisch) "verwandt" sind, sich am gleichen Ort befinden, oder allgemeiner, dass sie hinsichtlich als wesentlich erachteter Merkmale "ähnlich" sind. Dabei ist es natürlich wieder entscheidend, welche Merkmale in welcher Differenzierung herangezogen werden.

Eine seinerzeit sehr beliebte Überlegung war, dass Vergleichbarkeit von Massen voraussetzt, dass die Einheiten der beiden zu vergleichenden Massen dem gleichen "Ursachenkomplex" unterliegen und die Massen (bzw. Teilgesamtheiten) mithin "homogen" sein müssen. Auch hier bleibt eine Vagheit, weil man natürlich den Kreis der zu berücksichtigenden Einflussfaktoren zur Beschreibung des Ursachenkomplexes enger oder weiter ziehen kann. Dieser Gedanke der "Isolierung einer Ursache" spielte aus zwei Gründen für das gerade seinerzeit in Deutschland übliche Verständnis von Statistik eine Rolle, nämlich

1. beim Vergleich von Massen (etwa A und B) mit beschreibenden Statistiken (wie z.B. mit zwei Verhältniszahlen oder zwei Mittelwerten, etwa \bar{x}_A mit \bar{x}_B), und
2. bei der Konstruktion von Statistiken zur Beschreibung von Prozessen in der Zeit (z.B. bei der Bestimmung einer geeigneten Indexformel),

und dies betrifft nicht nur das Gedankengut der FSS, sondern ist auch weiterhin relevant. Denn beides sind, anders als die angedeuteten sophistischen Unterscheidungen von Gleichheitstypen, Betrachtungen die durchaus ihren Wert haben

Zu 1: hier geht es um das Vergleichbar-*machen* von Verhältniszahlen, indem man sie auf die gleiche Struktur bezieht. So werden z.B. standardisierte Sterberaten von zwei Ländern, A und B berechnet indem man die jeweiligen altersspezifischen Sterberaten mit der gleichen ("Standard-") Altersstruktur gewichtet. Das ist grundsätzlich ein übliches und (nicht nur von der

¹⁶ Dabei sind sogar sehr eigenartige Kriterien aufgestellt worden: Nach Zizek sind "streng genommen nur solche statistische Zahlen vergleichbar, die sich auf gleich grosse Beobachtungsmassen beziehen, bei denen also das Gesetz der grossen Zahlen gleichstark wirksam gewesen ist". Dies ist ein Zitat aus Zizeks Buch "Fünf Hauptprobleme der statistischen Methodenlehre" München und Leipzig 1922, S.34, was wir einer Besprechung dieses Buchs durch Alexander Tschuprow in Nordisk Statistisk Tidskrift, Vol II, 1923, S. 166 entnommen haben. Tschuprow hielt dagegen, dass es nur darauf ankomme, dass dieses Gesetz "hinreichend stark 'wirksam' " sei.

¹⁷ Für detailliertere Hinweise vgl. Klezl Norberg 1943/44.

¹⁸ Vergleichbarkeit weil sich die Massen in gleicher Weise in Teilgesamtheiten gliedern. Bei einem richtigen Verständnis des Konzepts der "Verteilung" in der Statistik ist es schwer nachzuvollziehen, was dies von Wesensgleichheit (die gleiche Verteilung, damals auch "Wesensform" genannt, besitzend) unterscheiden soll.

FSS) befürwortetes Verfahren zur Ausschalten vergleichsstörender Faktoren (in diesem Fall des Einflusses der Altersstruktur auf die Sterblichkeit).¹⁹ Für die rohen Sterberaten kann $m_A > m_B$ gelten, gleichwohl aber für die standardisierten Sterberaten $m_A^* < m_B^*$. Im ersten Fall (bei m) sprach man seinerzeit von einem "uneliminieren", und im zweiten (m^*) von einem "eliminieren" Vergleich.²⁰

Zu 2: hier geht es um einen Vergleich einer Größe im Zeitablauf, etwa um einen Preisindex mit der Zeitreihe $P_{01}, P_{02}, P_{03}, \dots$, und hier kommt das gerade auch in Deutschland (und v.a. von Laspeyres) stets sehr hoch gehaltene "Prinzip des reinen Preisvergleichs"²¹ (durch Konstanz der Mengen q) zum Tragen, auf das wir noch zurückkommen werden.

2. Die verfehlte Indextheorie von Paul Flaskämper

In diesem Teil versuchen wir zu zeigen, wie Flaskämpers Projekt, "die" einzig "logische" Indexformel aus einer "Logik des Vergleichs" zu deduzieren zu widersprüchlichen und nicht selten grob falschen Aussagen über Indexformeln führte und somit gescheitert ist. Es war bewusst als Antithese zu Irving Fishers nach wie vor sehr einflussreichen Theorie konzipiert, und dass es scheiterte ist nach unserer Auffassung eine notwendige Konsequenz der dargestellten Grundüberzeugungen der FSS. Dass man diese auch für andere bedauerliche Fehlentwicklungen verantwortlich machen kann, ist Gegenstand des abschließenden dritten Teils.

a) Das methodologische Programm

Flaskämper stand mit seiner Ablehnung von Irving Fishers "Making of Index Numbers" nicht allein da.²² Das Buch wurde von vielen als Paradebeispiel einer in Deutschland gerade *nicht* gewünschten Art von Statistik empfunden. Aber Flaskämper, für den es nur "vom logischen Standpunkt ein Sammelsurium" von Formeln war (Flaskämper 1928; S. 193, 198), ging darüber hinaus indem er einen Gegenentwurf zu dieser Indextheorie versuchte. Es ist hier nicht der Ort, die auch nicht unbedenkliche Methodologie Fishers darzustellen, und wir beschränken uns darauf, zu zeigen, dass Flaskämpers Methodologie schon im Ansatz verfehlt war.

In seiner Habilitationsschrift (Flaskämper 1928) schloss er eine Liste der "Bestimmungstücke" eines Preisindexes, wie Gegenstand (Großhandels- oder Verbraucherpreise usw.), Periodizität usw. mit der Bemerkung ab: "Daß die Formel hier unter den Bestimmungsstücken nicht aufgeführt ist, ist eine Konsequenz der Anschauung des Verfassers, ..., daß sich aus den oben genannten Bestimmungsstücken die Formel mit eindeutiger Notwendigkeit ergibt. Die Mannigfaltigkeit der Indexformeln bei einem und denselben Tatbestand beruht auf einer Verkenntnis des Indexproblems" (S. 51). Ähnlich radikal und basierend auf der gleichen verfehlten Vorstellung erfolgt seine Ablehnung von Axiomen (die früher oft als "Kriterien" bezeich-

¹⁹ Das Land A kann eine höhere "rohe" Sterberate als das Land B haben ($m_A > m_B$) einfach deshalb, weil seine Bevölkerung relativ älter (andere Verteilung des Alters) ist als die des Landes B.

²⁰ Der vergleichsstörende Faktor "Altersstruktur" galt so (durch die Standardisierung) als eliminiert.

²¹ Man könnte auch sagen, im ersten Fall (Querschnitte) geht es darum, die Wirkung von Strukturunterschieden rechnerisch auszuschalten, und im zweiten Fall darum, eine Zeitreihe y_t zu betrachten bei der sich aufeinanderfolgende Werte y_1, y_2, \dots, y_T nur in einem Aspekt unterscheiden, aber alle übrigen Einflüsse – auch hier wieder "rein rechnerisch" – konstant gehalten werden (*ceteris paribus*). Diese Zielsetzung, rechnerisch quasi Bedingungen eines Experiments zu schaffen, ist ganz deutlich erkennbar bei E. Laspeyres und seiner Indexformel.

²² Ganz ähnlich kritisierte auch Gottfried Haberler und Wilhelm Winkler Fishers "rein mechanische Einstellung zum Indexproblem" und das "Fehlen des Eingehens auf das innere Wesen des zu erforschenden Sachgebiets" (Winkler 1923; S. 576). Auch Winkler fordert die "Ableitung einer Indexformel aus dem Wesen des Sachgebiets" (S. 577). Ganz anders – und sehr weitsichtig – äußerte sich dagegen Ladislaus v. Bortkiewicz: "Deutscherseits verdient dieses Werk umso mehr Beachtung, als man hierzulande im Lauf des letzten halben Jahrhunderts sowohl in der Theorie wie in der Praxis der Preisindexzahlen merklich ins Hintertreffen geraten ist" (v. Bortkiewicz 1924; S. 853).

net wurden). Er hält sie für "eigentlich überflüssig"; und "Für uns liegt ein logisches Bedürfnis nach ihnen überhaupt nicht vor. Denn es ist ja eine unseres Erachtens sehr wichtige *These* unseres Standpunktes, daß Indexziffern ... sich logisch in nichts von Einzelindex unterscheiden" (S. 118).²³

In einer Besprechung von Gottfried Haberlers Buch "Der Sinn der Indexzahlen" (Flaskämper 1929a) wird er noch deutlicher: Die "Logik des Vergleichs" ist der einzig erfolgreiche Ariadnefaden "Wer so vorgeht, braucht keine Kriterien. Die richtigen Formeln ergeben sich für ihn aus eben dieser Logik" (S. 158). Haberlers axiomatische Betrachtungen einer Reihe von seinerzeit diskutierten Indexformeln²⁴ veranlassten ihn zu folgenden deutlichen Worten gegen eine mehr mathematische Richtung in der Statistik: "Aber dieses formalistische Vorgehen ist einer logisch und systematisch begründeten Wissenschaft unwürdig; es muß endlich einmal aufhören" (S. 154).²⁵

Es zeigt sich hieran, dass Flaskämpers Methodologie geprägt ist, von

1. der Erwartung allein durch "Logik" zur richtigen Indexformel zu gelangen, und dass es auch nur *eine* solche Formel geben kann (das ist, was wir sein "Programm" nennen wollen und was wir für schon im Ansatz verfehlt halten), und von
2. einer Fehleinschätzung der Rolle von Axiomen in der Indextheorie, weil er
 - keinen prinzipiellen Unterschied zwischen einem Index und einer Messzahl sieht,
 - verkennt, dass der Formulierung von Axiomen als Forderungen, die eine "sinnvoller" Indexformel erfüllen sollte selbst eine Interpretation (und damit praktisch eine "sachlogische" Dimension) zugrundeliegt und weil er glaubt, dass
 - die "Sachlogik" speziell in Gestalt der oft als deutsche Besonderheit gepriesenen "Logik des Vergleichs"²⁶ mindestens das leisten könne, was Axiome leisten.

Wir beginnen mit dem letzten Punkt und zeigen, dass diese "Logik" nicht das leistet, was man sich von ihr verspricht. Sie ist kein "Ariadnefaden", der uns zur angeblich allein richtigen Indexformel führt, und dass die "Mannigfaltigkeit der Indexformeln" keineswegs eine "Verkenning des Indexproblems" ist.²⁷

b) Axiome und die "Logik des Vergleichs"

Flaskämper erkennt nicht, dass Axiome als Forderungen, die eine "sinnvolle" Indexformel erfüllen sollte, selbst eine Interpretation darstellen.²⁸ Fordert man z.B. von der Formel Identität (I) oder die Mittelwerteigenschaft (M) als Axiome, so besagt dies, dass es nicht "sinnvoll" wäre, wenn der Index eine Preissteigerung anzeigt, obgleich sich kein Preis geändert hat oder eine Preissteigerung anzeigt, die größer ist als die größte oder kleiner als die kleinste Preis-

²³ Hervorhebung und "Einzelindex" im Original.

²⁴ Flaskämper spricht hier von einem "unfruchtbaren Rechnen mit Formeln" (S. 153).

²⁵ Die heftigen Worte gegen - nach Flaskämpers Meinung überflüssige - mathematische Betrachtungen in Haberlers Buch sind umso unverständlicher, als einmal diese Abschnitte sich auf sehr elementarem Niveau bewegen und zum anderen, weil sich auch Haberler dezidiert gegen den Formalismus von Fisher positionierte. Bemerkenswert an dieser Buchbesprechung ist auch, dass sie nur wenige Formeln enthält, darunter eine (auf S. 152), die schon auf dem ersten Blick als falsch erkannt werden sollte, nämlich $E_1/E_1 > \Sigma p_1q_1/\Sigma p_1q_1$, denn 1 kann ja wohl kaum größer als 1 sein.

²⁶ Alfred Jacobs 1940 in einer Rezension der Festschrift für Friedrich Zahn: "Die Mehrzahl der Beiträge zeigt ferner, daß die Logik des Vergleichs als Grundform aller Statistik in der Praxis der deutschen statistischen Forschung das Feld beherrscht" (S. 394).

²⁷ Flaskämper 1929a glaubt, dass allein schon der Umstand, dass v. Bortkiewicz mehrere Indexformeln analysierte ein Beweis sei "für den falschen Weg, auf den sich die Mehrzahl der Indextheoretiker befindet" (S. 158).

²⁸ Daher kommt wohl auch seine falsche Vorstellung, Axiome könnten nur dem Ausschluss unbrauchbarer Indexformeln dienen, nicht aber positiv der Begründung brauchbarer Formeln.

messzahl. Wenn Flaskämper demgegenüber meint, ein Index könne sinnvoll eine der beiden Mengen q_0 oder q_t (oder auch sogar $(q_0 + q_t)/2$)²⁹ enthalten "während alle jene Formeln, die die beiden Gewichte ... gleichzeitig berücksichtigen ... rein abstrakte Formeln ohne gegenständlichen Sinn sind" (Flaskämper 1929a, S. 152), so ist das von anderer Qualität, denn

- man kann zeigen (beweisen), dass eine Indexformel I oder M erfüllt (oder nicht erfüllt),
- aber wie zeigt man, dass ein Index nur mit Gewichten q_0 oder q_t einen "gegenständlichen Sinn" hat, eine Formel mit *beiden* Mengen q_0 und q_t aber nicht?³⁰

Die zuletzt angesprochene Behauptung verweist schon auf eine der gerade bei Flaskämper notorischen Widersprüchlichkeiten. Wenn $(q_0 + q_t)/2$ ein sinnvolles Gewicht sein soll, dann enthält es doch neben den Mengen q_0 auch Mengen q_t , und genau das soll "ohne gegenständlichen Sinn" sein.

Auf der gleichen Ebene befindet sich auch die Ablehnung einer Mittelung von Formeln (etwa ein Mittel aus Laspeyres, P^L und Paasche, P^P) bei gleichzeitig Befürwortung oder zumindest Tolerierung von gemittelten Gewichten, wobei sich doch in vielen Fällen leicht zeigen läßt, dass beides "formal" auf das gleiche hinausläuft.³¹ In der folgenden Gleichung wird der Index P^{ME} (Preisindex nach Marshall und Edgeworth) dargestellt als Index mit gemittelten Mengengewichten und als gewogenes Mittel aus den Preisindizes P^L und P^P

$$(1) \quad P_{0t}^{ME} = \frac{\sum p_t \cdot \frac{1}{2}(q_0 + q_t)}{\sum p_0 \cdot \frac{1}{2}(q_0 + q_t)} = \frac{1}{1 + Q_{0t}^L} \cdot P_{0t}^L + \frac{Q_{0t}^L}{1 + Q_{0t}^L} \cdot P_{0t}^P.$$

Es führt also zu nichts, die "Logik des Vergleichs" und Axiome gegeneinander ausspielen zu wollen, aber man kann sich fragen, wie weit man mit der Forderung der Vergleichbarkeit kommt, wenn man mit ihr eine Indexformel rechtfertigen oder verwerfen will.

Der Gedanke des "reinen Preisvergleichs" ("pure" price index), wonach eine Indexformel nur die Preisbewegung widerspiegeln sollte (und z.B. nicht auch die Veränderung der Mengen), spielt eine wichtige Rolle als Argument für einen (direkten) Laspeyres Preisindex und gegen Kettenindizes aller Art. Aus ihm kann aber nicht eine einzig richtige Formel deduziert werden; denn man kann zeigen, dass

- sich sogar gegensätzliche Meinungen darüber, welche Indexformeln sinnvoll sind gleichermaßen auf eine eigene "Theorie" der Vergleichbarkeit berufen, aus der dann mit *zwingender Notwendigkeit* die angeblich allein richtige Formel folgt,
- es nicht unumstritten ist, welche Folge von Indexzahlen $P_{01}, P_{02}, P_{03}, \dots$ Ausdruck einer "reinen" Preisentwicklung ist
- auch andere Forderungen an Indexformeln als reiner Preisvergleich berechtigt sind (weshalb es auch eine Illusion ist, es gäbe *eine* allein richtige Formel),

²⁹ In Flaskämper 1928 behauptet er, sich auf Hermberg (1923) berufend, es sei "logisch eindeutig bewiesen, daß alle drei sich daraus ergebende Formeltypen nebeneinander berechtigt sind" (S. 101; es bleibt offen, wie man so etwas überhaupt "beweisen" kann), also Laspeyres (P^L) mit q_0 , Paasche (P^P) mit q_t und Marshall-Edgeworth P^{ME} mit $(q_0 + q_t)/2$ gleichwertig sind. Andererseits hält er ein Mittel (wohl jede Art von Mittel) aus P^L und P^P "logisch" für "eine Unmöglichkeit" (S. 107).

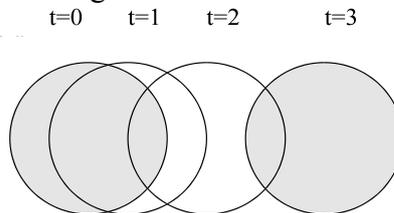
³⁰ Es ist sicher zu begrüßen, "... daß man sich bei jeder Formel genau nach ihrem logisch-sachlichen Sinn fragt" (Flaskämper 1929a, S. 153), aber es ist nicht immer einfach, in puncto "Sinn" zu exakten Feststellungen zu gelangen und es ist auch nicht richtig, dass man sich keine solchen Fragen stellt, wenn man axiomatische Betrachtungen anstellt. Das Verdammn von Formeln mit beiden Gewichten ist aus heutiger Sicht natürlich völlig verfehlt. Alle sog. "superlativen" Indexformeln im Sinne von Diewert (also die Formeln von Fisher, Törnqvist, und Walsh) berücksichtigen Mengen q_0 und q_t in einer symmetrischen Weise.

³¹ Diese Inkonsistenz hängt mit der von Flaskämper geradezu als fixe Idee verfolgten Überzeugung zusammen, dass ein Index eine Messzahl von Mittelwerten sein muss, nicht aber ein Mittelwert von Messzahlen.

- Laspeyres seine Formel durchaus auch mit Überlegungen (wie der Analogie zum Experiment) rechtfertigte, was eine Zeitreihe *vergleichbarer* Indexwerte widerspiegeln sollte und was nicht,³² dass man aber auch "pure" beim Vergleich von Preisindizes P_{01} , P_{02} , ... ganz anders verstehen kann und dass
- auch wenn keine Zeitreihe (mit $t = 1, 2, \dots$), also keine Folge P_{01} , P_{02} , ..., sondern "nur" ein binärer (bilateraler) Vergleich von zwei Perioden (0 und t) oder zwei Ländern (A und B) mit P_{0t} oder P_{AB} durchzuführen ist die Forderung des reinen Preisvergleichs nicht ausreichend spezifiziert ist, um daraus eine Indexformel bestimmen zu können.

Auf den zuerst genannten Punkt kann hier nur kurz eingegangen werden. Ein Beispiel hierfür ist die in v. d. Lippe (2010) dargestellte Kontroverse zwischen Werner Neubauer, der gegen Kettenindizes argumentierte und der der FSS zugerechnet wird³³ und Marco Martini, der versuchte, Kettenindizes als *einzig logisch vertretbare Methode* des Vergleichs darzustellen,³⁴ ausgehend von seiner Theorie der indirekten Vergleichbarkeit von direkt nicht Vergleichbarem, die er mit einem Schemas nach Art der folgenden Abb. 1 visualisierte.

Abb. 1: *Indirekter Vergleich der Zustände $t = 0$ und $t = 3$*



Wenn danach Verbrauchsstrukturen in 0 und 3 nicht "direkt" vergleichbar (weil völlig disjunkt) sind, dann können sie nach Martini gleichwohl "indirekt" verglichen werden, sofern es nur Überlappungen in jeweils zwei aufeinander folgenden Perioden gibt. Dabei ist nichts darüber gesagt, wie lang die Kette sein darf³⁵ und wie groß die Überlappungsbereiche mindestens sein müssen.³⁶

Man kann, wie man daran sieht, sich nicht nur auf völlig verschiedene Theorien der Vergleichbarkeit berufen, es ist auch sehr fraglich, ob die "Logik des Vergleichs" (welche auch immer es sein mag), oder generell die "Sachlogik" überhaupt so konkrete Folgerungen erlaubt, dass man mit ihr eine Indexformel (oder allgemein eine statistische Methode) beurteilen kann.

³² Er hielt nicht wie oft gemeint wird, einfach nur aus praktischen Gründen (weil aktuelle Gewichte nicht verfügbar sind) an konstanten Gewichten fest, oder weil sich zu seiner Zeit die Verhältnisse nicht so rasch änderten wie heutzutage (nach dem Motto: hätte er heute gelebt, dann hätte er anders über seine Formel gedacht).

³³ Es ist bemerkenswert, dass seine verbalen Feststellungen nach Art der FSS wie "Die Befürworter des Laspeyres-Index geben auf der Ebene des Idealtypus die Leitidee des 'reinen Preisvergleichs' als verbindlich vor. ... Die Befürworter des Kettenindex lassen sich im Regelfall auf eine Diskussion des Idealtypus der zu messenden Preisniveauveränderung nicht ein. Es dominiert die vordergründige Berufung auf das 'aktuelle' Gewichtungsschema für die Messung der aktuellen Preisniveauveränderungen" (Neubauer 1995) zwar richtig sind, aber kaum verstanden wurden und vor allem die Befürworter von Kettenindizes nicht umstimmte. Neubauer berief sich auch auf "die methodenleitende bzw. sachlogische Interpretation" und konnte sich damit nicht gegen Martini durchsetzen der mit einer Fülle von Martini-spezifischen Definitionen und Axiomen beeindruckte.

³⁴ Auch hier wurde wie bei der FSS mit "Logik" argumentiert, was inhärent dogmatisch ist: implizit wird damit dem Gegner mangelndes Denkvermögen vorgeworfen, weil er angeblich "zwingende" Schlüsse nicht zieht.

³⁵ L. v. Bortkiewicz 1932, der kein Befürworter von Kettenindizes war, wies mit Recht darauf hin, dass diese Logik einem Index zum "ewigen Leben" verhilft.

³⁶ In v. d. Lippe 2001 und 2010 haben wir darauf hingewiesen, dass es sich beim direkten und indirekten (durch Verkettung) Vergleich nicht um die gleiche Art "Vergleich" handelt. Das ist schon daran zu sehen, dass der direkte Vergleich nie pfadabhängig ist und dass auch niemand behauptet, dass er über ein beliebig langes Intervall möglich ist.

Was reiner Preisvergleich beinhalten könnte, mag am besten durch Vergleich einer Folge von Laspeyres (P^L) und Paasche (P^P) Indizes deutlich werden. Bei P^L

$$(2) \quad P_{01}^L = \frac{\sum p_1 q_0}{\sum p_0 q_0}, \quad P_{02}^L = \frac{\sum p_2 q_0}{\sum p_0 q_0}, \quad P_{03}^L = \frac{\sum p_3 q_0}{\sum p_0 q_0}, \dots$$

unterscheiden sich aufeinander folgende Werte nur durch die Preise, während sie sich in der Folge der P^P Indizes

$$(3) \quad P_{01}^P = \frac{\sum p_1 q_1}{\sum p_0 q_1}, \quad P_{02}^P = \frac{\sum p_2 q_2}{\sum p_0 q_2}, \quad P_{03}^P = \frac{\sum p_3 q_3}{\sum p_0 q_3}, \dots$$

auch durch beständig andere Mengen q_1, q_2, q_3, \dots unterscheiden, so dass sich die P^P Indizes (anders als die P^L Indizes) streng genommen nicht untereinander, sondern nur jeweils mit der Basisperiode 0 vergleichen lassen. Man kann daraus folgern, dass nur P^L , nicht aber P^P einen reinen Preisvergleich darstellt.

Dem wird allerdings in der dominierenden anglo-amerikanischen Indextheorie (z.B. von Diewert) vehement widersprochen. Aus dieser Sicht stehen nämlich P^L und P^P "logisch" auf der gleichen Stufe³⁷ und es wird deshalb auch gerne eine Mittelung der beiden Formeln empfohlen, z.B. mit einem ungewogenen geometrischen Mittel, wie beim bekannten "Idealindex" von Fisher³⁸

$$(4) \quad P_{0t}^F = \sqrt{P_{0t}^L P_{0t}^P}.$$

Dabei wird argumentiert, dass sich P^L und P^P nur durch Vertauschung der Rolle von q_0 und q_t unterscheiden. Vom Standpunkt des reinen Preisvergleichs stehen die beiden Indizes, wie ein Vergleich von (2) und (3) zeigt, aber keineswegs auf der gleichen Stufe, und es sind auch die Perioden 0 und t nicht von gleicher Qualität, denn 0 steht für eine über eine gewisse Zeitspanne konstante Periode, während t die Werte 1, 2, ... annimmt, bis wieder ein neues Basisjahr 0 definiert wird.

Entsprechend ist q_0 von anderer Qualität als q_t und Laspeyres hat sich wohl deshalb für q_0 entschieden. Es gibt allerdings kaum dezidierte Äußerungen von ihm, aus denen hervorgeht, warum er q_0 und nicht q_t (wie Paasche) als Gewichte wählte. Neben formalen Gründen, die sich aus seiner Kontroverse mit Drobisch ergaben, war es wohl der Gedanke, dass eine rechnerische Konstanz ein Experiment (*ceteris paribus*) simuliert, und dass Vergleichbarkeit (man

³⁷ Diese Vorstellung ist jenseits des Atlantiks schon lange dominierend, nicht erst seit Bruce Mudgetts (1951) einflussreichem Lehrbuch, in dem dies wieder und wieder betont wird. So sagte z.B. Walsh als Diskutant in Fisher 1921; S. 538 "Nothing can be offered in proof of the superiority of the one over the other". Die gegenteilige Auffassung (wie z.B. unsere), die hier einen Unterschied sieht wurde seinerzeit zwar auch in den USA geäußert, aber entsprechende Arbeiten werden kaum zitiert und sind in Vergessenheit geraten. Ein Beispiel hierfür ist Meeker 1921, der "perfect reversibility" und Fishers ideal index ablehnte ("Is it an index of price changes? Or is it an index of quantity changes? The answer seems clear that it is both and neither", S. 912). Ähnlich später auch Pfouts (zitiert in v. d. Lippe 2007, S. 362ff.). Auch W. Persons (1921) unterschied deutlich zwischen dem isolierten Zwei-Perioden-Vergleich (0 mit t) und der Betrachtung einer Zeitreihe ($t = 0, 1, \dots$) und er argumentierte für eine größere Berechtigung von P^L (gegenüber P^P) im Sinne des reinen Preisvergleichs.

³⁸ Die Position in Flaskämper 1928, S. 127 ist höchst widersprüchlich. Auch er hält P^L und P^P (und – wie gesagt – die Gewichte q_0 und q_t) für gleich gut begründet ("logisch gleichberechtigt" Flaskämper 1929a, S. 158). Andererseits hält er aber im gleichen Satz eine Mittelung von P^L und P^P für "nicht statthaft" und verwirft in seinen Schriften wiederholt vehement den Index P^F . Er hält zwar die laufende Veränderung des Nenners in (3) für einen Mangel von P^P und nennt diesen Index deshalb nur eine "uneigentliche Indexziffer" (S. 110), sieht darin aber nicht eine Verletzung des reinen Preisvergleichs und offenbar auch keinen relevanten Unterschied zu (2), also auch keinen Grund, P^L der Formel P^P vorzuziehen.

sprach zu seiner Zeit noch nicht von der "Logik des Vergleichs") dies erfordert, so wie auch Art und Qualität der Güter konstant gehalten werden müssen:

Um "den Charakter der Bewegung kennen zu lernen muß man nicht vorwärts, sondern lieber rückwärts schauen und diejenigen Objecte aussuchen, welche ausnahmsweise eine lange Zeit in vergleichbarer Qualität producirt wurden", und die "in die *Vergangenheit* recht weit zurückverfolgt werden können" (Laspeyres 1875, S. 18)

Man kann die Konstanz von q_0 in der Folge von P^L -Indizes als Inbegriff des reinen Preisvergleichs betrachten. Das ist die Position, die der Verfasser in v. d. Lippe 2001 und v. d. Lippe 2005 vertreten hatte, und dieses Verständnis von reinem Preisvergleich lag auch der in Deutschland seinerzeit ziemlich einhelligen Ablehnung von Kettenindizes zugrunde, also von Indizes, die gerade nicht *nur* von p_0 und p_t , sondern auch von p_1, p_2, \dots sowie von q_0, q_1, \dots bestimmt werden. Man kann aber auch (wie auch hier wieder Diewert), ausgehend von der "logischen Gleichberechtigung" von q_0 und q_t (und damit P^L und P^P) den Walsh Index P^W mit

$$(5) \quad P_{01}^W = \frac{\sum p_1 \sqrt{q_0 q_1}}{\sum p_0 \sqrt{q_0 q_1}}, \quad P_{02}^W = \frac{\sum p_2 \sqrt{q_0 q_2}}{\sum p_0 \sqrt{q_0 q_2}}, \quad P_{03}^W = \frac{\sum p_3 \sqrt{q_0 q_3}}{\sum p_0 \sqrt{q_0 q_3}}, \dots$$

als beste Option für einen "pure index" betrachten. Wie man sieht, unterscheidet sich diese Folge von (2) durch die Wurzel einer Mengenzahl $\omega_t = \sqrt{q_t/q_0}$

$$(5a) \quad P_{01}^W = \frac{\sum p_1 q_0 \omega_1}{\sum p_0 q_0 \omega_1}, \quad P_{02}^W = \frac{\sum p_2 q_0 \omega_2}{\sum p_0 q_0 \omega_2}, \quad P_{03}^W = \frac{\sum p_3 q_0 \omega_3}{\sum p_0 q_0 \omega_3}, \dots$$

die nicht konstant ist, sondern mit t variiert.

Wie wenig spezifisch das Postulat des reinen Preisvergleichs ist, ergibt sich schon allein daraus, dass danach im binären Vergleich auch ungewogene Preisindizes, wie

$$(6) \quad P_{0t}^C = \frac{1}{n} \sum_i \frac{p_{it}}{p_{i0}} \quad (\text{Carli}) \text{ oder}$$

$$(7) \quad P_{0t}^J = \sqrt[n]{\prod_i \frac{p_{it}}{p_{i0}}} \quad (\text{Jevons})$$

zu empfehlen wären, weil hier die Ergebnisse auch nur durch die Preise in 0 und t bestimmt werden und keine Mengenbewegungen (Unterschiedlichkeit von q_0 und q_t) den somit "reinen" Preisvergleich stören. Das ist nicht nur unbefriedigend, weil es gute Gründe gibt, gewogene Indizes den ungewogenen vorzuziehen (und auch andere Eigenschaften von Indexfunktionen von Interesse sind). Hinzu kommt auch, dass nach diesem Verständnis von "rein" auch *alle* ungewogenen Indizes (nicht nur P^C und P^J) *gleich gut* wären.

Aus allem folgt, dass man mit einer "Logik des Vergleichs" bei der Entwicklung und Beurteilung von statistischen Methoden nicht weit kommen kann. Man kann vielmehr zu Widersprüchlichkeiten und Fehlern kommen, und das ganz besonders dann, wenn noch Vorbehalte gegen Formalismus und Mathematik (oder unnötig viel Mathematik, wie immer man das definieren mag) hinzukommen, die in den Überzeugungen der FSS angelegt sind.³⁹

³⁹ Dass bei Flaskämper so viele mathematische Fehlleistungen zu beklagen sind, ist sicher kein Zufall, denn die Überzeugungen der FSS sind ja geradezu dazu angetan, dem Verächter und Ignoranten in puncto Mathematik ein Minderwertigkeitsgefühl zu nehmen: "Es ist deshalb auch eine völlige Verkennung der Dinge, wenn der Vertreter der reinen Mathematik auf die vom Standpunkt der Mathematik primitiven Verfahren des Sozialstatistikers mit Geringschätzung herabsieht, wie es ebenso unberechtigt ist, wenn der praktische Sozialstatistiker ... Minderwertigkeitsgefühle empfindet." (Flaskämper 1950, S. 160)

Selbst bei einer auf dem ersten Blick so einleuchtenden und berechtigten Forderung, wie die des "reinen Preisvergleichs",⁴⁰ die ja viel konkreter ist als eine bloße Beschwörung der "Sachlogik" zeigt sich auf dem zweiten Blick, dass sie schwer exakt zu fassen ist⁴¹ und als alleiniges Beurteilungskriterium für Indexformeln nicht ausreichen kann. Es gibt auch legitime andere Anforderungen an eine Indexformel. Und es ist Skepsis geboten, wenn angeblich "logisch" "zwingende" Folgerungen aus einer Theorie des Vergleichs gezogen werden. Das Beispiel von Martinis zeigt, dass dies auch dann gilt wenn der Autor einer solchen Theorie keine Vorbehalte gegen Formalismus hat und im Gegenteil mit Mathematik zu brillieren versucht.

c) Indizes und Messzahlen,

Mittelwerte von Messzahlen und Messzahlen von Mittelwerten

Eine als Ausgangspunkt in den Raum gestellte Grundüberzeugung Flaskämpfers war auch, dass ein Vergleich eine Verhältniszahl (wie z.B. eine Messzahl) erfordert und ein Mittelwert dagegen nicht die Aufgabe hat, etwas zu vergleichen, sondern einen typischen "Repräsentanten" für alle Einheiten einer Masse darzustellen. Ein solches teleologisches Denken, das einer beschreibenden Kennzahl (statistic) eine ihr spezifische Aufgabe, ein Ziel zuweist, mag nahe-liegend sein, wenn es um "Verstehen", "Sinnebung" usw. geht.⁴² Das Kategorisieren nach Zielen führte bei Flaskämper zu Aussagen, wie folgende

- dass kein Unterschied zu machen ist zwischen einer Preismesszahl (des Guts i) p_{it}/p_{i0} und einem Preisindex, bei dem nur an die Stelle von p_{it} und p_{i0} entsprechende Mittelwerte treten⁴³ (womit schlicht alle Probleme der Aggregation über n Güter ignoriert werden), und
- dass bei "relativen Zahlen", "Veränderungsgrößen", oder "dynamischen Größen" im Unterschied zu Zustandsgrößen das geometrische, nicht das arithmetische Mittel genommen werden muss;⁴⁴ was aber schon von Winkler 1930, S. 654 kritisiert wurde.

Der erste Punkt beinhaltet die abwegige Vorstellung, ein Index müsse alle Eigenschaften haben, die eine Messzahl schon per Konstruktion hat. Das ist eine besonders folgenreiche falsche Weichstellung gleich zu Beginn seiner Indextheorie; denn aus ihm folgerte er, dass Zeit- und Faktorkehrbarkeit ("Multiplikationssatz") oder gar Transitivität (circular test) usw. "bei einem echten Kollektivindex ebenso eine Selbstverständlichkeit wie bei einem Einzelin-

⁴⁰ Ich stimme mit Neubauer (1995) darin überein, dass es bei der Frage, ob man für oder gegen Kettenindizes ist im Grunde nur darum geht, für wie wichtig man den reinen Preisvergleich hält.

⁴¹ Zu einem entsprechenden Konkretisierungsversuch vgl. v. d. Lippe (2005).

⁴² Typisch für diese Art zu denken ist "Jedes in der Statistik angewendete Rechenverfahren ... muß einen sachlich-anschaulichen Sinn haben ... Für jede Situation gibt es deshalb strenggenommen, nur ein Verfahren" (Flaskämper 1931, S. 381).

⁴³ Das ist die bereits zitierte Aussage, dass sich Indizes und Messzahlen (Einzelindizes) "logisch in nichts ... unterscheiden" Flaskämper 1928, S. 118.

⁴⁴ Flaskämper 1927, S. 648 und 1928, S. 164f., 187 und 1931, S. 387, 399. Er glaubt dies "bewiesen" zu haben durch Mittelung von Wachstumsfaktoren p_{i1}/p_{i0} , p_{i2}/p_{i1} , p_{i3}/p_{i2} , ... ($i = 1, 2, \dots, n$) über die Zeit $t = 0, 1, 2, \dots, T$ und er sieht nicht den Unterschied zwischen der Mittelung über T Intervalle und der bei Indizes geforderten Mittelung über n Waren. Andererseits sei das geometrische Mittel von Mengen, wie in der Formel P^W nicht zu rechtfertigen (Flaskämper 1928, S. 109). Ähnlich zweifelhafte Aussagen Flaskämpfers über die Sinnhaftigkeit sind auch: Während das arithmetische Mittel immer dann (nur dann?) sinnvoll sei, wenn auch die Summe Sinn hat, hat für ihn das harmonische Mittel "keine oder kaum eine in der Statistik in Frage kommende Bedeutung" (Flaskämper 1931, S. 383) und ob ihm "eine sachlich-anschauliche Bedeutung entspricht, ist sehr zweifelhaft" (ebenda S. 392). Auch das ist schlicht nur Unsinn, zumal alle drei Mittelwerte Spezialfälle des "power mean" sind und es zumindest im Ausland schon zu Flaskämpfers Zeiten genügend Ausführungen über die Interpretation der drei Mittelwerte gab (zu nennen wäre z.B. Walsh 1901, ein Buch, das sich ausschließlich mit Indexzahlen beschäftigte und offenbar von Flaskämper bei seiner "Theorie der Indexzahlen 1928 völlig ignoriert wurde). Siehe auch Walsh 1901, S. 235 ff speziell zur verbalen Interpretation des harmonischen Mittels.

dex" sei (Flaskämper 1928, S. 119), und deshalb die Betrachtung von Axiomen überflüssig sei.

Faktorumkehrbarkeit ist für ihn "mit dem logischen Wesen der Indexzahlen und des Vergleichs überhaupt zwangsläufig verknüpft" (S. 61), und Zeitumkehrbarkeit hält er für "einen elementaren Sachverhalt, der mit dem Begriff des statistischen Vergleichs ohne weiteres gegeben ist" (S. 60), und dass Fisher davon abgerückt ist, den circular test (Transitivität) zu fordern ist für ihn ganz unverständlich (S. 57f, 122). Fisher habe sich durch "pseudologische Betrachtungen" damit "in verhängnisvolle Widersprüche verwickelt" (S. 60).

Eine zweite Folgerung wurde schon von seinen Zeitgenossen heftig kritisiert, nämlich dass ein Index "nicht ein irgendwie berechneter Durchschnitt von Einzelindices, sondern der Index des Durchschnitts" sein müsse (S. 79) und dass "Mittelwerte von Einzelindices" nur "Pseudoindices" seien (S. 118), denn so etwas "ist logisch keine Vergleichsgröße, sondern ein Durchschnitt, der sich auf einen Zeitpunkt bezieht" (S. 159). Dass sich die gemittelten Größen p_{it}/p_{i0} auf zwei Zeitpunkte beziehen macht ihn nicht skeptisch, weil – teleologisches Denken! – ein (zeitlicher) Vergleich eben nicht die Aufgabe eines Mittelwerts ist.

Es ficht ihn auch nicht an, dass nur wenige Indizes alle diese Forderungen erfüllen und z.B. auch P^F nicht transitiv ist. Er ist sogar der Meinung, dass P^L und P^P Zeit- und Faktorumkehrbarkeit erfüllen, wenn man diese "Kriterien" richtig versteht (S. 142)⁴⁵ und es überrascht auch nicht, dass er Kettenindizes für durchaus sinnvoll hält (S. 31).⁴⁶

Offensichtlicher Unsinn ist auch die von Flaskämper unermüdlich vorgetragene (und schon von seinen Zeitgenossen kritisierte) fixe Idee, dass die Gleichung

$$(8) \quad P_{0t}^L = \frac{\sum_i p_{it} q_{i0}}{\sum_i p_{i0} q_{i0}} = \sum_i \frac{p_{it}}{p_{i0}} \frac{p_{i0} q_{i0}}{\sum_i p_{i0} q_{i0}}$$

nicht zeige, dass ein Verhältnis von Ausgaben (und damit auch von Durchschnittspreisen) auch darstellbar ist als gewogenes Mittel von Preismesszahlen. Er wollte dies nur als ein Sonderfall verstanden wissen⁴⁷ und akzeptierte die linke Seite ($\sum p_t q_0 / \sum p_0 q_0$), lehnte aber die rechte Seite ab, weil er Ausgabenanteile $p_0 q_0 / \sum p_0 q_0$ für unzulässige Gewichte hielt. Damit sah er sich in der glücklichen Lage, Gl. 8 zwar als "zahlenlogisch" richtig anerkennen zu müssen, sie aber "sachlogisch" ablehnen zu können.⁴⁸ Wenn die Unterscheidung zwischen "Zahlenlogik" und "Sachlogik" überhaupt fruchtbar sein soll (einmal angenommen, man könne diese Begriffe hinreichend exakt definieren), dann müsste es genau so etwas geben, dass nämlich eine Gleichung nach der einen Logik richtig, nach der anderen aber falsch ist. Dass dies für uns

⁴⁵ Andererseits sieht er, dass P^L die Zeitumkehrbarkeit im Sinne von Fisher nicht erfüllt und er nennt deshalb den sonst auch von ihm sehr gelobten Index P^L nur eine "unechte oder Pseudoindexziffer" (S. 61). Andererseits ist aber auch P^P – wie gesagt (Fußnote 38) – nur eine "uneigentliche Indexziffer". Das alles beeinflusste aber sein Gesamturteil nicht, denn sein Fazit ist: Fisher wendet seine Kriterien (Axiome) "ganz schematisch an, ohne deren logischen Sinn zu Ende zu verfolgen" und sowohl P^L als auch P^P erfüllen alle *richtig verstandene* Kriterien (S. 144f).

⁴⁶ Flaskämper hält auch P^L und P^P für verkettbar und die richtige gegenteilige Aussage von L. v. Bortkiewicz für ein "auf Schematismus beruhendes Missverständnis" (1928, S. 129). Flaskämper betrachtete v. Bortkiewicz (neben Lexis) als Hauptrepräsentant der anderen, mathematischen Statistik und es ist bezeichnend, dass er bei abweichenden Ansichten regelmäßig schief lag und zu konträren Ansichten von v. Bortkiewicz fiel ihm meist nicht mehr ein als "Schematismus", "Formalismus" usw.

⁴⁷ Er glaubte, dies träfe nur beim arithmetischen Mittel und bei Ausgabengewichten $p_0 q_0 / \sum p_0 q_0$ zu (S. 181). Er schloß dies u.a. auch daraus, dass es ihm nicht gelang, den Paasche Preisindex als arithmetisches Mittel der Preismesszahlen mit den Gewichten $p_t q_t / \sum p_t q_t$ herzuleiten (Flaskämper 1927b, S. 647). Er ist offenbar nicht auf die Idee gekommen, hier das harmonische Mittel zu nehmen.

⁴⁸ Offenbar hatte Flaskämper keine Probleme, Gl. 8 "sachlogisch" abzulehnen. Er tat dies nicht nur mit dem Hinweis auf angeblich unzulässige Gewichte, sondern auch ganz pauschal: die Veränderung eines Durchschnitts und der Durchschnitt einer Veränderung "sind also logisch ganz verschiedene Begriffe" (S. 160).

heutzutage jedoch schwer nachvollziehbar ist zeigt, dass etwas im Argen liegt mit den drei Dichotomien der FSS (zwei Arten von Wissenschaft, von Statistik und von Logik).

Ein Sonderfall ist Gl. 8 keineswegs, was ja auch Flaskämper schon bei der Lektüre von Fishers "Making of Index Numbers" hätte sehen können. Viele Indizes erlauben beide Darstellungsarten, was v. Bortkiewicz (1927, S. 747) "Zwieförmigkeit" nannte und als ein denkbare Gütekriterium (Axiom) verstand, was wiederum Flaskämper als "ein völlig willkürlich und formal konstruiertes, also nicht aus der Logik des Vergleichs abgeleitetes Kriterium" abtat (1928, S. 182 und 1929a, S. 158).

Flaskämper sträubte sich – wie gesagt – auch wegen der Gewichte gegen Gl. 8. Richtige Gewichte konnten für ihn nur Mengenteile $q_{i0}/\Sigma q_{i0}$ sein (obgleich Σq_{i0} oft nicht zu definieren ist). Für ihn waren "Gewichte" offenbar nur vorstellbar als Häufigkeiten,⁴⁹ und Mittelwerte mussten (dynamischen Größen!) geometrische Mittel sein. Er hielt deshalb den Index

$$\sqrt[S]{\prod_i \left(\frac{p_{it}}{p_{i0}} \right)^{q_i}} \quad \text{mit } S = \Sigma q_i \text{ für erwägenswert, nicht aber } \sum_i \frac{p_{it}}{p_{i0}} \cdot \frac{q_i}{\sum q_i} \quad (\text{S. 177f}),$$

d) Bilder als Argumente und Beschwörung der "Logik"

Wir wollen es mit dieser Darstellung von Irrtümern bewenden lassen (man könnte sie noch erheblich verlängern) und auf ein weiteres Problem mit der Beschwörung der Sachlogik hinweisen. Die zuletzt wiederholt genannte Idee, das geometrische Mittel habe mit Dynamik zu tun scheint nicht nur eine Folge von dem zu sein, was wir "teleologisches Denken" in Bezug auf mathematische Formeln nannten, sondern auch Ausdruck eines Denkens in Bildern und Analogien. Man sieht dies an den Wortschöpfungen, die im Zusammenhang mit der Messung der "Dynamik" auftauchen, die angeblich Sache des geometrischen, nicht des arithmetischen Mittel sei. So ist dort z.B. die Rede von einem Maß der "Veränderungswucht" und (davon verschieden?) der "Veränderungsintensität" (Flaskämper 1928, S. 160).

Bilder werden auch als Argument gegen Formeln verwendet. So ist der Idealindex P^F von Fisher für Flaskämper ein "logisches Bastardprodukt, das keinem logisch eindeutigem Sachverhalt entspricht" (S. 197)⁵⁰ und er begründete dies mit einer statistisch höchst zweifelhaften Aussage "Man kann eben nicht zwei Vergleiche in einen zusammenziehen und zwei oder mehrere Erscheinungen ... gleichzeitig vergleichen wollen" (S. 108).⁵¹

Überhaupt sei ein Mittel aus P^L und P^P "logisch" eine Unmöglichkeit" (S. 107). Das Ergebnis der Verschmelzung von P^L und P^P sei "ein unlogischer Zwitter" (S. 197) und Flaskämper vergleicht folgerichtig Fishers "crossing of formulae"⁵² mit der Erzeugung von Artenvielfalt durch Züchtung und er vermisst ein klares "Zuchtziel" (S. 197).

Aus heutiger Sicht ist es überraschend, dass seinerzeit Vokabel wie "Bastardprodukt" als Argumente in einer wissenschaftlichen Auseinandersetzung akzeptiert wurden.⁵³ Das Beschwö-

⁴⁹ Sinnvoll als Gewichte sind allein "... die umgesetzten Mengen, nicht die umgesetzten Werte. Das geht einfach aus dem Begriff der Gewichte hervor, demzufolge diese die Zahl bedeuten, die angibt, wie oft ... der betreffende Preis, Lohn usw. berücksichtigt werden muß" (1928, S. 180).

⁵⁰ Einmal "logisch" hätte vielleicht auch schon gereicht.

⁵¹ Dass Flaskämper die Mittelung von Wägungsschemata (q_0 und q_i) akzeptiert, nicht aber die von Formeln P^L und P^P (obgleich sich P^L und P^P ja genau nur in diesen Gewichten unterscheiden) hängt wohl damit zusammen, dass er einer Formel eine Rolle, Funktion, "Aussage" beimisst, was wir teleologisches Denken nannten, einem Wägungsschema dagegen nicht. In der bildlichen Sprache geht es hier vielleicht um Organ vs. Organismus.

⁵² Fisher nutzte bekanntlich den time- und factor reversal test, sowie das crossing als "finders of formulae".

⁵³ Man findet auch in Schriften der damaligen Zeit diese Art der Argumentation nicht überall, sondern eher selten. Der von mir sehr geschätzte L. v. Bortkiewicz hat sich m. W. nie so ausgedrückt, obgleich er als sehr streitbarer Geist bekannt und berüchtigt war.

ren der "Logik" ist demgegenüber zwar im Stil, aber nicht in der Sache sehr viel besser.⁵⁴ Dabei dürfte der exzessive Gebrauch des Wortes "Logik" im Zusammenhang mit der (Deskriptiven) Statistik überhaupt deplatziert sein. Von "Logik" zu sprechen setzt voraus, dass über den "Wahrheitswert" (wahr/falsch) entschieden werden kann. Aber gibt es eine wahre und eine falsche Streuung, einen wahren und einen falschen Korrelationskoeffizient usw.?⁵⁵ Es ist zu vermuten, dass "logisch" im Sprachgebrauch der FSS, oder speziell von Flaskämper, eher gemeint sein dürfte im Sinne von verständlich, einleuchtend, (intuitiv) "überzeugend", plausibel, anschaulich, oder von etwas, was ein Gefühl vermittelt unbezweifelbar und sinnvoll zu sein.

Wie stark bei diesem Verständnis von Statistik eine Gefühlsdimension angesprochen ist, macht auch das folgende Zitat deutlich, in dem Flaskämper eine Methode völlig unangemessen würdigt: "Z.B. ist die Korrelationsrechnung doch weiter nichts als eine Fortsetzung dessen, was gefühlsmäßig auch der nichtmathematische Statistiker betreibt, wenn er die verschiedene Strenge des Zusammenhangs zwischen je zwei Zahlenreihen mehr oder weniger gefühlsmäßig auf Grund seines Zahleninstinkts oder durch das Augenmaß (wenn es sich um Kurvenpaare handelt) beurteilt ... Was tut aber die Korrelationsrechnung, die bis vor kurzem noch viele Statistiker mieden als etwas unnötig Kompliziertes, anderes, als daß sie an die Stelle dieser gefühlsmäßigen Vorstellungen ... ein eindeutiges und sicheres Maß setzt." (Flaskämper 1936, S. 14).

Auf ähnlich alberne Art wird auch das Gesetz der großen Zahl als trivial hingestellt. Flaskämper 1927a zog nach einer sehr stümperhaften verbalen Beschreibung dessen, was es nach seiner Meinung besagt,⁵⁶ die er aber als "logische Begründung des Gesetzes" verstand, den lapidaren Schluss: "Allerdings ist hiermit keine Einsicht gewonnen, die uns nicht auch schon der instinktive, gefühlsmäßig urteilende, also nicht mathematisch gebildete Verstand sagt" (S. 507).⁵⁷

Nachdem einer Formel ein Ziel (z.B. Vergleichsmaß, nicht Durchschnitt) und ein Bild (z.B. Bastard) zugeordnet wurde, haben wir es jetzt mit einem "Gefühl" zu tun. Es ist zu vermuten, dass hier "Logik" als Sammelbegriff für alle diese höchst unangemessene Arten, eine Formel, nicht nur eine Indexformel zu beurteilen, gebraucht wird. Wir verlassen damit die Ebene "Indextheorie" und kommen zurück zur Statistik ganz allgemein und fragen uns, was die FSS auf dieser Ebene geleistet hat?

⁵⁴ Sie war aber seinerzeit sehr verbreitet. Ein auffallendes Beispiel hierfür war Max Weber.

⁵⁵ Man kann von einem "wahren" Parameter (der Grundgesamtheit) und einer *geschätzten* Größe, der Schätz- oder Stichprobenfunktion für diesen Parameter sprechen, aber mit logisch "wahr" oder "falsch" hat das nichts zu tun.

⁵⁶ Danach ist es z.B. einmal eine Wahrscheinlichkeit (S. 506) und dann (auch?) eine Genauigkeit (S. 507), die mit dem Faktor $n^{1/2}$ wächst (womit es übrigens schnell den Wert 1 überschreiten würde, was man natürlich nicht sah). "Seinen ursprünglichen logischen Ort hat es in der Wahrscheinlichkeitsrechnung" und für diese sei die "sogenannte apriorische oder theoretische Wahrscheinlichkeit", die es aber in der sozialwissenschaftlichen Statistik nicht gäbe, der Ausgangspunkt, so dass u.a. auch deshalb das Gesetz dort nicht anwendbar sei. Und außerdem habe man in der *sozialwissenschaftlichen* Statistik ein anderes Erkenntnisziel. Nur an einem Platz hat Stochastik ein Existenzrecht: beim Ziehen einer Stichprobe.

⁵⁷ Man sprach seinerzeit auch von einem Unterscheid zwischen dem statistischen und dem mathematischen Gesetz der großen Zahl, oder z.B. auch von einer "inneren" (nicht berechenbaren) Wahrscheinlichkeit, auf die sich z.B. ein Kriminalist berufe, (so etwa noch nach dem Zweiten Weltkrieg Blind 1953, S. 310), die zwar anders, aber genauso wertvoll sei, wie die mathematische und die z.B. Winkler sehr an das oben angesprochene "Fingerspitzengefühl" erinnerte (Boustedt 1953, S. 346). Blind erwiderte darauf, dass er diesen angeblich allgemeineren Wahrscheinlichkeitsbegriff nicht reduziert wissen wollte "auf das Gebiet der Wahrscheinlichkeitsrechnung, sondern ... ihn darüber hinaus ausdehnen" wollte (nach Boustedt 1953, S. 358).

3. Die Philosophie der Frankfurter Schule und ihre Folgen

Vieles an den Auffassungen der FSS findet auch heute noch bei Statistikern (insbesondere aus den Ämtern) Anklang. Sie werden oft als ein notwendiger Blick über den Tellerrand der Mathematik hinaus gesehen und als solches begrüßt. Neben der Fokussierung auf "Vergleichbarkeit" scheint auch die Betrachtung des "Adäquationsproblems" auf dem ersten Blick ein nur zu berechtigtes Anliegen der FSS zu sein. Auf dem zweiten Blick ist jedoch auch hier Skepsis angebracht, wenn man sich fragt, was das Nachdenken hierüber gebracht hat oder bringen könnte, und ob dies überhaupt Themen sind, mit denen die Existenz einer eigenen (quasi zweiten) Art von Statistik gerechtfertigt werden kann. Während die "Adäquation" als Thema einer Art "Theorie der Wirtschaftsstatistik" auch in unserer Zeit viele fasziniert, verfangen Begriffe, wie "Wesen" und "Sachlogik" oder auch die Vorurteile gegenüber der Mathematik inzwischen kaum noch. Aber weil hierin die Basis für alles weitere zu suchen ist, sollten wir auch hiermit beginnen. Denn es ist auch die Basis für eine Neigung der FSS (zumindest in der Zeit bis etwa 1960), vieles an ausländischen, aber auch deutschen Arbeiten zu ignorieren, was aus der anderen Richtung in der Statistik, der mathematischen kam, die man für den kleineren und weniger relevanten Teil der Statistik hielt.

a) Was ist und was leistet die "Sachlogik"?

Nach der Auffassung von Flaskämper (die seinerzeit wohl die herrschende Lehre war) kann die "Sachlogik" speziell in Gestalt der oft als deutsche Besonderheit empfundenen "Logik des Vergleichs" mindestens das leisten, was (z.B. in der Indextheorie) Axiome leisten. Im zweiten Teil haben wir aber gesehen, dass nicht nur kein Beweis für diese verwegene These geliefert wurde, sondern dass sie stattdessen eine Quelle für krasse Widersprüche und Fehler war. Wir wollen nun zeigen, dass auch die Leistungen, die unter Berufung auf "Sachlogik" ganz allgemein erbracht wurden kritikwürdig sind. Das betrifft zunächst einmal

- die Definition von "Sachlogik" und die damit verbundene Einstellung zur Mathematik, und zum anderen
- die Vorstellungen über die Aufgabenteilung zwischen einer primär sachlogisch und einer primär zahlenlogisch orientierten der Statistik.

Statt Definitionen von Zahlen- und Sachlogik gibt es verschiedentlich Aufzählungen von Begriffen, die zu der einen oder der anderen Art von Logik gehören sollen. Danach sollen "sachlogische Begriffe" u.a. sein: "Bestands- und Bewegungsmassen, Vergleichbarkeit, Verhältniszahlen, statistische Trugschlüsse, wirtschaftliche Gesamtzahlen wie Volksvermögen ... Indexzahlen ..., alles Dinge, bei denen schwierige mathematische Gedankengänge keine oder nur eine sehr geringe Rolle spielen" (Flaskämper 1950, S. 153).⁵⁸

Der letzte Halbsatz zeigt deutlich, worum es bei der Unterscheidung eigentlich geht, nämlich um mehr oder weniger Mathematik. Denn zahlenlogischen Begriffe seien demgegenüber "Mittelwerte, Streuungsmasse, Korrelationskoeffizienten usw."⁵⁹ (letztere waren seinerzeit in puncto Mathematik das höchste der Gefühle). Als Aufgabe der Zahlenlogik wird auch ge-

⁵⁸ Ganz ähnliche Auflistungen findet man in Flaskämper 1933, S. 66, 70, Below 1956, S. 80, Blind 1952, S. 534. Um nicht den Eindruck entstehen zu lassen, dass ohne Mathematik nicht viel übrig bleibt, was schwierig ist, findet man Hinweise auf "höchst bedeutsame, logisch schwierige Entscheidungen" (Blind 1952, 533), oder "schwierigste sachlogische Erwägungen" (Flaskämper 1950, S. 158), ohne dass dies sonderlich konkret gemacht wird.

⁵⁹ Flaskämper 1933, S. 66. Weiter heißt es dort: "Ich rechne in letzter Hinsicht zu den sachlogischen Begriffen zunächst die verschiedenen Arten der Verhältniszahlen: die Verhältniszahl als allgemeine rechnerische Operation ist ein zahlenlogischer Begriff allereinfachster Art ... aber die einzelnen Arten der Verhältniszahlen sind sachlogische Begriffe." Mit den Arten sind Gliederungszahlen, Messzahlen usw. gemeint.

nannt "die mathematisch schlüssige Ableitung und Durchführung der rechnerischen Verfahren, die in der Statistik von der einfachsten Rechenoperationen bis zu den komplizierten Verfahren der Trend- und Korrelationsrechnung reichen" (Blind 1952, S. 534f), als ob richtiges Umformen von Gleichungen und richtiges Rechnen überhaupt ein Thema sein kann.

Sehr peinlich und dekuvierend wird es immer dann, wenn es nötig ist, sich im Zusammenhang mit "Zahlenlogik" über Gegenstände der anderen Art von Statistik, nämlich der "mathematischen" zu äußern. Das ist vor allem dann der Fall, wenn es um die Behauptung geht, die mathematische Statistik und insbesondere die Wahrscheinlichkeitsrechnung sei auf die Probleme der Sozialstatistiker nicht anwendbar weil man dort keine "zufallsentstellte Werte" (Blind 1952, S. 533) und auch meist keine Normalverteilung habe.⁶⁰ Es wird auch gern gesagt, sie sei nur in einem kleinen Teilgebiet, z.B. in der Bevölkerungsstatistik oder bei der mengenmäßigen Erfassung der Produktion anwendbar (Below 1956; S. 81). Diese Vorstellung der begrenzten und dienenden Rolle der Mathematik verbunden mit Ignoranz auf diesem Gebiet lässt sich bei Flaskämper und Blind gut belegen, und zwar mit ganz ähnlichen Äußerungen vor und nach dem Zweiten Weltkrieg (wir kommen darauf zurück).

In Flaskämper 1936 heißt es, das Ziel sei nicht ein "System von mathematischen Formeln, sondern ein Gebäude sinnhafter Zusammenhänge" (S. 4) und "Wir werden es vermeiden, mathematische Statistik um ihrer selbst willen zu betreiben, was als Sonderzweig der Mathematik natürlich immer seine Berechtigung hat" (Flaskämper 1936, S. 14). Blind 1952 beschreibt das Bemühen um eine "Einschmelzung der Stochastik in die sozialwissenschaftliche Statistik", die gerade Flaskämpfers beständiges Anliegen gewesen sei (S. 537). Dabei behielt man sich vor, *was* "eingeschmolzen" werden sollte.⁶¹ Die Idee war also, dass der "Sachlogiker" entscheidet, was ihm von der mathematischen Richtung "sinnvoll" erscheint, von einer Richtung, von der er nichts versteht und auch nach seinen Grundüberzeugungen nichts verstehen muss. Er bedient sich der Erkenntnisse aus einem angeblich nur sehr begrenzt relevanten Teilgebiet der Statistik und *er* ist der Chef bei diesem Unternehmen: "Die demgegenüber von der Stochastik vertretene Auffassung von der Statistik als einer ganz allgemein ... mit den Methoden der Wahrscheinlichkeitsrechnung arbeitenden Verfahrensweise ... erscheint dagegen schon wegen ihrer Unvereinbarkeit mit der lebendigen Wirklichkeit unterlegen" (S. 537).

Man beachte, dass dieses Zitat von Blind aus dem Jahre 1952 stammt. Vor dem Zweiten Weltkrieg wurden Vorurteile gegen Mathematik von Seiten der Statistiker noch viel offener geäußert (viel davon war auch von der Sorte Praktiker vs. Theoretiker).⁶²

Man sollte die (inakzeptable) Zuweisung einer dienenden Rollen der mathematischen Statistik unter der Herrschaft der "Sachlogiker" im Auge behalten, wenn man den Umstand würdigt, dass Flaskämper schon vor dem Zweiten Weltkrieg (und natürlich erst recht danach) die lange

⁶⁰ Diese sei angeblich vorauszusetzen, wenn man ein arithmetisches Mittel oder einen Korrelationskoeffizienten berechnen möchte (vgl. Blind 1953, S. 306f, 308f). Auf Flaskämpfers Argumente gegen eine "Überschätzung" des Gesetzes der großen Zahlen wurde bereits hingewiesen. Aus Platzgründen verzichten wir auf weitere entsprechende Zitate.

⁶¹ Nach Blind 1952 plädierte Flaskämper "für eine sorgsame Einschmelzung der dafür geeigneten komplizierten mathematischen, insbesondere auch der stochastischen Methoden in die sozialwissenschaftliche Statistik" (S. 537). Die Maßstäbe der "Eignung" wurden aber nie richtig offen gelegt.

⁶² Johannes Müller, seinerzeit Direktor des Thüringischen Statistischen Landesamts, stellte sogar den Grundsatz auf: "je weniger mathematische Formeln in der Wirtschafts- und Sozialstatistik zur Anwendung kommen, desto größeren Anspruch auf Richtigkeit können die gegebenen Zahlen machen" (Müller 1923/24, S. 445). Den Ausdruck "Begriffsstatistik" wählte er in Anlehnung an "Begriffsjurisprudenz", um die Wirklichkeitsfremdheit dieser Art von Statistik zu bezeichnen. Es findet sich auch die seinerzeit ebenfalls verbreitete Auffassung, das Geschäft des Praktikers, "Erhebungen zu organisieren und über Definitionen zu grübeln" sei "unendlich viel mühseliger ... als vom sicheren Schreibtisch aus elegante Formeln zu entwerfen" (S. 455). Flaskämper und viele andere sprachen auch oft von einer mindestens genauso großen "Dignität" dieses nicht-mathematischen Geschäfts.

"ungebührlich" betriebene Vernachlässigung der mathematischen Statistik in Deutschland für falsch gehalten hat (1933, S. 67).⁶³

Welche Themen sind es nun, die primär in den Kompetenzbereich der "Sachlogik" gehören, und bei denen man glaubt, eigenständige Leistungen erbringen zu können, die von der mathematischen Statistik nicht zu erwarten sind? Flaskämper 1950, S. 159 nennt folgende Aufgaben der Statistik ganz generell (es ist bezeichnend, dass eine Liste speziell der Aufgaben der mathematischen Statistik fehlt):⁶⁴ "... die Herausarbeitung der Begriffe, die das zu untersuchende Objekt erst zählbar machen sollen. Weiter gehören hierher aber auch logische Betrachtungen über den überaus problematischen Begriff der Vergleichbarkeit. Hierher gehören ferner Erwägungen über die sachliche Bedeutung eines mit irgendeinem mathematischen Verfahren gewonnenen Ergebnisses, Erwägungen weiter über die verschiedene sachlogische Bedeutung der einzelnen mathematischen Begriffe und Verfahren, z.B. der verschiedenen Mittelwertbegriffe." Der Sachlogiker erklärt also die Bedeutung der mathematischen Verfahren, und dabei lieferte doch die Philosophie der FSS die Rechtfertigung dafür, dass deutsche Statistiker lange Zeit guten Gewissens die Mathematik links liegen lassen konnten.

Bei welchen Verfahren untersuchte man nun deren "sachlogische Bedeutung"? Seit Zizek waren Betrachtungen über die verschiedenen Mittelwerte beliebt; kompliziertere Maßzahlen wurden m.W. nicht betrachtet. Wenn man schon keine mathematischen Leistungen vollbringen wollte, was konnte man dann mit "logischen" Studien über Mittelwerte vorweisen? Liest man entsprechende Schriften (etwa Flaskämper 1931), so fällt auf, dass die Leistungen vor allem darin bestanden, umständlich (und deshalb z.T. schwer verständlich) in Worten zu sagen, was die Formel ohnehin schon besagt. Man gewinnt den Eindruck, dass "sachlich", "sinnvoll" etc. meist nur einfach im Sinne von "anschaulich" zu verstehen ist. Dabei werden z.T. Mittelwerte gelobt, die uns heute eher exotisch vorkommen, nur weil man sie anschaulich finden mag. Um nur ein Beispiel zu nennen: Es heißt, der Scheidewert (R bei Flaskämper) habe "eine anschauliche Bedeutung" (S. 395) vor allem in der Einkommensstatistik; denn "... hier hat die Bestimmung jenes Einkommenswertes einen Sinn, der die Reihe der Einkommensempfänger so in zwei Abschnitte zerlegt, daß alle unterhalb dieses Wertes liegenden Einkommensempfänger in ihrer Gesamtheit genau soviel beziehen wie die oberhalb desselben gelegenen" (S. 395). Das ist nicht mehr als eine Beschreibung der Formel. Warum es aber einen Sinn haben soll, das Einkommen $x = R$ zu kennen, bei dem gilt $\sum_{x \leq R} xh(x) = \sum_{x > R} xh(x)$, so wie beim Median $x = Z$ gilt $\sum_{x \leq Z} h(x) = \sum_{x > Z} h(x) = 0,5$ wird nicht gesagt.⁶⁵ Dabei wäre das eigentlich eine naheliegende Frage, wenn man den Anspruch stellt, eine der Sozialwissenschaft besser dienende Art von Statistik zu liefern

Wenn die "sachlogische Bedeutung" vor allem (oder gar nur) dann gegeben ist, wenn eine Formel in Worten anschaulich gemacht werden kann überrascht es auch nicht, dass Flaskämper offenbar Schwierigkeiten hatte, den "Sinn" des harmonischen Mittels, etwa im Unterschied zu Median und Modus zu erkennen.⁶⁶

⁶³ Ihm war insbesondere auch bewusst, dass demgegenüber die angelsächsische Statistik gerade die mathematische Richtung pflegte, die sie jedoch, wie er meinte überschätzte und "übertrieben" anzuwenden geneigt war. Ihm lag durchaus auch daran, "zu einer fruchtbaren Vereinigung beider Strömungen zu gelangen" (sehr deutlich in Flaskämper 1940, S. 66*), allerdings wohl nur unter den oben genannten Bedingungen.

⁶⁴ Die im folgenden zuerst genannte Aufgabe betrifft das im nächsten Abschnitt behandelte Adäquationsproblem

⁶⁵ Was der Median (Zentralwert) bezüglich der relativen Häufigkeiten h ist (kumuliert bis zur 50% Linie), das ist R bezüglich der kumulierten Produkte xh . Ein Vergleich der Dichtefunktionen h und xh findet statt bei der Lorenzkurve. Das Größenverhältnis von Z und R ist also Ausdruck der Disparität (relativen Konzentration).

⁶⁶ Nach Flaskämper 1929a, S. 155 *fehlt* beim harmonischen Mittel ein "logisch-sachlicher Sinn"; an anderer Stelle schreibt er zumindest, dass er sehr zweifelhaft sei (1931, S. 392); vgl. Fußnote 44.

Noch 1962 unternahm Blind, ganz in der Tradition der FSS, den Versuch, "einer den wirklichen Bedürfnissen gerecht werdenden sachbezogenen Methodenlehre der sozialwissenschaftlichen Statistik" (Blind 1962, S. 232) am Beispiel der Mittelwerte zu demonstrieren. Offenbar wurden auch dort diese wirklichen Bedürfnisse (und wohl auch die Sinnhaftigkeit) primär im Sinne von Anschaulichkeit verstanden. Er konstruierte dazu einen Dualismus in der Zielsetzung, die nämlich "entweder auf eine materiell bedeutsame Aussage über ein sozialwissenschaftliches Phänomen oder auf die Erschließung eines Verteilungsgesetzes gerichtet sein" kann. Für die erste Zielsetzung seien "einfache, anschauliche und sachlich interpretierbare Maßzahlen" gefragt und "abstrakte Maßzahlen" müssten "mangels entsprechender Deutungsmöglichkeiten" ausscheiden. Zu den abstrakten Maßzahlen zählt er das arithmetische Mittel \bar{x} , das für ihn nur Wert hat als Schätzwert für das μ der Grundgesamtheit (also für die genannte "Erschließung eines Verteilungsgesetzes"), aber auch das geometrische Mittel und die Varianz s^2 . Es wird demgegenüber die mittlere absolute Abweichung gelobt und nicht gesehen, dass nicht nur diese, sondern auch die Standardabweichung s ein (quadratisches) Mittel von Abweichungen ist. "Formale Beziehungen" zwischen Maßzahlen und mathematische Eigenschaften von Maßzahlen werden, auch wenn sie nur bescheidene Ansprüche an das Abstraktionsvermögen stellen, als uninteressant hingestellt, wenn nicht gar lächerlich gemacht.⁶⁷

Was die der "Sachlogik" zu verdankenden Einsichten betrifft, so ist mir hier nur die These, bekannt, dass Median und Modus besser (und z.B. für die Beurteilung der Verteilung von Löhnen, Einkommen usw. wichtiger) seien als die erwähnten abstrakten Mittelwerte.⁶⁸ Das ist keine eindrucksvolle Demonstration der Leistungsfähigkeit der sachlogischen Interpretation von Mathematik: man kommt über umständliche Verbalisierungen von Formeln nicht hinaus, wertet letztlich allein aufgrund eines nicht immer nachvollziehbaren Urteils über die "Anschaulichkeit" und konstruiert Dualismen (neben den bereits im Teil 1 genannten), wie den zwischen "materiell bedeutsame" Aussagen und "Erschließung eines Verteilungsgesetzes", die höchst zweifelhaft sind.

b) Das "Adäquationsproblem"

Anders als das Thema "Mittelwerte" ist die "Adäquation" ein Problem, das hervorgehoben zu haben auch heute noch von vielen als besondere Leistung der (deshalb positiv zu würdigen) FSS gelobt wird. Man hat Adäquation auch oft als *das* zentrale Thema einer möglichen "Theorie der Wirtschaftsstatistik" gesehen.⁶⁹ Ein Interesse daran, die Bedeutung und Problematik der Adäquation herauszustellen besteht auch auf Seiten der amtlichen Statistik, die sich inzwischen vermehrt gedrängt sieht, den wissenschaftlichen Charakter ihrer Leistungen zu

⁶⁷ "Was nützt es ihm (gemeint ist: dem Wirtschaftler und Sozialpolitiker, v. d. L.) da z.B. zu wissen, daß bei einer Lohnverteilung die Summe der Quadrate der Abweichungen aller Individuallöhne vom Durchschnittslohn ein Minimum ist?" (Blind 1962, S. 218).

⁶⁸ Diese These findet sich bei Zizek, Flakämper und auch in Blind 1952, S. 354. Die Anschaulichkeit verneint Blind 1962 bei einem "abstrakten" Mittelwert, wie dem arithmetischen Mittel, "weil bei seiner Bestimmung die einzelnen empirischen Werte sozusagen miteinander verrechnet werden und im Ergebnis untergehen" (S. 233). An diesem Maßstab gemessen müsste das meiste, was die Statistik an Ergebnissen liefert, unbrauchbar sein.

⁶⁹ Ich muss zugeben, dass ich auch selbst lange solchen Gedanken gegenüber aufgeschlossen war (vgl. v. d. Lippe 1996, S. 27). Es gab eine Zeitlang die Idee, es könne neben der Stochastik als Fundament der statistischen Methoden auch eine solche spezielle "Theorie" für die Wirtschaftsstatistik geben (Menges, auch ein Flakämper Schüler, prägte dafür den Begriff "Deigmatik" und es wurde auch der Begriff "Syllepsis" diskutiert). Hinter solchen, inzwischen sicher überholten Ideen steht natürlich die in der FSS hochgehaltene These von den zwei Arten von Statistik, die ich – wie die obigen Ausführungen zeigen – inzwischen ablehne. Diese These wurde wohl auch nicht nur in der FSS in Deutschland lange hochgehalten. Nach Rinne 2011, S. 81 schien man in der Deutschen Statistischen Gesellschaft nach dem Zweiten Weltkrieg noch lange dem Gedanken der zwei Arten von Statistik anzuhängen: "Die Repräsentanten der Gesellschaft übernahmen auch nicht die sich international durchsetzende, aus dem angelsächsischen Bereich stammende Auffassung von einer einheitlichen Methodenlehre für die sozial- und naturwissenschaftliche Statistik."

betonen.⁷⁰ Es gibt viel Literatur zu diesem Thema⁷¹ und auch Aufsätze, in denen beispielhaft gezeigt wird, wie bestimmte Konzepte auf diese Weise messbar gemacht werden.⁷² Es ist auch unbestritten, dass sich in der Arbeit der Statistiker und auch in der Diskussion unter den Nutzern der Statistik viel darum dreht, ob mit einem Erhebungsbegriff oder Messkonzept wirklich das erfasst wird, was gezählt, bzw. gemessen werden sollte. Man denke nur an die ständigen Diskussionen darüber, dass die "wahre" Zahl der Arbeitslosen viel höher (oder auch niedriger) sei als die in der Statistik ausgewiesene Zahl.

Dass der FSS die Betonung des Adäquationsproblem als besondere Leistung zugutegehalten wird ist gleichwohl etwas zu relativieren, denn

- inhaltlich wird viel von dem, was hier der Gegenstand ist, auch unter anderen Begriffen reflektiert ohne dass damit die fragwürdige erkenntnistheoretische Position der FSS verbunden wird;
- das Adäquationsproblem, so wie es von der FSS dargestellt wird betrifft nur einen Teil der mit statistischen Daten verbundenen Konkretisierungs- und Messprobleme, nämlich solche, die durch eine (operationale) Definition eines Begriffs zu lösen sind, d.h. indem man angibt, was darunter subsumiert werden soll (nach diesem Muster kann aber nicht bei allen Adäquationsproblemen verfahren werden).⁷³

Der bei der Adäquation implizierte Gedanke einer Annäherung eines Beobachtungsbegriffs an ein ideelles "Konstrukt" (was eigentlich intendiert ist) wird auch unter dem Begriff "Validität" (z.B. einer Messung) diskutiert und es gibt auch Methoden, die "Validität" zu quantifizieren.⁷⁴ Als Qualitätskriterium in der amtlichen Statistik kennen wir auch die "Relevanz", wonach eine Statistik einem gewünschten Zweck und Ziel gerecht werden soll. Es ist auch davon auszugehen, dass auch solche Produzenten und Konsumenten von Statistiken, die nie vom Adäquationsproblem gehört haben,⁷⁵ genau die hiermit gemeinten Überlegungen anstellen. Man kann sich entsprechende Gedanken machen, ob eine statistische Zahl wirklich das zum Ausdruck bringt, was mit dem entsprechenden Begriff gemeint ist, auch ohne deshalb einen gewissen Agnostizismus gutzuheißen, wie: "Also an das eigentliche Wesen sozialer Tatbestände kommen wir mit Zahl und Maß nicht heran, sie sind im Kern nicht quantifizierbarer Natur".⁷⁶

Dass man trotzdem mit Zahlen operieren kann, also auch Statistik eine Existenzberechtigung hat, bedeutet "daß sie erst zähl- und meßbar gemacht werden müssen",⁷⁷ was die Aufgabe der Adäquation ist. Diese Aufgabe wird allerdings genau dann kaum lösbar, wenn die Zielgröße "das eigentliche Wesen" oder der "Kern" sein soll. Mit Recht beklagt deshalb auch Litz 1990, S. 429 die "mangelnde Stringenz und Operationalität des Adäquationsbegriffs"⁷⁸ und das Feh-

⁷⁰ Adäquation wird in diesem Zusammenhang ausdrücklich erwähnt in Egeler, Wöll, Zwick 2012, S. 277.

⁷¹ Grohmann 1985, Litz 1990, v. d. Lippe 1996, S. 28ff.

⁷² Erwähnt seien Aufsätze von Neubauer und Härtel im Band 60 (1976) S. 285 – 482 des Allgemeinen Statistischen Archivs, der als Gedenkband zum 90. Geburtstag von Flaskämper und zum 70. Geburtstag von Blind erklärt wurde.

⁷³ Tatsächlich lassen sich sogar die wenigsten Dinge durch Erarbeitung einer ultimativen Definition messen, und oft ist auch eine begriffliche Adäquation gar nicht nötig um etwas beobachten und messen zu können.

⁷⁴ Die hier im Verständnis der FSS implizierte Annäherung, die wohl schrittweise und allein durch Nachdenken erzielt wird, würde es eigentlich nahelegen, an ein Maß des Abstands zwischen dem operationalen Begriff und dem Ziel-Begriff zu denken, das dann für die "Güte" einer Adäquation bestimmend sein könnte. Es ist mir nicht bekannt, dass sich jemand im Umkreis der FSS mit solchen Überlegungen beschäftigt hatte.

⁷⁵ Dies dürfte wohl für mehr oder weniger alle Ausländer und bald wohl auch für alle Deutsche gelten.

⁷⁶ Flaskämper 1933, S. 60.

⁷⁷ Flaskämper 1933, S. 61 und S. 65. Fast wörtlich genauso bei Blind 1952, S. 531. Beklagt wird bei Flaskämper auch das "wesensmäßige Auseinanderfallen von statistischen Begriffen und Wirklichkeitsbegriffen" (S. 65).

⁷⁸ Nach Litz 1990 ist wohl viel von den Überzeugungen der FSS Max Weber zu verdanken. Weber unterschied danach "Sinnadäquanz und Kausaladäquanz" und erstere wird wie folgt definiert: "Sinnhaft adäquat ist was durchschnittlichen Denk- und Gefühlsgewohnheiten als typisch und richtig erscheint" (Weber nach Litz 1990, S.

len einer Adäquationstheorie. Es gibt sicher von Fall zu Fall gute Betrachtungen von Statistiker auf der Ebene von Begriffen, aber es ist fraglich, ob sie verallgemeinerungsfähige Erkenntnisse liefern, die bei der Lösung anderer Adäquationsprobleme helfen. Es scheint hier mehr nur ein situativ bestimmtes (und beispielhaft demonstriertes) Vorgehen zu geben und (noch) nicht eine Kunst- oder gar Methodenlehre zu existieren.

Das Muster eines Adäquationsprozesses, das die Autoren der FSS offenbar vor Augen haben ist, dass zunächst "Begriffe der Gegenstandswissenschaft (Nationalökonomie, Betriebswirtschaftslehre, Soziologie) 'zurecht geschliffen' werden, damit der Statistiker mit ihnen arbeiten kann" (Flaskämper 1950, S. 157). Aber so vorzugehen, d.h. zuerst auf beobachtbare Sachverhalte beziehend definieren und dann zählen, bzw. messen funktioniert nicht überall. So kann man verfahren bei einer Systematik (z.B. Definition von Waren, Berufen etc.), oder bei der Definition von Erhebungseinheiten (Haushalt, Familie, Arbeitslose etc.), oder auch bei Größen, die durch Aufzählung der einzubeziehenden Positionen definiert werden können, wie z.B. Vermögen, Geldvolumen etc. Aber man kann z.B. nicht den Begriff der Intelligenz so "zurecht schleifen", dass man danach feststellen kann, wie intelligent jemand ist. Das gleiche gilt bei der Wohlstandsmessung. Hier ist ein anderes Vorgehen gefragt und auch üblich.⁷⁹

Man sieht daran, dass das Adäquationsproblem kein Alleinstellungsmerkmal der FSS ist, und dass es von dieser nur eingeschränkt und mit wenig greifbaren Ergebnissen behandelt worden ist. Ein Problembewusstsein in dieser Hinsicht ist sehr zu wünschen, aber es kann auch gefördert werden, ohne dass dabei das Gedankengut der FSS mitgeliefert werden muss.

c) Der deutsche Sonderweg

Es wird schon seit einiger Zeit gesagt, dass die FSS obsolet ist,⁸⁰ was wohl auch ganz besonders für die dahinterstehende idealistische Philosophie gelten dürfte.⁸¹ So etwas fasziniert kaum noch. Allerdings scheint durchaus ein Bedürfnis danach zu bestehen, Methoden nicht nur "formal", sondern vermehrt auch verbal zu reflektieren, sie "verstehen" zu wollen und nach ihrem "Sinn" zu fragen. Das ändert aber nichts daran, dass die FSS Vergangenheit ist. Was rechtfertigt es dann, sich mit ihr zu beschäftigen. Zwei Gründe sprechen dafür, dass es sich auch heute noch lohnt, sich mit dieser obsoleten Schule zu beschäftigen

- die *Tatsache*, dass Statistiker hierzulande für eine längere Zeit (wohl für Jahrzehnte) den Anschluss an das Ausland verloren hatten und
- die durch Erfahrung mit der FSS untermauerte *Erwartung*, dass es auch in Zukunft schwierig sein dürfte der dominierenden "rein formalen" Betrachtungsweise und Fokussierung auf entsprechende Themen in der Statistik etwas entgegenzusetzen.

444). Kausaladäquanz ist aus dem Strafrecht gut bekannt und zu verstehen als ein Ablauf, der der Erfahrung entspricht. Es ist kausaladäquat dass man stirbt, wenn man ein Beil auf den Kopf bekommt, wenn man nur über ein Beil stolpert und dann stirbt, wäre das nicht kausaladäquat.

⁷⁹ In v. d. Lippe 1996, S. 31 habe ich Messen mit einem *Modell* (z.B. Faktorenanalyse bei Intelligenz oder das Modell der stationären Bevölkerung bei der Lebenserwartung) und Messen mit *Indikatoren* unterschieden. Indikatoren (mit oder ohne addierte Punktzahlen versehen) sind sicher eine unbefriedigende Lösung. Man verfährt so bei der Wohlstandsmessung oder z.B. auch, wenn die wissenschaftliche Qualität eines "papers" daran gemessen wird, wie oft es zitiert wird. Das mag angreifbar sein, aber der Versuch, die Variable "wissenschaftliche Qualität" begrifflich "zurecht zu schleifen" wird vermutlich auch nichts Besseres liefern.

⁸⁰ Günter Menges schrieb 1976 in dem erwähnten Band 60 des Allgemeinen Statistischen Archivs, diese Schule "schien eine Zeitlang obsolet zu sein" (S. 293). Dass sie es aber spätestens jetzt, 36 Jahre später tatsächlich sein dürfte (und nicht nur zu sein scheint), wird wohl allgemein so gesehen.

⁸¹ Wie erwähnt sind Überzeugungen der FSS, wie die Existenz prinzipiell nicht erkennbarer geistiger Entitäten, wie z.B. das "Wesen" und die Vorstellung, dass alles aus *einem "letzten"* Grund heraus deduktiv ableitbar ist, Erbe der Philosophie des "Deutschen Idealismus" (Fichte, Hegel, Schelling).

Beginnen wir mit den Tatsachen, auch wenn sie "nur" Historie sind. Wir haben bewusst sehr viel aus Flaskämper 1950 zitiert. Bemerkenswert an diesem Text ist nämlich, dass er fast wörtlich das alles wiederholte, was bereits in den Schriften von 1931 und 1933 stand. Auch die Aussagen in Blind 1952 und 1953 waren praktisch alle bereits in Blind 1936 zu lesen. Man findet in Flaskämper 1950, also nach dem Krieg lediglich etwas stärker als früher ein Bedauern, darüber dass man die Mathematik und ausländische Schriften vernachlässigt hat: "Es kann nicht bestritten werden, daß die Sozialstatistik in Deutschland schon seit langem die mathematische Weiterentwicklung ihrer Methoden stark vernachlässigt hat und dadurch gegenüber dem Ausland ... in bedauerlicher Weise ins Hintertreffen geraten ist" (S. 152).⁸² Darüber, dass seine vor dem Krieg publizierten Gedanken maßgeblich dazu beitrugen, sich guten Gewissens ins Hintertreffen zu begeben, verlor er jedoch kein Wort. Dabei war der entstandene Schaden nicht unerheblich. "Die wissenschaftliche Statistik musste sich aus der jahrelangen Isolierung lösen und den Anschluss an die Entwicklung im Ausland suchen. Vergangenheitsbewältigung hat sie allerdings nicht oder so gut wie nicht betrieben, obwohl es dazu mehr als einmal Gelegenheit gegeben hat"⁸³ (Rinne 2011, S. 80). Nach Rinne war auch die Einstellung der meisten Statistiker in Deutschland zu den mathematischen Methoden noch "über viele Jahrzehnte bis in die 1970er Jahre kritisch" (S. 81).

Dass Deutschland ins "Hintertreffen geraten ist" kann auch wieder recht eindrucksvoll am Beispiel der Indextheorie gesehen werden. Nach hoffnungsvollen Anfängen mit Drobisch, Laspeyres, Paasche und Lehr, Ende des 19. Jahrhunderts⁸⁴, die auch alle im Ausland sehr beachtet wurden kamen, bis v. Bortkiewicz 1923/24, Haberler 1927 und (unbeachtet) Flaskämper 1928 kaum nennenswerte Beiträge zu diesem Gebiet aus Deutschland.⁸⁵ Ganze Entwicklungen, wie z.B. die (ältere) stochastische und die sich allmählich heranbildende ökonomische Theorie der Indexzahlen⁸⁶ sind weitgehend unbemerkt an uns vorbeigegangen. In dieser Zeit wurden in den USA jeweils zwei Bücher über Indexzahlen von Correa Moylan Walsh (1901 und 1921) und von Irving Fisher (1911 und 1922) publiziert, ferner über ein Dutzend Aufsätze von F. Y. Edgeworth sowie mehrere Arbeiten von Bowley, Knibbs und anderen englischsprachigen Autoren⁸⁷. Während v. Bortkiewicz, aber auch Haberler solche Arbeiten zitierten, gilt dies kaum für Flaskämper, der z.B. zwar die auch heute noch vielbeachtete Arbeit "Zweck und Struktur einer Preisindexzahl" von Ladislaus v. Bortkiewicz aus den Jahren 1923 und 1924 beiläufig zitierte, aber ohne auf Inhalte einzugehen und nur um zu sagen, dass die vielen Formeln darin ein Beweis "für den falschen Weg" dieser Art Indextheorie sei.⁸⁸

Das Ignorieren fremder und anderer (mehr mathematischer) Arbeiten erfolgte also absichtlich und aus Prinzip. Das "Hintertreffen" war nicht ein bedauerliches Missgeschick oder Folge einer unglücklichen politischen Situation in Deutschland, sondern es entstand aus der Illusion, auf einem besseren Weg zu sein. Das erklärt es auch, warum sich nach dem Zweiten Weltkrieg die Einsicht, auf den falschen Weg gewesen zu sein, nicht so schnell verbreitete: "Die Integration moderner, stärker formal-mathematische ausgerichteter Ansätze, wie sie insbesondere in den angelsächsischen Ländern schon weiter entwickelt worden waren, gelang al-

⁸² Er lobte, dass "... in der deutschen Statistik im Gegensatz zur ausländischen die Herausarbeitung der sachlogischen Begriffe gepflegt worden ist", räumte aber ein, dass dies "allerdings unter zeitweise sehr starker Vernachlässigung der mathematischen Seite ..." geschah (Flaskämper 1950, S. 159).

⁸³ Rinne spielte damit auf die NS-Vergangenheit des langjährigen Vorsitzenden der Deutschen Statistischen Gesellschaft, Friedrich Zahn an. Inzwischen ist auch die NS-Vergangenheit von Flaskämper gut dokumentiert.

⁸⁴ Die Zeitspanne der Arbeiten der genannten Autoren zum Thema Preisindizes ist genaugenommen sehr klein, nämlich zwischen 1871 (die erste Arbeit von Drobisch) und 1885 (das Buch von Lehr).

⁸⁵ Vgl. hierzu und zu den folgenden Ausführungen v. d. Lippe 2013.

⁸⁶ Es ist v. Bortkiewicz zu verdanken, dass die Arbeit des russischen Ökonom Konüs weltweit bekannt wurde.

⁸⁷ Es ist diesen Autoren weit mehr als deutschen Autoren zu verdanken, dass Namen wie Laspeyres und Paasche nicht in Vergessenheit geraten sind und in Verbindung mit Indexformeln noch heute bekannt sind.

⁸⁸ Flaskämper 1929a, S. 158, vgl. auch Fußnote 27. Ganz ähnlich wurde auch – wie gesagt – Fisher zwar zitiert, aber hauptsächlich um ihn als Paradebeispiel eines abzulehnenden Ansatzes hinzustellen.

lerdings zunächst nur zögernd. Lediglich moderne Stichprobenverfahren fanden bald vielfach Beachtung und Anwendung" (Rinne 2011, S. 81).

Noch auf der Jahrestagung der Deutschen Statistischen Gesellschaft 1953 in Heidelberg wurde zwischen der Frankfurter Schule (vertreten durch Flaskämper, Blind, Hartwig und Statistiker aus den Ämtern⁸⁹) und der "Münchener Schule" (Anderson, Kellerer, Strecker) heftig gerungen. Das ist sehr ausführlich in dem heute noch lesenswerten Bericht Boustedt 1953 dokumentiert. Dass es dabei auch um das beklagte "Hintertreffen" ging macht eine Bemerkung von Hans Kellerer deutlich: "Den Versuch, das Gesamtgebäude der statistischen Methodik in zwei Teile aufzuspalten, halte ich ... für ein unglückliches Unterfangen ... Wollen wir wirklich den Anschluß an die internationale Forschung wiedergewinnen, dann müssen sich auch die Sozialstatistiken gründlich mit den Fragen der sog. mathematischen Statistik vertraut machen." (nach Boustedt 1953, S. 348).

Falls sie dieses in der Zwischenzeit zu gründlich getan haben sollten, kann es sein, dass wir es am Ende vielleicht mit einem Dilemma zu tun haben. Denn es ist auch nicht unbedenklich, wie inzwischen Statistik an den Hochschulen gelehrt wird und dass man heutzutage eher mit wirren Darstellungen, die sehr mathematisch sind oder zumindest so erscheinen punkten kann als mit verbalen (oder gar "sachlogischen") Ausführungen.⁹⁰ Es ist interessant zu sehen, dass "abgehobene" Mathematik in der Statistik auch schon vor über sechzig Jahren von einem berühmten Statistiker beklagt wurde.⁹¹ Aber die nicht nachahmenswerten deutschen Erfahrungen mit der FSS (quasi Nicht-Mathematik als Programm) lehren uns, dass sich daran wohl auch in Zukunft so schnell nichts ändern wird.

Literatur

Abkürzung: AStA = Allgemeines Statistisches Archiv

Below, Fritz (1956), Über den Erkenntniswert sozialstatistischer Daten, Schweizerische Zeitschrift für Volkswirtschaft und Statistik, Bd. 92, S. 78 – 86.

Blind, Adolf (1936), Der Ganzheitscharakter der Volkswirtschaft und die Statistik, in Flaskämper und Blind 1936, S. 15 – 34.

Blind, A. (1952), Die neue Entwicklungsrichtung der Sozialwissenschaftlichen Statistik, Zeitschrift für die gesamte Staatswissenschaft, Bd. 108, S. 528 – 537.

Blind, A. (1953), Probleme und Eigentümlichkeiten sozialstatistischer Erkenntnis, AStA, Bd. 37, S. 301 – 313.

Blind, A. (1962), Die zusammenfassende Kennzeichnung von Häufigkeitsverteilungen auf Grund quantitativer Merkmale in der sozialwissenschaftlichen Statistik, AStA Bd. 46, S. 217 – 235.

Boustedt, Olaf (1953), Die 24. Jahresversammlung der Deutschen Statistischen Gesellschaft am 28. und 29. Oktober 1953, AStA Bd. 37, S. 314 – 361.

Egeler, R., T. Wöll und M. Zwick (2012), Perspektiven für die amtliche Statistik, Wirtschafts- und sozialstatistisches Archiv, Bd. 5, H. 4, S. 269 – 284.

Fisher, I. (1922/1927), The Making of Index Numbers, Boston, New York (1st edition 1922, 2nd 1927).

⁸⁹ Von dieser Seite wurde die FSS gerade deshalb sehr gelobt, weil sie traditionell und aus verständlichen Gründen den größten Wert auf anschauliche und leicht verbal zu vermittelnde Methoden legt, die damit auch mehr den Nutzern und Geldgebern der amtlichen Statistik zusagen.

⁹⁰ Ein Beispiel hierfür sind die oben angedeuteten kontroversen Stellungnahmen zu Kettenindizes von Werner Neubauer und Marco Martini vor Eurostat 1995.

⁹¹ Selbst ein so prominenter Vertreter der mathematischen Statistik wie M. G. Kendall klagte (Kendall 1950, S. 135): "...I feel sometimes, in wrestling with papers on statistical matters with a full emblazonment of modern mathematics, sets of points, Lebesgue integration, asymptotic summability and so forth, that work which achieves a slight gain in generality at the expense of being incomprehensible is defeating its own purpose... One of the few doctrinal points on which all statisticians seem to agree is the necessity for keeping one foot firmly on the ground."

- Flaskämper, Paul (1927a), Die Statistik und das Gesetz der großen Zahlen, AStA Bd. 16, S. 501 – 514.
- Flaskämper, P. (1927b), Bemerkungen zum Indexproblem, Eine Erwiderung, AStA Bd. 16, S. 645 – 648.
- Flaskämper, P. (1928), Theorie der Indexzahlen, Beitrag zur Logik des sozialen Vergleichs, Berlin 1928.
- Flaskämper, P. (1929a), Der Sinn der Indexzahlen, Betrachtungen zu Gottfried Haberlers Buch gleichen Titels, AStA Bd. 18, S. 149 – 159.
- Flaskämper, P. (1929b), Das Problem der Gleichartigkeit in der Statistik, AStA Bd. 19, S. 205 – 234.
- Flaskämper, P. (1931), Beitrag zur Logik der statistischen Mittelwerte, AStA Bd. 21, S. 379 – 404.
- Flaskämper, P. (1933), Die Bedeutung der Zahl für die Sozialwissenschaften, AStA Band 23 (1933/34), S. 58 – 71.
- Flaskämper, P. u. A. Blind (1936), Beiträge zur deutschen Statistik, Festgabe für Franz Zizek zur 60. Wiederkehr seines Geburtstages, Leipzig 1936.
- Flaskämper, P.(1936), Gegenwarts- und Zukunftsaufgaben der Statistik in Deutschland, in Flaskämper u. Blind (1936), S. 1- 14.
- Flaskämper, P. (1940), Zur gegenwärtigen Lage der Statistik, Weltwirtschaftliches Archiv, Bd. 52, S. 65* – 72*.
- Flaskämper, P. (1950), Mathematik und Statistik, AStA Bd. 34, S. 105 – 161.
- Grohmann, Heinz 1985, Vom theoretischen Konstrukt zum statistischen Begriff, Das Adäquationsproblem, AStA, Bd. 69 (1985), S. 1 - 15.
- Hermberg, Paul (1923), Die richtige Form der Indexziffer, Weltwirtschaftliches Archiv Bd. 19, S. 585 – 594.
- Jacobs, Alfred (1940), Neue Methoden in der deutschen Statistik, AStA Bd. 29, S. 392 – 403.
- Kendall, Maurice G. (1950), The Statistical Approach, *Economica* Vol. 17, No. 66, pp. 127 – 145.
- Klezl-Norberg, Felix (1943/44), Die Theorien der Gleichartigkeit in der Statistik, AStA Bd. 32, S. 136 – 155.
- Klein, Ingo (2004), Grundlagenstreit in der Statistik, Diskussionspapier 60/2004 der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Friedrich Alexander Universität Erlangen-Nürnberg.
- Laspeyres, Etienne (1875), Die Kathedersocialisten und die statistischen Congresse, Gedanken zur Begründung einer nationalökonomischen Statistik und einer statistischen Nationalökonomie, Berlin 1875.
- Litz, Hans Peter (1990), Statistische Adäquation und Idealtypus, AStA Band 74, S. 429 – 456.
- Martini, Marco (1995), The frequency of updating of the base of consumer price index numbers (expertise for Eurostat, unpublished) 1995.
- Meeker, Royal (1921), On the Best Form of Index Number, *Journal of the American Statistical Association*, Vol. 17, No. 135, pp. 909 – 915
- Mudgett, Bruce D. (1951), *Index Numbers*, New York (John Wiley).
- Müller, Johannes (1923/24), Begriffsstatistik, AStA Band 14, S. 438 - 455
- Neubauer, Werner (1995), Konzeptionelle Vor- und Nachteile eines verketteten Verbraucherpreisindex, hektographiert, 29. Juni 1995.
- Rinne, Horst (2011), Das Allgemeine Statistische Archiv, in H. Grohmann, W. Krämer u. A. Steger (Hrsg.), *Statistik in Deutschland, 100 Jahre Deutsche Statistische Gesellschaft*, Berlin, Heidelberg, S. 77 - 86.
- von Bortkiewicz, Ladislaus 1923, Zweck und Struktur einer Preisindexzahl, *Nordisk Statistisk Tidskrift*, Vol. II, pp. 369 – 408 (Teil I), Teil II *Nordisk Statistisk Tidskrift*, Vol III, 1924, 208 – 252, Teil III *Nordisk Statistisk Tidskrift*, Vol III, 1924, 494 – 516
- von Bortkiewicz, Ladislaus (1924), Buchbesprechung von Irving Fisher, *The making of index numbers*, *Archiv für Sozialwissenschaft und Sozialpolitik*, Bd. 51, S. 848 – 853.
- von Bortkiewicz, L. (1927), Der gegenwärtige Stand des Problems der Geldwertmessung, in *Handwörterbuch der Staatswissenschaften*, 4. Aufl. 1927 Bd. IV S. 743 -752.
- von Bortkiewicz, L. (1932), Die Kaufkraft des Geldes und ihre Messung, *Nordic Statistical Journal* 1932, pp. 1 – 68.
- von der Lippe, Peter (1996), *Wirtschaftsstatistik*, 5. Auflage, Stuttgart.

- von der Lippe, P. (2001), Chain Indices, A Study in Index Theory, Frankfurt, Publication Series Spectrum of Federal Statistics, Vol 16, Wiesbaden, Statistisches Bundesamt (Hrsg.).
- von der Lippe, P. (2005), Das Ideal des "reinen Preisvergleichs", Jahrbücher für Nationalökonomie und Statistik Bd. 225, S. 499 -509.
- von der Lippe, P. (2007), Index Theory and Price Statistics, Frankfurt, 2007
- von der Lippe, P. (2010), Neuere Entwicklungen in der Preisstatistik und Indextheorie, Wie mit der traditionellen deutschen Theorie und Praxis gebrochen wurde, Jahrbuch für Wirtschaftswissenschaften, Band 61/2010, H. 2, S. 171 – 195.
- von der Lippe, P. (2013), Recurrent Price Index Problems and Some Early German Papers on Index Numbers; Notes on Laspeyres, Paasche, Drobisch and Lehr, angenommen für Jahrbücher für Nationalökonomie und Statistik.
- Winkler, Wilhelm (1923), Die "beste Indexformel" Bemerkungen zu Irving Fishers "The Making of Index Numbers", Jahrbücher für Nationalökonomie und Statistik Bd. 121, S. 571 -581.
- Winkler, W. (1930), Buchbesprechung von Flaskämper Paul, Theorie der Indexzahlen, Berlin u. Leipzig 1928, Zeitschrift für die gesamte Staatswissenschaft, Bd. 88, 652 – 655.
- Zizek, Franz (1931), Der statistische Vergleich, AStA Bd. 21, S. 525- 550.