

Regionale Unterschiede der Kriminalitätsbelastung in Westdeutschland

Zur Kontroverse um eine Nord-Süd-Gefälle der Kriminalität

von

Peter Wetzels und Christian Pfeiffer

Januar 1996

Regionale Unterschiede der Kriminalitätsbelastung in Westdeutschland - Zur Kontroverse um ein Nord-Süd-Gefälle der Kriminalität -

von Peter Wetzels und Christian Pfeiffer

Zusammenfassung: Seit langem weist die Polizeiliche Kriminalstatistik ein starkes Nord-Süd Gefälle der registrierten Kriminalität auf. Es war lange Zeit unklar, inwieweit dies reale Divergenzen der Kriminalitätsbelastung widerspiegelt oder aber Resultat von Unterschieden in Anzeigeverhalten und polizeilicher Registrierungspraxis ist. In einer aktuellen Publikation haben Kury et al. ausgeführt, die Ergebnisse von zwei repräsentativen Opferbefragungen des MPI hätten zeigen können, daß dieses Gefälle auf tatsächlichen regionalen Unterschieden der Viktimisierung beruhe. Diese Daten von Kury et al. werden einer methodischen Kritik sowie einer Reanalyse unterzogen. Im Ergebnis zeigt sich, daß aufgrund methodischer Unzulänglichkeiten und widersprüchlicher Befunde die von ihnen aufgestellte Hypothese empirisch auf der Grundlage ihrer eigenen Daten weder gestützt noch verworfen werden kann. Ferner werden Ergebnisse der KFN-Opferbefragung 1992 zu dieser Frage vorgestellt. Sie zeigen, auf der Basis einer repräsentativen Befragung von 8966 deutschsprachigen Bewohnern von Privathaushalten in Westdeutschland, daß eine Höherbelastung des Nordens sowie des Bundeslandes Schleswig-Holstein im Bereich von - vor allem leichteren - Eigentumsdelikten festzustellen ist, nicht aber für Gewaltdelikte. Ferner werden Hinweise zur Bedeutung des Anzeigeverhaltens für eine partielle Erklärung von Nord-Süd-Divergenzen vorgelegt. Das feststellbare höhere Viktimisierungsrisiko der Region Nord wird jedoch überschätzt, wenn die PKS als Indikator des Viktimisierungsrisikos von Privatpersonen fehlgedeutet wird. Es werden die methodischen Erfordernisse sowie Vor- und Nachteile kriminalstatistischer Analysen auf der Basis von Opferbefragungsdaten und die Möglichkeiten und Grenzen ihres Vergleichs mit der polizeilichen Kriminalstatistik diskutiert.

1. Einleitung

Zum Jahreswechsel 95/96 würden die Bürger Schleswig-Holsteins durch dramatisch klingende Pressemeldungen aufgeschreckt. Überregional und regional waren Schlagzeilen wie „Kriminalität: Schleswig-Holsteiner leben am gefährlichsten“ (WamS vom 01.01.96) oder „Schleswig-Holstein ist am gefährlichsten“ (Lübecker Nachrichten vom 31.12.1995) in der Presse zu lesen. Die Meldungen beriefen sich auf eine Studie des Max-Planck-Institutes Freiburg (Kury, Obergfell-Fuchs & Würger, 1995). Die Autoren hatten darin für Norddeutschland, insbesondere Schleswig-Holstein, im Vergleich zu Bayern und Baden-Württemberg eine wesentlich höhere Kriminalitätsbelastung festgestellt. Gestützt auf die Daten zweier repräsentativer Opferbefragungen berichteten sie, das Viktimisierungsrisiko sowohl für Kontaktdelikte als auch für Kfz-Delikte sei in Schleswig-Holstein „etwa doppelt so hoch“ (Kury et al., 1995 S. 771) wie in den beiden südlichen Bundesländern.

Diese Forschungsarbeiten und ihr vehementes Presseecho hatten auch politische Folgen. Die CDU-Landtagsfraktion sah bestätigt, daß in Schleswig-Holstein ein erhebliches Defizit der inneren Sicherheit bestehe. Trotz finanzieller Engpässe müsse der Sicherheit der Bürger nun politische Priorität eingeräumt werden. Die Landesregierung könne sich an dem Erfordernis zusätzlicher Polizeistellen nicht mehr vorbeimogeln (Pressemitteilung Nr. 1/96 der CDU Landtagsfraktion vom 03.01.1996). Der Landesinnenminister bezeichnete demgegenüber die Studie wegen einer für sein Bundesland unzureichenden Datenbasis als methodisch unzulängliche „Kaffeersatzleserei“ sowie „unseriös“. Sie sei lediglich zu „Panikmache und Dramatisierung“ geeignet (Pressestelle der Landesregierung: Erklärung vom 02.01.1996). Auf diese Vorwürfe erwiderte Kury in einem Leserbrief an die FAZ (12.01.1996), die Kritik einer unzureichenden Datenbasis träfe nicht zu. Es seien bundesweit 13.000 Personen befragt worden. Damit sei es gelungen, eine für Schleswig-Holstein im Vergleich zum Süden Deutschlands beträchtlich höhere Kriminalität

tätsbelastung nachzuweisen. Während man bisher *vermutet* habe, daß die in der Polizeilichen Kriminalstatistik (PKS) schon seit längerem auftauchenden Nord-Süd-Divergenzen auf regionale Unterschiede sowohl des Anzeigeverhaltens der Bevölkerung als auch der Interventions- und Registrierungspraxis der Polizei zurückgehen könnten, habe die MPI-Studie das Gegenteil gezeigt.

Mit diesen *Vermutungen* spricht Kury Thesen an, die wir 1994 publiziert haben (vgl. Pfeiffer & Wetzels, 1994). Auf der Basis einer vom KFN Anfang 1992 durchgeführten repräsentativen Opferbefragung waren wir zu der Einschätzung gelangt, daß die PKS die Unterschiede der Kriminalitätsbelastung zwischen nördlichen und südlichen Bundesländern überzeichnet. Zwar hatten auch wir für den Norden eine höhere Opferbelastung festgestellt. Die regionalen Divergenzen fielen jedoch bei weitem nicht so groß aus, wie in der PKS. Zudem betrafen die feststellbaren Unterschiede fast ausschließlich Eigentumsdelikte.

Analysen regionaler Unterschiede der Kriminalitätsverteilung, etwa der Zusammenhang zwischen Kriminalitätsbelastung und Merkmalen städtischer und ländlicher Wohnbedingungen oder sozialen und ökonomischen Faktoren, haben in der Kriminologie eine lange Tradition. Sie stellen einen Ansatz dar, die Relevanz ökologischer Faktoren für eine Erklärung des sozialen Phänomens Kriminalität zu prüfen. Werden solche ökologischen Analysen auf politische Einheiten wie Bundesländer bezogen, so entwickeln sie allerdings auch unmittelbare praktische, kriminalpolitische Brisanz. Gerade dies macht eine sorgfältige wissenschaftliche Absicherung besonders dringlich, da fehlerhafte Befunde nicht nur theoretische, wissenschaftsinterne Konsequenzen, sondern auch unmittelbare politisch-praktische, für das Gemeinwesen und die Bürger eventuell einschränkende oder kostenintensive Folgen haben können. Dies gibt uns besondere Veranlassung, die empirische Basis dieser Kontroverse genauer zu beleuchten.

Im ersten Schritt werden wir prüfen, inwieweit die MPI-Studien tatsächlich hinreichende empirische Evidenz für ein starkes Nord-Süd-Gefälle der Kriminalitätsbelastung bieten. Danach werden wir unsere eigenen zu dieser Frage erarbeiteten Forschungsergebnisse breiter darstellen. Im Zusammenhang damit werden wir jeweils erörtern, welche methodischen Voraussetzungen unseres Erachtens erfüllt sein müssen, um zu regionalen Unterschieden der Kriminalitätsbelastung überhaupt empirisch gesicherte Aussagen treffen zu können.

2. Methodische Kritik der MPI Forschung

Kury und Kollegen (Kury, 1995; Kury, Obergfell-Fuchs & Würger, 1995) stützen ihre Thesen auf die Ergebnisse von zwei Opferbefragungen: eine 1989 in Westdeutschland und West-Berlin durchgeführte telefonische Opferbefragung, welche Teil des International Crime Survey war (van Dijk, Killias & Mayhew, 1990), im folgenden als ICS 89 bezeichnet, sowie eine 1990 in den alten und neuen Bundesländern von MPI und BKA durchgeführte Opferbefragung (Kury, Dörmann, Richter & Würger, 1992), im folgenden als MPI 90 bezeichnet.

Solche repräsentativen Studien basieren auf Stichprobenbefragungen, von deren Ergebnissen ausgehend die Viktmisierungsrate in der jeweiligen Grundgesamtheit (z. B. der Bevölkerung Schleswig-Holsteins) geschätzt wird. Die Genauigkeit und Zuverlässigkeit solcher Schätzungen hängt von dem *Erhebungsinstrument* (entscheidend für Reliabilität und Validität der Messung), der *Stichprobenqualität* (z.B. deren Repräsentativität) sowie der *Stichprobengröße* ab. Ferner ist ein *adäquates inferenzstatistisches Vorgehen* unerlässlich, um den Schluß von der Stichprobe auf die Grundgesamtheit wahrscheinlich-

keitstheoretisch abzusichern. In allen vier genannten Punkten lassen sich gegen die MPI-Studie erhebliche Einwände erheben.

Zum *Erhebungsinstrument* ist festzustellen, daß dieses nicht differenziert genug ist. So kann keine getrennte Prävalenzrate für leichte und schwere Gewalt berechnet werden; sexuelle Gewalt betrifft unterschiedslos Belästigungen wie auch Vergewaltigung. Gravierender erscheint aber, daß die Befragung zu jedem Delikt eingeleitet wurde mit der Formulierung: „Ist es Ihnen persönlich während der letzten fünf Jahre einmal passiert, daß es ...“. Die zeitliche Verortung erfolgte also nicht in freier Erinnerung, sondern der Referenzzeitraum wurde vorgegeben. Dieser Abfragemodus ist besonders anfällig für einen Fehler, der in der Fachliteratur als Telescopingeffekt bezeichnet wird (vgl. Skogan, 1981; Rubin & Baddley, 1989; Bradburn, Rips & Shevell, 1987). Insbesondere der Vorgang, daß frühere Delikte von den Befragten in diesen Zeitraum hinein verlagert werden, führt zur Überschätzung von Viktimisierungsraten.¹ Die von Kury et al. vorgelegten Analysen zum Regionalvergleich basieren auf den Prävalenzraten des Fünfjahreszeitraums vor der Erhebung. Bei Verwendung eines solch langen Referenzzeitraumes schlagen sich Probleme des Vergessens und der Veränderung von kognitiven Repräsentationen - für die einzelnen Delikte unterschiedlich, abhängig von Deliktschwere und subjektiver Relevanz des Erlebnisses - besonders stark nieder. Dies kann für die Untersuchung regionaler Unterschiede des Opferrisikos dann zu bedeutsamen Effekten führen, wenn - wie im hier vorliegenden Fall - die verglichenen Bundesländer z. B. in der Rate der von Armut betroffenen Personen ebenfalls Divergenzen aufweisen. So ist beispielsweise nicht auszuschließen, daß wohlhabende Personen bagatelhafte Eigentumsdelikte schneller vergessen als ärmere Befragte, für die dieses Erlebnis einen fühlbar stärkeren Verlust bedeutet.

Zur *Stichprobenqualität* merken die Autoren selbst an, daß der ICS 89 als Datenbasis angreifbar ist, da die Rücklaufquote dieser telefonischen Befragung lediglich 32% betrug. Es ist denkbar, daß dadurch in der Stichprobe diejenigen überrepräsentiert sind, die dem Interviewer etwas berichten wollten - die Opfer von Straftaten. Auffallend ist jedenfalls, daß die für den Fünfjahreszeitraum 1984 bis 1988 berechneten Opferraten des ICS 89 teilweise um das doppelte bis sechsfache über den Opferraten des zwei Jahre später durchgeführten MPI 90 liegen (Referenzzeitraum 1986 bis 1990).² Für diese Diskrepanz, die angesichts der sich überlappenden Zeiträume nicht als Abbild von Realität gewertet werden kann, bieten die Autoren keine Erklärungen. Da unklar ist, inwieweit die geringe Rücklaufquote zu einer systematischen Verzerrung der Stichprobe geführt hat, erscheint diese Studie für Schätzungen von Viktimisierungsraten in der Grundgesamtheit unseres Erachtens nicht geeignet.³ Die Rücklaufquote von MPI 90 ist demgegenüber mit 70,1% erheblich besser.

Für beide Studien gilt aber, daß die *Stichprobengröße* für einen Vergleich auf der Ebene einzelner Bundesländer nicht ausreicht. So beträgt die Anzahl der Befragten in Schleswig-Holstein bei MPI 90 lediglich n=94 (beim ICS 89 sind es n=207). Schätzungen

¹Eine Möglichkeit der Verringerung dieses Effektes ist die freie zeitliche Verortung eines Geschehens ohne Vorgabe eines Zeitraums, wie dies z. B. in der KFN-Opferbefragung (vgl. Wetzels, Greve, Mecklenburg, Bilsky & Pfeiffer, 1995) und auch in einer Schweizer Opferstudie (vgl. Killias, 1989) umgesetzt wurde.

² Würden die Daten dieser beiden Studien Realität abbilden, dann müßten sich innerhalb von weniger als zwei Jahren dramatische Veränderungen ereignet haben, in deren Folge sich die Opferraten der sich überschneidenden Fünfjahreszeiträume für Kontaktdelikte in Baden-Württemberg auf ein Drittel (von 23,3% auf 8,1%) und in Bayern auf die Hälfte (von 24,7% auf 12,4%) reduziert hätten. Die Rate für Einbruchsdelikte wäre in Schleswig-Holstein auf ein Sechstel (von 13,5% auf 2,1%) zurückgegangen, während nur beim Fahrraddiebstahl keine derart extremen Veränderungen stattgefunden hätten.

³Der Hinweis, daß die geringe Rücklaufquote für die einzelnen Bundesländer ähnlich gewesen sei, weshalb sich dies „nicht auf mögliche Unterschiede hinsichtlich der Regionalverteilung“ ausgewirkt haben „dürfte“ (Kury et al., 1995, S. 770) ist allenfalls eine Vermutung, da Angaben über die Verweigerer nicht vorliegen.

auf der Ebene von Bundesländern erfordern jedoch bereits dann, wenn eine Viktimisierungswahrscheinlichkeit von ca. 15% zu vermuten ist und die Schätzung eine Genauigkeit von nur $\pm 5\%$ haben soll, einen Stichprobenumfang von mindestens 400 Befragten⁴ (vgl. Yamane, 1981, S. 824).⁵ Für seltenere Ereignisse, wie etwa Gewaltdelikte, ist aber eine Genauigkeit von $\pm 5\%$ nicht mehr akzeptabel, da bei einer geschätzten Opferrate von 5% eine reale Verdopplung nicht mehr nachweisbar wäre. Wird eine Veränderung um ein Fünftel für relevant erachtet, so ist für ein Delikt mit einer vermuteten Wahrscheinlichkeit von 5% die Schätzgenauigkeit auf $\pm 1\%$ festzulegen. In diesem Falle läge die erforderliche Stichprobengröße für jede zu vergleichende Grundgesamtheit (z. B. Bundesland) bereits bei $N=1900$. Dies zeigt, daß Stichproben mit ca. 2000 Befragten zwar ausreichen, um Schätzungen für die BRD-West vorzunehmen. Gesicherte Aussagen über einzelne Bundesländer sind damit aber nicht mehr möglich.

Bei einer Stichprobenziehung über ein sogenanntes „random-route“ Verfahren wie in MPI 90 ergibt sich zusätzlich das Problem, daß die Anzahl der Sample Points bei kleinen Stichproben zu gering wird, was sich negativ auf die Stichprobenqualität auswirkt.⁶ Es treten räumliche Klumpungen auf, die dazu führen, daß die Stichprobe hinsichtlich soziodemographischer Merkmale (hier regionale Besonderheiten *innerhalb* eines Bundeslandes) nicht mehr in ausreichendem Maße ein repräsentatives Abbild der Grundgesamtheit darstellt. Verzerrungen der Stichprobe können zwar bis zu einem gewissen Grad durch Gewichtung ausgeglichen werden. Ist jedoch die Anzahl der Sample-points zu klein, kann dies, weil bestimmte Gebiete in der Stichprobe gar nicht vertreten sind, auch über Gewichtungen nicht mehr korrigiert werden.

Auch das *inferenzstatistische Vorgehen* bietet Anlaß zur Kritik. Die Autoren nehmen Punktschätzungen vor und verwenden zur statistischen Kontrolle ihrer Stichprobenergebnisse mehrfache χ^2 Tests. Sie geben deren Signifikanz unter Verwendung unterschiedlicher Irrtumswahrscheinlichkeiten von 5%, 1% und 0,1% an. Abgesehen davon, daß das Signifikanzniveau a priori festgelegt werden sollte, ist hier zu beachten, daß Kury et al. 4 Variablen zwischen 7 westlichen Bundesländern (Flächenstaaten) vergleichen, um festzustellen, ob sich die Bundesländer hinsichtlich ihrer Opferbelastung signifikant unterscheiden. Dies bedeutet, daß je Variable $(7 \times (7-1))/2 = 21$ Vergleiche durchgeführt wurden, insgesamt also 84 zum Teil abhängige Vergleiche. (Werden nur die Länder Schleswig-Holstein, Niedersachsen, Bayern und Baden-Württemberg verglichen, so sind dies immerhin noch $6 \times 4 = 24$ Vergleiche.) Bei solchen mehrfachen Vergleichen kommt es zu einer Erhöhung des Alpha-Fehlerrisikos (vgl. Stelzl, 1982). So beträgt bei 24 Vergleichen unter Beibehaltung einer nominellen Irrtumswahrscheinlichkeit von $\alpha = 0.05$ die tatsächliche Irrtumswahrscheinlichkeit $\alpha' = 1 - (1 - 0,05)^{24} = 0.69$, bei 84 Vergleichen gar 0.98.⁷ Zur Kontrolle eines derart hohen Fehlerrisikos sollte eine explizite α -Adjustierung vorgenommen werden, bei 24 zum Teil abhängigen Vergleichen nach der Formel

⁴ Das Risiko, daß die Rate für die Grundgesamtheit außerhalb dieses Intervalls liegt und die Stichprobe zufallsbedingt eine andere Schätzung liefert, wurde in diesem Fall bei 5% festgelegt.

⁵ Der erforderliche Stichprobenumfang nimmt dann, wenn der zu schätzende Prozentwert 50% ist, ein Maximum an. Bei geringeren Viktimisierungswahrscheinlichkeiten sind daher zwar die für eine Schätzung erforderlichen Stichproben kleiner, allerdings bei gleichbleibender Breite des Konfidenzintervalls (hier $\pm 5\%$) und gleichbleibendem Konfidenzkoeffizienten (hier 95%). Soll die Breite des Konfidenzintervalls verringert werden, wächst der erforderliche Stichprobenumfang (bei gleichem Konfidenzkoeffizienten und gleichbleibender Wahrscheinlichkeit des Ereignisses) quadratisch an (vgl. Bortz, 1989, S. 138). So beträgt, bei einer Wahrscheinlichkeit von 15% und einer Absicherung auf dem 95% Niveau, bei Halbierung des Konfidenzintervalls auf $\pm 2,5\%$ die erforderliche Stichprobengröße $N=1600$.

⁶ Bei durchschnittlich 5 Interviews je Sample Point (vgl. Kury et al. 1992, S.460) ergeben sich für Schleswig-Holstein ca. 19 Startadressen, d. h. Wohnumgebungen bzw. Straßenzüge.

⁷ Das Risiko eines nur zufallsbedingt signifikanten Ergebnisses beträgt dann nicht 5% sondern 69% bzw. 98%.

$\alpha'_{\text{adjust}} = 1/24 \times 0,05 = 0,002$ (bei 84 Vergleichen wäre $\alpha'_{\text{adjust}} = 0,0006$). Danach wäre bei Akzeptanz einer nominellen 5% Irrtumswahrscheinlichkeit das Alphaniveau auf mindestens .002 festzulegen. In MPI 90 würde dies dazu führen, daß bei Einschränkung des Vergleichs auf 4 Bundesländer nur noch der Unterschied zwischen Schleswig-Holstein und Baden-Württemberg im Bereich der Kontaktdelikte statistisch signifikant ist. (Wird der Vergleich auf 7 Bundesländer ausgedehnt, so ist keiner der berichteten Unterschiede auf Basis MPI 90 statistisch signifikant).⁸

Ein unseres Erachtens transparenteres Verfahren ist die Angabe von Intervallschätzungen über die Berechnung von Konfidenzintervallen. Diese lassen die stichprobenabhängige Schätzungsgenauigkeit unmittelbar ersichtlich werden.

Bezeichnen wir die Stichprobengröße mit N, die Viktimisierungsrate mit p, und mit q die Rate der Nichtviktimisierten (wobei $p+q=1$ und somit $q=1-p$, sowie $p=n$ der Viktimisierten/N), dann können (unter der Voraussetzung, daß $p \times q \times n \geq 9$) die Grenzen des 95% Vertrauensintervalls nach folgender Formel approximativ berechnet werden: $p_{\text{krit}1,2} = p \pm 1,96 (p \times q / n)^{1/2}$ (vgl. Bortz, 1989, S. 131 ff.). Die „wahre Rate p“ der Grundgesamtheit befindet sich dann mit einer Wahrscheinlichkeit von 95% in den Grenzen des Intervalls $p_1 < p < p_2$. Überlappen sich diese Konfidenzintervalle wechselseitig nicht, so ist mit einer Irrtumswahrscheinlichkeit von .0025 davon auszugehen, daß sich auch die Raten in den Grundgesamtheiten unterscheiden (das Fehlerrisiko entspricht also dem o. a. Erfordernis der Adjustierung bei mehrfachen Vergleichen).

3. Reanalyse der MPI-Daten

Die MPI-Daten wurden von uns entsprechend dem angegebenen Verfahren zur Berechnung von Konfidenzintervallen reanalysiert. Danach ergibt sich, daß die These einer im Vergleich zu den süddeutschen Ländern besonders hohen Kriminalitätsbelastung Schleswig-Holsteins nur sehr begrenzt bestätigt werden kann. So haben Kury et al. (1995, S. 771) in bezug auf Kontaktdelikte ihre Ergebnisse dahingehend zusammengefaßt, daß die Rate der Opfer in Schleswig-Holstein „etwa doppelt so hoch“ sei wie in Bayern oder Baden-Württemberg.

Tabelle 1: Viktimisierungsraten und Konfidenzintervalle nach Bundesland, Fünffjahreszeitraum, Daten aus MPI 90

	N	Kontaktdelikte		Einbruchsdelikte	
		Opferrate	95% - Intervall	Opferrate	95% - Intervall
SH	94	21,2%	12,9-29,5	2,1%	n zu klein
Nieds.	242	14,9%	10,4-19,4	2,9%	n zu klein
NRW	529	14,4%	11,4-17,4	3,4%	1,9-4,9
HE	173	10,4%	5,9-14,9	4,6%	n zu klein
Rhld.Pflz.	91	12,1%	5,4-18,8	5,5%	n zu klein
BAWÜ	305	8,1%	5,0-11,2	3,9%	1,7-6,1
BAY	361	12,4%	9,0-15,7	4,2%	2,1-6,3

Tabelle 1 zeigt jedoch überlappungsfreie Intervalle bei Kontaktdelikten nur beim Vergleich von Schleswig-Holstein und Baden-Württemberg. Ein statistisch bedeutsamer Unterschied im Vergleich zu Bayern besteht nicht. Wie gravierend sich die mangelnde Stichproben-

⁸Es ist zu beachten, daß mit einer Adjustierung des Alphaniveaus gleichzeitig das Beta-Risiko steigt (vgl. Bortz, 1984). Dieses ist abhängig von Teststärke, Effektgröße, Stichprobenumfang und Alphaniveau. Um z. B. einen Unterschied der Opferraten von 5% zwischen zwei Bundesländern identifizieren zu können, wenn die Opferrate für BRD-West 10% beträgt (also z. B. 7,5% in BAY und 12,5% in SH), ist schon bei $\alpha = .01$ und $\beta = .20$ eine Stichprobengröße von $N=868$ erforderlich.

größe auswirkt, zeigt sich deutlich daran, daß für Schleswig-Holstein die Intervallbreite 16,6% ausmacht. Das heißt, die Viktimisierungsrate wird so ungenau geschätzt, daß z. B. eine reale Verdopplung von 13% auf 26% durch diese Art von Studie nicht nachweisbar wäre.

Betrachtet man die Einzeldelikte, welche die Aggregatvariable der Kontaktdelikte bilden, so zeigt sich, daß ein statistisch bedeutsamer Unterschied wegen der geringen Stichprobengröße für die Delikte tätlicher Angriff/Bedrohung und Diebstahl nicht mehr nachweisbar ist. Für den Raub liegt das Konfidenzintervall in Baden-Württemberg so, daß schon die Punktschätzung für Schleswig-Holstein in dieses Intervall fällt, ein Unterschied also schon bei 5% Irrtumswahrscheinlichkeit nicht anzunehmen ist. Der signifikante Unterschied der beiden Bundesländer in der Aggregatvariable kommt erst durch die Kumulation der Einzeldelikte zustande. Die Einbeziehung der Kategorie der sexuellen Belästigung hat dabei einen erheblichen Effekt; hier bestehen die größten absoluten Unterschiede. Gerade deren Operationalisierung ist aber problematisch. So zeigten Kury et al (1992, S. 131), daß mehr als ein Drittel dieser Fälle ein strafrechtlich möglicherweise unerhebliches „freches Benehmen“ betreffen. Auch die Kategorie „tätlicher Angriff/Bedrohung“ ist nicht unproblematisch. Opfererlebnisse reichen hier von der Körperverletzung mit Waffengewalt bis hin zu Bedrohungen ohne jede physische Gewaltanwendung. Es ist unklar, ob letztere dem Straftatbestand der Nötigung entsprechen würden. Bei der Analyse der Deliktschwere für 84 derartige Vorfälle stellten Kury et al. (1992, S. 142) jedenfalls fest, daß 33,3% auf Bedrohungen entfielen, Vorfälle mit Waffen/Gegenständen machten nur 8,3% aus.

Für Einbruchsdelikte finden sich bereits in der Stichprobe von MPI 90 der Hypothese eines Nord-Süd-Gefälles entgegenstehende Raten. Wegen der geringen Stichprobengröße sind Intervallschätzungen jedoch mit Ausnahme von Nordrhein-Westfalen, Baden-Württemberg und Bayern nicht möglich. Die Studie läßt eine Analyse von Nord-Süd Unterschieden für dieses Delikt gar nicht zu. Die Autoren sehen dies auch, behalten jedoch unter Verweis auf die Ergebnisse des ICS 89 die Hypothese eines auch insoweit bestehenden Nord-Süd-Gefälles bei (vgl. Tabelle 2). Wird aber ungeachtet der mangelhaften Rücklaufquote der ICS 89 für hinreichend reliabel und valide erachtet, so hätte auch festgestellt werden müssen, daß bei Kontaktdelikten kein statistisch bedeutsamer Unterschied besteht. Weiter ist die Opferrate Niedersachsens in der Stichprobe ICS 89, anders als in MPI 90, am niedrigsten. Zu kritisieren ist ferner, daß die Erhebung der Delikte in ICS 89 und MPI 90 nahezu identisch erfolgte, die Autoren aber nicht bzw. nur am Rande die erheblichen Unterschiede der festgestellten Viktimisierungsraten problematisieren. Das Vorgehen, im Befund wie in der Zuverlässigkeit (Rücklaufquote) divergente Stichproben selektiv oder zumindest jedenfalls unsystematisch heranzuziehen, kann unseres Erachtens die Glaubwürdigkeit der so belegten Ergebnisse und Argumente nicht eben erhöhen.

Tabelle 2: Viktimisierungsraten und Konfidenzintervalle nach Bundesland
Fünfjahreszeitraum, Daten aus ICS 89

	N	Kontaktdelikte		Einbruchsdelikte	
		Opferrate	95%-Intervall	Opferrate	95%-Intervall
SH	207	27,5%	21,5-33,5	13,5%	8,8-18,2
Nieds.	652	21,3%	18,1-24,5	8,9%	6,7-11,1
BAWÜ	786	23,3%	20,0-26,6	5,6%	4,0-7,2
BAY	881	24,7%	21,9-27,5	6,7%	5,0-8,4

Bei der Sammelkategorie der Kfz-Delikte (Tabelle 3) wird die Interpretation der Befunde dadurch erschwert, daß diese sich aus sehr unterschiedlichen Einzeldelikten zusammensetzt (Diebstahl von Kfz, Diebstahl aus Kfz, Kfz-Sachbeschädigung und Motorrad Diebstahl), bei denen erhebliche Unterschiede von Versicherungsgrad, Schadens-

summen, aber auch Opferraten bestehen. So führen Kury et al. (1995, S. 772) aus, daß in Schleswig-Holstein und Niedersachsen im Vergleich zu den beiden südlichen Ländern „etwa doppelt so viele Fahrzeuge gestohlen oder beschädigt“ wurden. Auf Ebene der Einzeldelikte ist aber festzustellen, daß in Schleswig-Holstein kein einziger Befragter aus MPI 90 Opfer eines Kfz-Diebstahls war (vgl. Kury et al., 1992, S. 214). Betrachtet man die Ergebnisse aus MPI 90, so fällt wiederum die enorme Breite des Schätzintervalls von $\pm 9,2\%$ für Schleswig-Holstein auf. Infolgedessen überlappen sich die Intervalle aller vier Bundesländer.

Tabelle 3: *Viktimisierungsraten und Konfidenzintervalle nach Bundesland Kfz-Delikte im Fünfjahreszeitraum, Daten aus ICS 89 und MPI 90*

	SH		Nieds.		BAWÜ		BAY	
ICS 89	(N=182)		(N=541)		(N=676)		(N=729)	
	Rate	95%-Intervall	Rate	95%-Intervall	Rate	95%-Intervall	Rate	95%-Intervall
Kfz-Delikte	47,3%	40,0-54,6	33,8%	29,8-37,8	33,0%	29,5-36,5	32,8%	29,9-35,7
MPI 90	(N=74)		(N=162)		(N=245)		(N=286)	
	Rate	95%-Intervall	Rate	95%-Intervall	Rate	95%-Intervall	Rate	95%-Intervall
Kfz-Delikte	20,3%	11,1-29,5	24,6% ¹	18,0-31,2	15,9%	11,3-20,5	18,2%	13,8-22,6

¹ Für Niedersachsen ist die Stichprobengröße nicht eindeutig; nach Kury et al., 1992 müßten dies 185 Befragte sein, nach Kury et al., 1995 jedoch 162.

Zudem zeigt sich erneut, daß die Ergebnisse der beiden Studien zu verschiedenen Schlüssen führen: Während sich auf der Basis von MPI 90 Unterschiede nicht nachweisen lassen, finden sich im ICS 89 signifikante Divergenzen. Trotz identischer Fragen und eines sich weitgehend überlappenden Befragungszeitraums liegen die Viktimisierungsraten des ICS 89 teilweise um mehr als das Doppelte über denen aus MPI 90. Danach wären die Kfz-Delikte im Vergleich der beiden sich überlappenden Fünfjahreszeiträume in Schleswig-Holstein in kurzer Zeit auf weniger als die Hälfte zurückgegangen. In den beiden südlichen Bundesländern hätte es ebenfalls eine erhebliche Reduktion gegeben: Wie Kury selbst gezeigt hat, können Divergenzen in dieser Größenordnung nicht auf den Unterschied von telefonischer versus persönlich-mündlicher Befragung zurückgeführt werden (vgl. Kury, 1994). Trotzdem verzichten die Autoren darauf, diese Widersprüche zu kommentieren.

Tabelle 4: *Viktimisierungsraten und Konfidenzintervalle nach Bundesland Fahrraddiebstahl im Fünfjahreszeitraum, Daten aus ICS 89 und MPI 90*

	SH		Nieds.		BAWÜ		BAY	
ICS 89	(N=172)		(N=570)		(N=622)		(N=669)	
	Rate	95%-Intervall	Rate	95%-Intervall	Rate	95%-Intervall	Rate	95%-Intervall
Fahrraddieb.	26,2%	19,6-32,8	18,6%	15,4-21,8	12,7%	10,1-15,3	12,4%	9,9-14,9
MPI 90	(N=67)		(N=186)		(N=214)		(N=245)	
	Rate	95%-Intervall	Rate	95%-Intervall	Rate	95%-Intervall	Rate	95%-Intervall
Fahrraddieb.	23,9%	13,7-34,1	18,8%	13,2-24,4	11,2%	7,0-15,4	14,3%	9,9-18,7

Ein nur teilweise ähnliches Ergebnis zeigt sich für den Fahrraddiebstahl (Tabelle 4). Obwohl in der Stichprobe die Opferrate für Schleswig-Holstein doppelt so hoch ist wie jene

für Baden-Württemberg, ist ein Nord-Süd-Gefälle nicht mit der erforderlichen Sicherheit nachweisbar. Wegen der enormen Breite des Konfidenzintervalls für Schleswig-Holstein von $\pm 10,2\%$ in MPI 90 liegt dessen unterer Intervallwert von $13,7\%$ unter der oberen Intervallgrenze Baden-Württembergs von $15,4\%$. Zwar zeigen sich beim Fahrraddiebstahl nicht die extremen Unterschiede der Opferraten zwischen ICS 89 und MPI 90. Angesichts der zu den anderen Delikten jedoch auftauchenden Zweifel an der Qualität des ICS 89 wäre es allerdings verfehlt, in diesem Fall, wo das Ergebnis scheinbar „paßt“, daraus zu folgern, man habe einen empirischen Beleg für einen Nord-Süd Unterschied gefunden.

Fassen wir die zentralen Ergebnisse unserer Reanalysen zusammen, so ist festzustellen:

1. Die Daten aus MPI 90 führen zu dem Ergebnis, daß lediglich im Bereich der Kontaktdelikte ein Unterschied zwischen Schleswig-Holstein und Baden-Württemberg anzunehmen ist. In allen anderen Deliktsbereichen sind signifikante Unterschiede nicht nachweisbar. Der Unterschied im Bereich der Kontaktdelikte ist wiederum primär Folge der Aggregation von Einzeldelikten. Getrennte Analysen für die Einzeldelikte haben keine signifikanten Abweichungen zwischen Schleswig-Holstein und Baden-Württemberg nachweisen können. Die von Kury et al. hervorgehobenen Divergenzen auf Basis von Punktschätzungen können nur als Hinweise auf einen möglichen Unterschied zwischen Nord und Süd gewertet werden. Eine Entscheidung über Annahme oder Verwerfung dieser Hypothese ist auf Basis ihrer Daten jedoch nicht möglich.
2. Die Daten des ICS 89 sind für die Schätzung von Opferraten, wegen niedriger Rücklaufquoten, unzuverlässig. Die Zweifel an der Datenqualität werden durch die hohen Unterschiede der Opferraten bestätigt, die sich im Vergleich von ICS 89 und MPI 90 zu denselben Bundesländern zeigen. Hierfür wird ebensowenig eine plausible Erklärung angeboten wie für die widersprüchlichen Ergebnisse, die zu Kontakt und Einbruchsdelikten in diesen beiden Studien gefunden wurden.
3. Zwei zentrale Mängel, die niedrige Rücklaufquote von ICS 89 und die zu geringen Stichprobengrößen für einen Vergleich von Bundesländern, wurden durch den Landesinnenminister Schleswig-Holsteins in seiner Stellungnahme vom 2.1.96 angesprochen. Wenn Kury in seiner Antwort vom 12.01.1996 darauf verweist, die Kritik sei unberechtigt, weil in beiden Studien insgesamt 13.000 Menschen befragt worden seien, so weicht er den entscheidenden Punkten der Kritik aus. Mit der schlichten Addition zweier Stichproben, von denen eine zum größten Teil Befragte aus den neuen Bundesländern enthält, mag er zwar eine beeindruckende Gesamtzahl von Befragten demonstrieren. Es bleibt aber das Problem der mangelhaften Rücklaufquote des ICS 89 bestehen sowie der Umstand, daß die Stichproben für Schleswig-Holstein in *beiden* Studien mit 207 bzw. 94 Befragten für eine Analyse von Viktimisierungsrisiken erheblich zu klein sind.

4. Die KFN-Opferbefragung 1992

Die Hypothese eines Nord-Süd-Gefälles der Opferbelastung wird durch die beiden von Kury et al. analysierten Opferbefragungen nicht hinreichend gestützt, kann aber - da dies vor allem auf methodische Unzulänglichkeiten zurückzuführen ist - auch nicht verworfen werden. Immerhin fanden sich in der Stichprobe Unterschiede, die - obschon nicht signifikant - die Möglichkeit offenlassen, daß bei methodisch adäquaterem Vorgehen die Hypothese eines solchen Gefälles der Kriminalitätsbelastung erhärtet werden könnte. Die Daten der Polizeilichen Kriminalstatistik geben ebenfalls Anlaß dazu, einen solchen Unterschied als Möglichkeit zu betrachten. Ferner sprechen auch theoretische Überlegungen

für eine solche Annahme. Wir stimmen mit Kury et al. darin überein, daß die in den nördlichen Bundesländern im Vergleich zum Süden größeren Anteile von Menschen in wirtschaftlich und sozial angespannter Lebenslage, ausgedrückt über die Anteile der Sozialhilfeempfänger, Einkommensunterschiede und höhere Arbeitslosenraten, sowie die höhere Ausprägung von Indikatoren sozialer Desorganisation dazu beitragen können, daß sich die Kriminalitätsbelastung dieser Regionen unterscheidet. So haben zwei am KFN durchgeführte Untersuchungen zu regionalen Unterschieden der Kriminalitätsbelastung von Landkreisen und kreisfreien Städten in Nordrhein-Westfalen und Niedersachsen einen signifikanten Zusammenhang polizeilich registrierter Kriminalität mit Merkmalen sozialer Desorganisation und Deprivation nachweisen können (vgl. Ohlemacher & Pfeiffer, 1995; Ohlemacher, 1995; Pfeiffer, 1995, S. 91 ff.).⁹

Anomie- und kontrolltheoretische Überlegungen lassen insoweit aufgrund der sozialen und wirtschaftlichen Unterschiede, wie sie Kury et al. auf Basis amtlicher Statistiken für die Bundesländer zutreffend beschrieben haben¹⁰, im Norden der Bundesrepublik höhere Viktimisierungsraten erwarten. Offen ist allerdings die Frage, ob dies für alle Deliktsbereiche gilt und welches Ausmaß diese Unterschiede aufweisen. Diesen beiden Fragen werden wir im folgenden auf Basis der Daten aus der 1992 durchgeführten repräsentativen Opferbefragung des KFN nachgehen (vgl. Wetzels et al., 1995). Gegenüber MPI 90 sowie ICS 89 erscheint die Datenbasis der KFN-Opferbefragung (im folgenden KFN 92) in mehreren Punkten zuverlässiger:

- Die Gesamtstichprobe ist mit 8966 Befragten in Westdeutschland wesentlich größer, Schätzungen sind also genauer als bei den beiden MPI-Studien. In Schleswig-Holstein wurden so z. B. 377 Befragte erfaßt. Die Anzahl der Sample Points ist in diesem nach dem Saarland kleinsten Flächenland mit 72 ebenfalls wesentlich höher als in MPI 90 (19 Sample Points), wodurch das Risiko der Klumpung deutlich geringer ist.
- Im Gegensatz zum ICS 89 ist die Rücklaufquote von KFN 92 mit 73,8% in den alten Bundesländern sehr zufriedenstellend. Ein Vergleich mit sozialstatistischen Daten der Grundgesamtheit zeigt darüber hinaus, daß diese Stichprobe als repräsentativ angesehen werden kann (vgl. Wetzels et al., 1995). Gewichtungen können die Schätzung von Opferraten im Vergleich zu ungewichteten Daten nicht weiter verbessern, ein weiterer Hinweis auf die Datenqualität.
- Der Abfragemodus ist so gestaltet, daß Referenzzeiträume nicht vorgegeben werden, die zeitliche Verortung vielmehr in freier Erinnerung erfolgt, wodurch Telescopingeffekte reduziert werden.
- Die Befragung fand ab Januar 1992 statt, wodurch der Jahreswechsel als Ankerpunkt der zeitlich Verortung gewählt werden konnte, was Fehler in der zeitlichen Einordnung ebenfalls verringert (vgl. Loftus & Marburger, 1983).
- Der Abfragemodus erlaubt die Berechnung von Prävalenzraten für beliebige ganzjährige Referenzzeiträume zwischen 1987 und 1991.

⁹Gestützt werden diese Annahmen auch durch die Ergebnisse einer von Pfeiffer (1995) durchgeführten Analyse der Kriminalitätsbelastung verschiedener Altersgruppen der west- und ostdeutschen Bevölkerung. Danach haben die Tatverdächtigenziffern seit Ende der 80er Jahre bei den Gruppen am stärksten zugenommen, deren soziale Situation sich in diesem Zeitraum am deutlichsten verschlechtert hat, den Jugendlichen und Heranwachsenden.

¹⁰Diese Unterschiede lassen sich im übrigen auch in den Daten der KFN-Opferbefragung 1992 nachweisen. So hat die Stichprobe für die nördlichen Bundesländer die höchste Arbeitslosenrate (8,5%; Mitte: 6,3%; Süd 5,5%), den höchsten Anteil an Haushalten mit einem pro-Kopf-Nettoeinkommen von unter 800 DM (Nord: 10,7%; Mitte: 5,9%; Süd: 5,9%), während der Süden den höchsten Anteil an Befragten mit einem pro-Kopf-Nettoeinkommen des Haushaltes von über 3000 DM aufweist (Süden: 8,2%; Mitte: 7,2%; Norden: 5,9%).

- Die Opfererfahrungen wurden differenzierter, anstelle über 11 wie in Kury et al., über 16 verschiedene Deliktsarten erhoben. So wurde z. B. explizit zwischen Vergewaltigung/sexueller Nötigung einerseits und sexueller Belästigung andererseits unterschieden und die Körperverletzung unterteilt in solche mit und ohne Waffen/Gegenständen.

Wir haben an anderer Stelle bereits darauf hingewiesen, daß nach unseren Ergebnissen zwar Nord-Süd Unterschiede in der Stichprobe bestehen, diese aber wesentlich niedriger ausfallen, als nach der PKS anzunehmen wäre (Pfeiffer & Wetzels, 1994). In der damaligen Publikation haben wir uns auf Analysen gestützt, bei denen wir die BRD-West in eine nördliche, südliche und mittlere Region unterteilten. Durch diese Zusammenfassung von Bundesländern waren die Schätzungen der regionalen Viktimisierungsraten wegen der größeren Stichproben relativ genau. Um jedoch die strittige Frage einer besonderen Belastung Schleswig-Holsteins zu prüfen, werden wir im folgenden zunächst die Auswertung auf der Ebene einzelner Bundesländer vornehmen. Wir sind uns allerdings bewußt, daß die Anzahl der Befragten in Schleswig-Holstein, obschon deutlich höher als in MPI 90 und ICS 89, dennoch problematisch klein ist, was zu entsprechend großen Intervallen bei den Schätzungen führt. In einem zweiten Schritt werden wir daher die Ergebnisse für den Nord-Süd Vergleich detaillierter darstellen.

4.1 Analyse der Prävalenzraten für ausgewählte Bundesländer

In Anlehnung an die von Kury et al. vorgelegte Studie werden im folgenden die Opferraten des Fünfjahreszeitraums von 1987 bis 1991 von Schleswig-Holstein im Vergleich zu Bayern und Baden-Württemberg für einzelne Deliktskategorien dargestellt (Tabelle 5).

Schwere Gewalt umfaßt dabei die Raubdelikte sowie Vergewaltigung und Körperverletzung mit Waffen; leichte Gewalt enthält die einfache Körperverletzung sowie Bedrohung/Nötigung. Der schwere Diebstahl umfaßt Diebstahl von/aus Kfz, Motorrad Diebstahl sowie Einbruchdelikte. Fahrraddiebstahl, sonstiger Diebstahl, Betrug und Kfz-Sachbeschädigung werden als Einzeldelikte dargestellt. Die sexuelle Belästigung wurde wegen ihrer ungenauen Abgrenzung in die Analyse nicht einbezogen.¹¹ Um den Vergleich mit den Ergebnissen aus Kury et al. (1995) zu erleichtern wird der Einbruchdiebstahl (Teilmenge des schweren Diebstahls) zusätzlich gesondert dargestellt. Ferner werden auch die Kfz-Delikte (Diebstahl von/aus Kfz, Motorrad Diebstahl sowie Kfz-Sachbeschädigung) in einer Kategorie zusammengefaßt und dargestellt.

Von einer einheitlich höheren Opferbelastung Schleswig-Holsteins ist offenbar in bezug auf den Fünfjahreszeitraum nicht auszugehen. Vielmehr ist zwischen verschiedenen Delikten zu unterscheiden. So erscheint Schleswig-Holstein im Bereich der schweren Gewaltdelikte signifikant weniger belastet als Bayern, wohingegen sich bei den leichten Gewaltdelikten kein signifikanter Unterschied mehr nachweisen läßt. Bei schwerem Diebstahl sowie den zu dieser Kategorie zählenden Einbruchdelikten findet sich ebenfalls kein für die Grundgesamtheit relevanter Unterschied. In der Stichprobe ist Bayern sogar etwas stärker belastet. Beim Betrug sind schon die Opferraten der Stichproben annähernd gleich. Anders verhält es sich demgegenüber mit den leichteren Eigentumsdelikten. So ist Schleswig-Holstein im Vergleich zu beiden südlichen Bundesländern statistisch signifikant höher mit Fahrraddiebstahl und Kfz-Sachbeschädigungen belastet. Beim sonstigen Dieb-

¹¹ Hier liegen die Raten Schleswig-Holsteins (5,3%) unter denen von Bayern (8,0%) und Baden-Württemberg (8,2%). Ihre Nichtberücksichtigung wirkt sich also entgegen unserer Hypothese aus. Beachtet man weiter, daß derartige Delikte im Norden auch häufiger angezeigt werden (10,3% vs. 8,3% im Süden), dann ist festzustellen, daß auch unsere Hypothese eines höheren Dunkelfeldes im Süden gegenüber dem Norden somit schwieriger zu bestätigen sein wird.