

# Gibt es eine Lohnkurve in den neuen Bundesländern?

Herbert S. Buscher\*

Mit Hilfe von Querschnittsregressionen und Panelschätzungen wird für den Zeitraum von 1998 bis 2002 geprüft, ob eine Lohnkurve in den neuen Bundesländern existiert. Die Daten stammen aus dem IAB-Betriebspanel und beschränken sich auf das Produzierende Gewerbe in Ostdeutschland. Insgesamt zeigen die Ergebnisse, dass von der Existenz einer Lohnkurve ausgegangen werden kann. Jedoch liegen die geschätzten Lohnelastizitäten in einem Bereich, der absolut gesehen das 2- bis 3fache beträgt, was üblicherweise in der Literatur publiziert wird.

## Gliederung

- 1 Einleitung
  - 2 Die wirtschaftliche Entwicklung in den neuen Bundesländern
  - 3 Empirische Evidenz für Deutschland
  - 4 Modelltheoretische Erklärungen der Lohnkurve und Spezifikation der Gleichungen
  - 5 Die Daten
  - 6 Ergebnisse der Schätzungen
    - 6.1 Schätzergebnisse für den gepoolten Datensatz, 1998-2002
    - 6.2 Schätzergebnisse für Random-Effects-Modelle, 1998-2002
  - 7 Schlussfolgerungen
- Literatur
- Anhang 1: Weitere Regressionsergebnisse
- Anhang 2: Daten

## 1 Einleitung

Die zentrale Aussage der „Lohnkurve“ lautet, dass in Regionen mit hoher Arbeitslosigkeit das Niveau des Nominallohnes niedriger ist als in Regionen mit vergleichsweise geringer Arbeitslosigkeit, also ein negativer Zusammenhang zwischen der regionalen Arbeitslosigkeit und dem regionalen Lohnniveau auf empirischer Basis besteht. Blanchflower und Oswald (1994b) kommen in ihrer Monografie zu dem Schluss, dass ein derart stabiler Zusammenhang zwischen den beiden zentralen Größen regionale Arbeitslosigkeit und regionales Lohnniveau besteht, dass praktisch von einer „Gesetzmäßigkeit“ gesprochen werden kann<sup>1</sup>.

Seit den ersten Veröffentlichungen von Blanchflower und Oswald (1990, 1994b, 1995) zur Lohnkurve setzte eine Vielzahl von weiteren Studien zur Lohnkurve ein, die den behaupteten empirischen Zusammenhang für verschiedene Länder und Regionen überprüften. Für Westdeutsch-

land wurde die Lohnkurve für unterschiedliche regionale Abgrenzungen und Zeiträume geschätzt. Im Gegensatz zu den empirischen Ergebnissen für angelsächsische Länder fiel hier die empirische Evidenz eher gemischt aus<sup>2</sup>. Als ein Grund hierfür wird angeführt, dass die Lohnverhandlungen in (West-)Deutschland zwar dezentral durchgeführt werden, aber praktisch über ihre Signalwirkung dann im Wesentlichen für ganz Deutschland gelten. Dieses Argument wird zusätzlich durch die Möglichkeit der Allgemeinverbindlichkeitserklärung von Tarifverträgen durch das (frühere) Bundesarbeitsministerium bzw. dem heutigen Bundesministerium für Wirtschaft und Arbeit (BMWA) erhärtet.

Die für Deutschland vorliegende empirische Evidenz zur Lohnkurve bezieht sich ganz überwiegend auf die alten Bundesländer<sup>3</sup>; nur wenige Arbeiten beschäftigen sich mit ostdeutschen Regionen<sup>4</sup>. Ein Grund hierfür mag sein, dass noch kein ausreichender Datensatz zur Verfügung stand und/oder dass infolge des Anpassungsprozesses der ostdeutschen an die westdeutsche Wirtschaft – insbesondere bezüglich des Lohnniveaus – eine solche Überprüfung als wenig aussichtsreich erschien. Mittlerweile je-

\* Dr. Herbert S. Buscher ist Leiter der Abteilung für Arbeitsmarkt des Instituts für Wirtschaftsforschung Halle. Der Beitrag liegt in der alleinigen Verantwortung des Autors. Er wurde im September 2003 eingereicht und nach der Begutachtung und einer Revision im Januar 2004 zur Veröffentlichung angenommen.

Für die Unterstützung bei der Aufbereitung der Daten bedanke ich mich bei Lutz Bellmann, Barbara Schwengler und Uwe Blien, jeweils IAB Nürnberg, bei Frau Antonia Milbert und Steffen Maretzke, beide BBR Bonn, bei Gisela Hardt und Frau A. Toth sowie Michael Barkholz, jeweils IWH Halle, für die Hilfe bei der Durchführung der Untersuchungen. Zwei anonymen Gutachtern sei an dieser Stelle ausdrücklich für ihre wertvollen Hinweise gedankt.

<sup>1</sup> In seiner Buchbesprechung führt Card (1995) zwar einige Bedenken und Einwände gegenüber dieser Schlussfolgerung an, kommt jedoch ebenfalls zu dem Ergebnis, dass von einem negativen Zusammenhang zwischen beiden Variablen ausgegangen werden muss.

<sup>2</sup> Siehe u.a. Wagner (1994, 1996a, b), Bellmann/Blien (2001), die verschiedenen Beiträge in MittAB 29 (3) / 1996, Rendtel/Schwarze (1995), Gerlach/Schettkat (Hrsg.) (1995), Schwarze (1995), Schmidt (1992).

<sup>3</sup> Siehe z.B. Gerlach/Wagner in Gerlach/Schettkat (Hrsg.) (1995) sowie die Beiträge in MittAB 29 (3) / 1996, S. 455-502.

<sup>4</sup> Siehe z.B. Schwarze (1995), Pannenberg/Schwarze (1996a, b), Baltagi/Blien/Wolf (2000) sowie Elhorst/Blien/Wolf (2003).

**Tabelle 1: Bruttolöhne und -gehälter je Arbeitnehmer (Inland), alle Wirtschaftsbereiche**

Zeit	Messeinheit	Westdeutschland (mit Berlin)	Westdeutschland (ohne Berlin)	Ostdeutschland (mit Berlin)	Ostdeutschland (ohne Berlin)
1991	Euro	21 805	21 962	12 616	11 071
1995		25 099	25 155	19 888	18 793
2000		26 436	26 477	21 470	20 404
2002		27 282	27 329	22 212	21 140
1991	1995=100	86.9	87.3	63.4	58.9
1995		100.0	100.0	100.0	100.0
2000		105.3	105.3	108.0	108.6
2002		108.7	108.6	111.7	112.5
1991	Deutschland = 100	110.0	110.8	63.6	55.8
1995		104.4	104.6	82.7	78.1
2000		103.7	103.8	84.2	80.0
2002		103.5	103.7	84.3	80.2

Quelle: Arbeitskreis Volkswirtschaftliche Gesamtrechnung der Länder

doch liegen Datensätze vor, die es erlauben, die Existenz einer Lohnkurve auch für Ostdeutschland zu überprüfen, ohne dass hierbei auf die Daten in der ersten Hälfte der neunziger Jahre rekurriert werden muss.

Die Frage, ob für Ostdeutschland auf empirischer Basis eine Lohnkurve gefunden werden kann, ist aus mehreren Gründen interessant. Im Gegensatz zu Westdeutschland ist der Anteil tarifgebundener Betriebe in Ostdeutschland erheblich niedriger. Eine Folge hieraus ist, dass viele Arbeitnehmer zum Teil deutlich unter Tarif bezahlt werden. Als „Alternative“ zur Arbeitslosigkeit sind sie bereit zu einem untertariflichen Lohn zu arbeiten, wenn im Gegenzug hierzu ihr Beschäftigungsverhältnis bestehen bleibt. Folglich dürfte hier, im Gegensatz zu den westlichen Bundesländern, das Argument einer „praktisch“ zentralisierten Lohnverhandlung – und somit eine zu geringe Variation zwischen den Löhnen auf regionaler sowie auf Branchenebene – nur begrenzt zutreffen. Andererseits liegt die Arbeitslosigkeit in den neuen Bundesländern deutlich über dem West-Niveau, was im Wesentlichen durch die strukturellen Anpassungsprozesse bedingt ist. Dies bedeutet, dass sich die ostdeutschen Regionen in einem Lohn-Arbeitslosigkeits-Diagramm relativ weit rechts auf der Abszisse bewegen, sodass nicht a priori die „typischen“ Ergebnisse für eine Lohnkurve erwartet werden können. Ferner wurde und wird ein Teil der ostdeutschen Arbeitslosigkeit in die alten Bundesländer „exportiert“, was durch die hohen Pendlerzahlen zwischen Ost und West sowie durch die teils kräftigen Abwanderungen arbeitsfähiger Personen belegt werden kann. Schließlich weisen die neuen Bundesländer im Durchschnitt eine höhere Frauenerwerbsquote als die westlichen auf. Dies führt zu einem höheren Arbeitspotenzial und zum anderen zu teilweise niedrigeren Löhnen, da Frauen vielfach niedriger als ihre männlichen Kollegen entlohnt werden. Diese wesentlichen Unterschiede in den Entwicklungen zwischen West- und Ostdeutschland lassen es sinnvoll erscheinen zu überprüfen, ob eine ostdeutsche Lohnkurve besteht.

Ein Unterschied zu bestehenden Untersuchungen zur ostdeutschen Lohnkurve besteht in dem verwendeten Datensatz. Baltagi/Blien/Wolf (2000) und Elhorst/Blien/Wolf (2003) schätzen die Lohnkurve mit Daten der Beschäftigtenstatistik für 114 Landkreise; Pannenberg/Schwarze (1996a, b) hingegen benutzen die Daten des Sozio-ökonomischen Panels (SOEP) und erfassen alle Branchen der Wirtschaft. Für Ostdeutschland liegen meines Erachtens keine Schätzergebnisse vor, die auf der Grundlage des IAB-Betriebspanels Ost erzielt wurden. Dieser Datensatz, mit einer Beschränkung auf das ostdeutsche Produzierende Gewerbe, findet in dieser Arbeit Anwendung.

Die Arbeit ist wie folgt gegliedert: Im nächsten Abschnitt wird ein kurzer Überblick über wichtige ökonomische Daten für Ostdeutschland gegeben. Danach folgt eine Übersicht über die bisherigen empirischen Ergebnisse zur Lohnkurve in Ostdeutschland. Abschnitt 4 beschreibt den Datensatz und stellt die zu testenden Spezifikationen vor, deren Ergebnisse dann im fünften Abschnitt präsentiert werden. Der Beitrag schließt mit einigen Schlussfolgerungen, die aus den empirischen Ergebnissen folgen.

## 2 Die wirtschaftliche Entwicklung in den neuen Bundesländern

Bereits kurz nach der Wiedervereinigung setzte ein rascher Anpassungsprozess der ostdeutschen Löhne und Gehälter an das westdeutsche Niveau ein, dessen Dynamik aber seit 1995 deutlich abflachte. Tabelle 1 zeigt diese Entwicklung für die Bruttolöhne und -gehälter je Arbeitnehmer über alle Wirtschaftsbereiche<sup>5</sup>.

Von 1991 bis 1995 stiegen die Bruttolöhne und -gehälter je Arbeitnehmer in Ostdeutschland (ohne Berlin) um 41,1

<sup>5</sup> Die Lohn- und Gehaltsentwicklung für das Produzierende Gewerbe sind im Tabellenanhang in den Tabellen A1 und A2 dargestellt.

**Tabelle 2: Lohnentwicklung im Produzierenden Gewerbe 1995–2001<sup>1</sup>**

	Westdeutschland	Ostdeutschland
<b>Bruttostundenverdienst</b>		
Nominal	11,9 %	15,4 %
Real (Verbraucherpreise) <sup>2</sup>	2,3 %	4,3 %
Real (Erzeugerpreise)	6,7 %	10,0 %
<b>Bruttomonatsverdienst</b>		
Nominal	13,1 %	20,3 %
Real (Verbraucherpreise)	3,4 %	8,8 %
Real (Erzeugerpreise)	7,8 %	14,7 %

Quelle: SVR JG 2002/03, eigene Berechnungen.

<sup>1</sup>) Verbraucherpreise = Preisindex für die Lebenshaltung aller privaten Haushalte, 1995=100; Erzeugerpreise = Erzeugerpreise, insgesamt, gewerbliche Produkte. Bei den Verbraucherpreisen wurden jeweils die Indizes für West- und Ostdeutschland verwendet, für die Erzeugerpreise wurden beide Regionen gleich behandelt.

<sup>2</sup>) Bruttostunden- und Bruttomonatsverdienste wurden jeweils mit dem Verbraucherpreisindex bzw. dem Erzeugerpreisindex deflationiert, um die Konsumenten- bzw. Produzentenlöhne zu bestimmen.

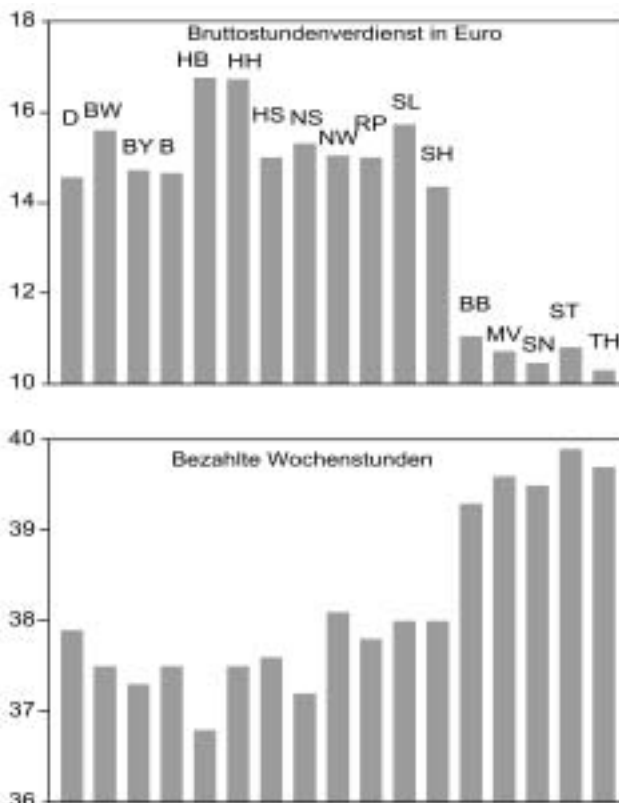
Indexpunkte; seitdem verlangsamte sich dieser Anstieg von 1995 bis 2002 auf insgesamt 12,5 Indexpunkte. Gleichwohl stiegen in den neuen Bundesländern seit 1995 die Löhne und Gehälter etwas stärker als in den alten Bundesländern, wo der Anstieg nur 8,7 Indexpunkte ausmacht. Verglichen mit dem Bundesdurchschnitt erreichten die ostdeutschen Löhne und Gehälter 2002 ein Niveau von etwa 80 %.

Verglichen mit den Nominallohnsteigerungen fielen die Realloohnerhöhungen zwischen 1995 und 2001 in beiden Teilen Deutschlands wesentlich moderater aus. Wie Tabelle 2 zeigt, stiegen die Produzentenlöhne in Westdeutschland um 6,7 % und in Ostdeutschland um 10,0 %, bezogen auf die Bruttostundenverdienste im Produzierenden Gewerbe. Für den gleichen Zeitraum belaufen sich die Nominalloohnerhöhungen auf 11,9 % im Westen und auf 15,4 % im Osten. Bemerkenswert ist, dass die Real-löhne aus Sicht der Konsumenten deutlich geringer anstiegen (2,3 % im Westen und 4,3 % im Osten) als die Produzentenlöhne. Für alle ausgewiesenen Veränderungs-raten gilt jedoch, dass die Lohnsteigerungen in den neuen Bundesländern zwischen 1995 und 2001 deutlich über den Raten in den westlichen Bundesländern liegen.

Abbildung 1 zeigt die Bruttostundenverdienste und die bezahlten Wochenstunden nach Bundesländern untergliedert. In beiden Grafiken ist das Ost-West-Gefälle deutlich sichtbar: Während in den westlichen Bundesländern die bezahlten Wochenarbeitsstunden zwischen 37 und 38 Stunden schwanken, befinden sie sich für Ostdeutschland im Bereich zwischen 39 und 40 Wochenarbeitsstunden.

Andererseits liegt der Bruttostundenverdienst im Westen mit einem Stundenlohn zwischen 14 und 16 Euro (wenn man von den Stadtstaaten absieht) deutlich über dem Ost-niveau. In den neuen Bundesländern werden im Durch-schnitt zwischen 10 und 11 Euro die Stunde gezahlt.

**Abbildung 1: Bruttostundenverdienste und bezahlte Wochenstunden im Produzierenden Gewerbe, 2003**



Quelle: Statistisches Bundesamt (Stand März 2003); eigene Berechnungen

### 3 Empirische Evidenz für Deutschland

Seit der Publikation des Buches „The Wage Curve“ von Blanchflower und Oswald im Jahre 1994 wurde dieser Ansatz von vielen anderen Arbeitsmarktökonomern aufgegriffen und für verschiedene Länder empirisch überprüft.<sup>6</sup> Dieser Abschnitt beschränkt sich auf eine knappe Wieder-gabe der empirischen Ergebnisse für Deutschland, wobei insbesondere die neuen Bundesländer berücksichtigt wer-den. Gleichwohl bezieht sich die überwiegende Mehrzahl empirischer Arbeiten auf Westdeutschland. Diese Be-schränkung ist vor allem darauf zurückzuführen, dass in-folge des Anpassungsprozesses in den neuen Bundeslän-dern nicht damit zu rechnen war, eine negative Beziehung zwischen der Lohnhöhe und der Arbeitslosigkeit zu fin-den, wie sie für die entwickelten Marktwirtschaften ty-pisch zu sein scheint.

Die nach Wissen des Autors bisher ersten Arbeiten zur Lohnkurve, die auch Ostdeutschland berücksichtigen, wurden von Pannenberg und Schwarze (1996a, b) und Baltagi, Blien und Wolf (2000) veröffentlicht. Pannenberg und Schwarze (1996a, b) schätzen die Lohnkurve für Ostdeutschland auf der Grundlage der Paneldaten des

<sup>6</sup> Einen aktuellen Überblick über die empirischen Ergebnisse zur Lohnkurve gibt Blien (2004) in diesem Band.

SOEP für den Zeitraum 1992 – 1994, wobei sie zum einen eine gepoolte OLS-Schätzung mit fixen regionalen Effekten und zum anderen ein Random-Effects-Modell schätzen. Für beide Schätzverfahren wird einmal eine Standard-Lohnkurve geschätzt und zum anderen eine erweiterte Lohnkurve, die als erklärende Variable anstelle der Arbeitslosenquote die Quote der Beschäftigungssuchenden (Arbeitslose + Teilnehmer an FuU-Maßnahmen) und den „accomodation ratio“ (konditionale Wahrscheinlichkeit, arbeitslos zu werden) enthält. Für die Standard-Spezifikation müssen die Autoren für den Untersuchungszeitraum die Existenz einer Lohnkurve in Ostdeutschland verwerfen, nicht jedoch für den erweiterten Lohnkurvenansatz. Die geschätzte Elastizität im Random-Effects-Modell beträgt  $-0.1$  und ist auf dem 10% Signifikanzniveau von Null verschieden.

Mit Daten der Beschäftigtenstatistik für den Zeitraum von 1993 – 1998 schätzen Baltagi, Blien und Wolf (2000) den Zusammenhang zwischen regionaler Lohnhöhe und Arbeitslosigkeit für 114 Landkreise in Ostdeutschland. Die Verfasser berücksichtigen zeitliche und regionale Effekte und unterscheiden zwischen männlichen und weiblichen Erwerbspersonen. Sowohl für alle Arbeitnehmer sowie unterschieden nach männlich und weiblich finden die Autoren signifikante negative Elastizitäten der Arbeitslosigkeit bezüglich der Lohnhöhe.

Die neueste Arbeit zur Lohnkurve in Ostdeutschland wird von Elhorst, Blien und Wolf (2003) vorgestellt. Über den Zeitraum von 1993 bis 1999 verwenden die Autoren einen Paneldatensatz, der aus insgesamt 114 Kreisen besteht. An den herkömmlichen Schätzverfahren zur Lohnkurve kritisieren die Verfasser, dass insbesondere räumliche Beziehungen zwischen den Regionen vernachlässigt wurden, was zu einem Effizienzverlust bei der Schätzung der Lohnelastizität führen kann. Diesen räumlichen Abhängigkeitsstrukturen tragen die Autoren Rechnung, indem sie eine  $(114 \times 114)$ -Matrix in ihre Schätzungen aufnehmen, die die Wanderungsbewegungen zwischen den Regionen erfasst. Dieser Ansatz ist vergleichbar zu den Schätzungen, die Büttner (1999) für Westdeutschland spezifiziert.

Der zweite Kritikpunkt an den traditionellen Lohnkurvenschätzungen bezieht sich auf die ungeklärte Frage, was durch eine Lohnkurve verkörpert wird: handelt es sich bei dieser Beziehung um eine langfristige oder um eine kurzfristige Relation? Sind eher die kurzfristigen Effekte von Interesse, dann sollte die Schätzung für fixe Effekte zwischen den Regionen kontrollieren, sind hingegen eher die langfristigen Effekte von Belang, dann sollte entsprechend für die zeitlichen Effekte kontrolliert werden, so die Interpretation von Elhorst et al. Schließlich, in Ländern mit hoher Arbeitslosenquote ( $> 16\%$ ) kann die Situation eintreten, dass die geschätzte Elastizität des Lohnes bezüglich der Arbeitslosigkeit zusätzlich den so genannten „downward effect“ der nationalen Arbeitslosenquote erfasst.

Für den Untersuchungszeitraum erhalten Elhorst et al. die folgenden Ergebnisse: Betrachtet man die Lohnkurve als eine langfristige Beziehung, dann beträgt die Elastizität

der Arbeitslosigkeit bezüglich des Lohnes zwischen  $-0.064$  und  $-0.042$ , abhängig davon, ob in den Koeffizientenschätzungen der „downward effect“ enthalten ist (erster Wert) oder nicht. Analog erhalten Elhorst et al. Elastizitäten zwischen  $-0.112$  und  $-0.008$ , wenn die Lohnkurve als eine kurzfristige Beziehung interpretiert wird.<sup>7</sup> In all diesen Fällen wurde der räumlichen Abhängigkeit zwischen den Kreisen Rechnung getragen.

#### 4 Modelltheoretische Erklärungen der Lohnkurve und Spezifikation der Gleichungen

In ihrem Buch sowie in verschiedenen Zeitschriftenbeiträgen zeigen Blanchflower und Oswald, dass die „empirische“ Lohnkurve mit verschiedenen Theorien der Lohnbildung kompatibel ist<sup>8</sup>. Zu nennen sind hier insbesondere der so genannte Verhandlungs- (wage bargaining approach) sowie der Effizienzlohnansatz<sup>9</sup> in seinen unterschiedlichen Varianten<sup>10</sup>. Eine Kombination beider Ansätze stellen Beck, Büttner und Fitzenberger (2001) vor. Demgegenüber werden „labor contract models“ sowie Ungleichgewichtsansätze deutlich seltener zur Begründung einer Lohnkurve herangezogen.

Im Effizienzlohnansatz wird davon ausgegangen, dass es in Zeiten einer Rezession den Unternehmen möglich ist, den Beschäftigten deutlich niedrigere Löhne zu zahlen als in Zeiten einer Vollbeschäftigung, da Arbeitslosigkeit von den Arbeitnehmern als eine Bedrohung angesehen wird. Dieser potenziellen Bedrohung kann jedoch nicht durch eine unmittelbare Abwanderung in eine andere Region begegnet werden, da mit einer Migration erhebliche Kosten verbunden sind. Ein Arbeitnehmer kann arbeitslos werden, wenn er versucht, seine Arbeit nur teilweise auszuüben (Drückeberger, shirking) und hierbei auffällt. Dieser Ansatz wird häufig zur theoretischen Begründung einer Lohnkurve in Westdeutschland verwendet, wobei dann unterstellt wird, dass innerhalb einer Periode keine Wanderungen zwischen den Regionen stattfinden. Für die neuen Bundesländer scheint diese Annahme jedoch etwas zu restriktiv, da es insbesondere zwischen 2000 und 2002 zu erheblichen Abwanderungen in die alten Bundesländer kam. Aus diesem Grunde scheint der Verhandlungsansatz (bargaining) zur Erklärung einer Lohnkurve in Ostdeutschland eher geeignet zu sein.

<sup>7</sup> Die von Elhorst et al. benutzte Unterscheidung zwischen kurz- und langfristigen Effekten ist wenig glücklich gewählt, da dann normalerweise zusätzlich eine dynamische Spezifikation erforderlich ist, die den Zusammenhang zwischen der kurz- und langfristigen Entwicklung herstellt. Eine Möglichkeit hierzu bieten Kointegrationsstrukturen in Paneldaten. Dieser Ansatz wird von den Autoren jedoch nicht gewählt. Zu diesem Punkt merkt einer der Gutachter an: „Gemeint ist eventuell, dass bei der Random-Effects-Schätzung auch Querschnittsinformation bei der Schätzung genutzt wird und diese die „langfristigen“ Zusammenhänge eher widerspiegelt als die bei der Fixed-Effects-Schätzung ausschließlich genutzte Variation über die Zeit. Diese Interpretation setzt aber natürlich voraus, dass die Querschnittsdaten eine „Gleichgewichtssituation“ widerspiegeln.“

<sup>8</sup> Blanchflower/Oswald (1990, 1994b, 1999), Muysken/van Veen (1996).

<sup>9</sup> Für eine empirische Überprüfung der Implikationen des Effizienzlohnansatzes für Westdeutschland siehe u.a. Blanchflower/Oswald (1996).

<sup>10</sup> Zu nennen sind hier der Shirking-Approach, der Turnover-Approach, der Selection-Approach und der Sociological Approach. Für eine Übersicht über diese Ansätze siehe u.a. Blien (1986).

Im Verhandlungsansatz<sup>11</sup> wird von der Annahme ausgegangen, dass die „outsider“, d.h. die Arbeitslosen, die Verhandlungsposition der Beschäftigten schwächen und hierdurch den Anteil am Gewinn, den die Beschäftigten sich aneignen wollen, verringern („rent-sharing“). Folglich sollte auch hier in Regionen mit hoher Arbeitslosigkeit ein niedrigeres Lohnniveau vorherrschen.

Da jedoch für alle theoretischen Erklärungsansätze bezüglich der Lohnentwicklung bzw. -höhe eine Lohnkurve hergeleitet werden kann, eignet sich die Lohnkurve nicht direkt zur Diskriminierung hinsichtlich der Frage, welcher Ansatz am ehesten mit den vorhandenen Datensätzen kompatibel ist. Für die neuen Bundesländer scheint der Verhandlungsansatz jedoch eher zuzutreffen als die Erklärung über Effizienzlöhne.

Die grundlegende Spezifikation jeder Lohnkurve erklärt das Niveau der Nominallöhne (in Logarithmen,  $\log(w)$ ) in einer Region  $j$  durch die in dieser Region vorherrschende Arbeitslosenquote,  $\log(u_j)$ , sowie weiterer erklärender Variablen, die im Vektor  $X$  zusammengefasst sind und – je nach Ansatz – personen- oder firmencharakteristische Spezifika erfassen. Bezeichnet  $t$  die Zeit,  $j$  die Region und  $i$  den  $i$ -ten Betrieb, dann lautet die Grundspezifikation

$$(1) \log(w_{ijt}) = f(u_{jt}, X_{it}),$$

wenn zunächst von einem Störterm abgesehen wird. Da die Lohnkurve im  $(w, u)$ -Raum konvex und nichtlinear verläuft, werden üblicherweise für die Wahl der Arbeitslosenquote verschiedene Formulierungen überprüft wie z.B.

$f(u_{jt}, X_{it})$  oder  $f(u_{jt}, u_{jt}^2, X_{it})$  oder  $f(\log(u_{jt}), X_{it})$  oder  $f((1/u_{jt}), X_{it})$  oder  $f(\log(u_{jt}), (\log(u_{jt}))^3, X_{it})$ .

Die Ergebnisse zeigen jedoch, dass eine spezielle Form für die Arbeitslosenquote  $u$  nicht die grundlegenden Ergebnisse ändert. Im Anhang sind einige Schätzungen unter Berücksichtigung regionaler Effekte dargestellt.

Als Lohnvariable dient die Bruttolohn- und -gehaltsumme je Beschäftigten in einem Betrieb. Als Elemente des Vektors  $X$  wurden aus dem IAB-Betriebspanel Ost die folgenden Variablen ausgewählt: die Betriebsgröße in Logarithmen,  $grln$ , gemessen als Anzahl der Beschäftigten je Betrieb, eine (0, 1)-Variable,  $brat$ , die angibt, ob in dem Betrieb ein Betriebsrat besteht; die Variable nimmt den Wert 1 an, wenn ein Betriebsrat besteht und ist ansonsten Null. Eine weitere (0,1)-Variable,  $tarif$ , informiert darüber, ob ein Betrieb tarifgebunden ist oder sich an den bestehenden Tarifverträgen orientiert, ohne ausdrücklich tarifgebunden zu sein. Eine eins wird zugewiesen, wenn eine tarifvertragliche Bindung besteht oder sich der Betrieb an Tarifverträgen orientiert. Da in einigen Branchen des Produzierenden Gewerbes der Frauenanteil an den Beschäftigten erheblich über dem Durchschnitt liegt, z.B. im Nahrungs- und Genussmittelgewerbe, wurde der Anteil der Frauen an den Beschäftigten,  $frau$ , ebenfalls als erklärende Variable berücksichtigt. Schließlich wurde der Anteil der Teilzeitbeschäftigten an den gesamten Beschäftigten eines Betriebes erfasst,  $tz$ .

Eine weitere Variable, die aber nur für den Zeitraum von 1998 bis 2000 zur Verfügung steht, erfasst die Quote der Langzeitarbeitslosen in den Regionen. Diese Daten wurden von der BBR Bonn zur Verfügung gestellt.

Ist ein Betrieb tarifgebunden oder orientiert er sich an den bestehenden Tarifverträgen, dann sollte hierdurch ein positiver Effekt auf den Nominallohn ausgehen, d.h., die Löhne und Gehälter in diesen Betrieben dürften oberhalb der Löhne und Gehälter liegen, die in Betrieben ohne tarifliche Bindung gezahlt werden. In die gleiche Richtung sollte auch die Existenz eines Betriebsrates in einem Betrieb gehen. Von Betriebsräten ist u.a. zu erwarten, dass sie die Interessen der Belegschaft vertreten und versuchen, ein bestimmtes Lohnniveau zu halten, sofern es mit dem Ziel der Beschäftigung vertretbar erscheint. Folgt man diesen Überlegungen, dann sollten die beiden Koeffizienten positive Vorzeichen aufweisen.

Frauen werden häufig niedriger als ihre männlichen Kollegen entlohnt. Daher ist zu erwarten, dass Betriebe mit einem hohen Anteil an weiblichen Beschäftigten eine niedrigere Lohn- und Gehaltsumme aufweisen als Betriebe, in denen überwiegend Männer beschäftigt sind. Da dieser Effekt aber nicht der Lohnkurve zugerechnet werden kann, ist hierfür gesondert zu kontrollieren. Für den zu schätzenden Koeffizienten wird deshalb ein negatives Vorzeichen erwartet.

Für Teilzeitbeschäftigte kann ebenfalls erwartet werden, dass sie eine geringere Entlohnung als Vollzeitbeschäftigte erhalten, insbesondere, wenn sie gering qualifizierte Tätigkeiten ausüben. Zusätzlich ist zu berücksichtigen, dass die Bruttolohn- und -gehaltsumme durch die Anzahl der Beschäftigten dividiert wurde, ohne dass zuvor die Teilzeitbeschäftigten in Vollzeitäquivalente umgerechnet wurden. Deshalb ist zu erwarten, dass Betriebe, in denen viele Teilzeitbeschäftigte tätig sind, eine zu niedrige Lohn- und Gehaltsumme pro Kopf ausweisen. Um diesen Effekt in der Regression zu erfassen, wurde ebenfalls die Quote der Teilzeitbeschäftigten als Regressor aufgenommen.

Zusätzlich zu diesen Variablen wurden vier Blöcke von Dummy-Variablen aufgenommen, die für zeitliche Effekte kontrollieren, der Branchenzugehörigkeit der Betriebe Rechnung tragen, die Betriebsgröße über Beschäftigtenklassen erfassen (als Alternative zur Anzahl der Beschäftigten) und schließlich für die regionale Ansiedlung der Betriebe kontrollieren. Details über diese Variablen sind im Anhang der Arbeit dargestellt.

## 5 Die Daten

In dieser Arbeit werden die Daten des IAB-Betriebspanels verwendet, also von einer Unternehmenssicht aus das Problem analysiert. Bisherige Studien zur Überprüfung einer möglichen Lohnkurve in den neuen Bundesländern verwendeten entweder die Datensätze des Sozio-

<sup>11</sup> Siehe z. B. Blanchflower/Oswald (1994b), Kapitel 3, 83 – 91.

ökonomischen Panels (SOEP) oder der IAB-Beschäftigtenstatistik auf Kreisebene. Daten des IAB-Betriebspanels wurden meines Wissens nach bisher nur für Westdeutschland verwendet, sodass hier mit einem „neuen“ Datensatz, der sich zudem auf die betriebliche Sicht konzentriert, überprüft werden soll, ob sich empirisch eine Lohnkurve in den neuen Bundesländern ermitteln lässt. Im Gegensatz zu den Daten des SOEP bzw. der Beschäftigtenstatistik können jedoch keine individuellen Charakteristika wie Berufserfahrung, Qualifikation etc. berücksichtigt werden. Stattdessen können aber firmenspezifische Informationen in die Modellspezifikation eingehen wie Tarifbindung, übertarifliche Entlohnung, Anteil der Frauen an den Beschäftigten etc. Ein Nachteil des IAB-Betriebspanels für Ostdeutschland besteht darin, dass von den befragten Betrieben nur die Bruttolohn- und -gehaltssumme mitgeteilt wird. Diese Informationsverdichtung verhindert, dass direkt oder indirekt Stundenlöhne berechnet werden können, wie dies im Allgemeinen mit individuellen Datensätzen möglich ist. Stattdessen konnte nur das monatliche Bruttoeinkommen je Beschäftigten ermittelt werden.

Eine weitere Einschränkung bezüglich der Lohnvariablen betrifft die Unterscheidung in Vollzeit- und Teilzeitarbeit sowie die zugehörigen Lohn- und Gehaltssummen. Da auch hier keine ausreichenden Informationen zur Verfügung stehen, wird versucht, über den Anteil der Teilzeitbeschäftigten in einem Betrieb diesem Problem zu begegnen. Da jedoch der Anteil der Teilzeitarbeit in den neuen Bundesländern deutlich niedriger ausfällt als in Westdeutschland, ist zu vermuten, dass dieser Effekt auf den berechneten Lohn von geringerer Bedeutung sein wird. Für diese Annahme spricht auch, dass sich die Untersuchung hier auf das Produzierende Gewerbe beschränkt, in dem der Anteil der Teilzeitarbeitenden traditionell geringer ausfällt als in den Bereichen Dienstleistungen, Handel und Staat.

Für die Beschränkung auf Betriebe des ostdeutschen Produzierenden Gewerbes lassen sich – mit Einschränkungen – die folgenden Gründe nennen: Nach wie vor leistet das Produzierende Gewerbe in Ostdeutschland den höchsten Beitrag zur Bruttowertschöpfung. Insbesondere die Bereiche Handel und Dienstleistungen sind im Vergleich zu Westdeutschland immer noch unterproportional entwickelt. Ein weiterer Grund für die Konzentration auf die Betriebe des Produzierenden Gewerbes ist darin zu sehen, dass sich besonders die Industrie in den letzten Jahren in den neuen Bundesländern sehr dynamisch entwickelt hat – sowohl in einzelnen Branchen als auch in einzelnen Regionen – und somit am ehesten von einer Angleichung an das westdeutsche Niveau ausgegangen werden kann. Allerdings sollte eine weiterführende Untersuchung zukünftig auch die anderen Sektoren der ostdeutschen Wirtschaft in die Analyse einbeziehen.

Aus Gründen des Datenschutzes kann aus dem IAB-Betriebspanel Ost nicht ermittelt werden, in welchem Kreis ein Betrieb angesiedelt ist. Um gleichwohl einen regionalen Bezug herstellen zu können, werden anstelle von Kreisen siedlungsstrukturelle Kreistypen gebildet, die sich nach dem Verdichtungsgrad einer Region klassifizieren lassen, beginnend mit „Agglomeration Kernstadt“ bis hin zu „Ländlicher Raum: Ländlicher Kreis geringerer Dichte“. Insgesamt wurden neun unterschiedliche siedlungsstrukturelle Kreistypen gebildet, wie sie von der BBR zur Verfügung gestellt werden. Diesen neun siedlungsstrukturellen Kreistypen wurden nun die 112 Kreise Ostdeutschlands (ohne Berlin) zugeordnet. Diese Art der Zuordnung impliziert, dass Kreise mit vergleichbaren Charakteristika zu einem siedlungsstrukturellen Kreistyp zusammengefasst werden, obwohl sie räumlich sehr weit voneinander entfernt liegen können. Aufgrund dieser Informationsverdichtung ist es nicht möglich, räumliche Abhängigkeiten im Sinne unmittelbarer Nachbarschaftseffekte zu erfassen, wie es beispielsweise in der Arbeit von Elhorst et al. (2003) geschieht.

ren lassen, beginnend mit „Agglomeration Kernstadt“ bis hin zu „Ländlicher Raum: Ländlicher Kreis geringerer Dichte“. Insgesamt wurden neun unterschiedliche siedlungsstrukturelle Kreistypen gebildet, wie sie von der BBR zur Verfügung gestellt werden. Diesen neun siedlungsstrukturellen Kreistypen wurden nun die 112 Kreise Ostdeutschlands (ohne Berlin) zugeordnet. Diese Art der Zuordnung impliziert, dass Kreise mit vergleichbaren Charakteristika zu einem siedlungsstrukturellen Kreistyp zusammengefasst werden, obwohl sie räumlich sehr weit voneinander entfernt liegen können. Aufgrund dieser Informationsverdichtung ist es nicht möglich, räumliche Abhängigkeiten im Sinne unmittelbarer Nachbarschaftseffekte zu erfassen, wie es beispielsweise in der Arbeit von Elhorst et al. (2003) geschieht.

**Tabelle 3: Anteile der Betriebe in den siedlungsstrukturellen Kreistypen**

Siedlungsstrukturelle Kreistypen	PANEL98 (1998 – 2002)
Agglomeration Kernstadt (RT1)	10,5
Agglomeration Hochverdichteter Kreis (RT2)	1,9
Agglomeration Verdichteter Kreis (RT3)	6,2
Agglomeration Ländlicher Kreis (RT4)	8,4
Verstädterte Kernstadt (RT5)	8,1
Verstädterter Verdichteter Kreis (RT6)	15,4
Verstädterter Ländlicher Kreis (RT7)	20,2
Ländlicher Raum: Ländlicher Kreis höherer Dichte (RT8)	13,8
Ländlicher Raum: Ländlicher Kreis geringerer Dichte (RT9)	15,4
<b>Anzahl Unternehmen</b>	<b>3145/5 = 629 pro Welle</b>

In einem nächsten Schritt wurden dann die aus dem IAB-Betriebspanel Ost entnommenen Betriebe den siedlungsstrukturellen Kreistypen zugeordnet. Tabelle 3 zeigt für den Paneldatensatz die Anzahl/Anteil der Betriebe je Kreistyp.

**Tabelle 4: Mittlere Arbeitslosenquoten für die siedlungsstrukturellen Kreistypen**

Jahr/Regionstyp	1998	1999	2000	2001	2002
RT1	15,5	16,5	16,6	17,2	17,4
RT2	19,8	18,3	17,5	17,9	18,8
RT3	20,1	18,6	18,1	18,9	18,2
RT4	17,2	17,2	17,2	17,7	18,1
RT5	19,5	19,1	18,5	18,6	19,0
RT6	20,6	19,7	19,6	20,1	20,7
RT7	20,6	19,8	19,3	19,8	20,2
RT8	18,1	17,7	17,3	17,1	17,4
RT9	20,7	19,9	20,2	21,1	22,2

Quelle: BBR

Arbeitslosenquoten liegen sowohl für die einzelnen Arbeitsamtsbezirke vor als auch auf Kreisebene. Um auch hier Arbeitslosenquoten für die einzelnen siedlungsstrukturellen Kreistypen zu erhalten, wurden die Arbeitslosenquoten der einzelnen Kreise wieder den Kreistypen zugeordnet und pro Kreistyp eine mittlere Arbeitslosenquote pro Jahr berechnet; d.h., für jedes Jahr wird jedem Kreistyp eine Arbeitslosenquote zugeordnet. Da im IAB-Betriebspanel-Ost keine Arbeitslosenquoten enthalten sind, wurden diese freundlicherweise vom BBR zur Verfügung gestellt. Tabelle 4 zeigt die mittleren Arbeitslosenquoten für die neun siedlungsstrukturellen Kreistypen.

Insgesamt zeigen die Arbeitslosenquoten für die einzelnen Regionen wenig Variation über die Zeit; beispielsweise beträgt die Spannweite der Arbeitslosenquoten für die Raumregion 4 weniger als einen Prozentpunkt über fünf Jahre. Die größte Variation mit 2,3 Prozentpunkten über die Zeit ist für die Raumregion 9 (ländlicher Kreis geringer Dichte) festzustellen. Ob die in den Arbeitslosenquoten vorhandene Varianz ausreicht, um eine statistisch gesicherte Lohnkurve für Ostdeutschland zu ermitteln, wird weiter unten diskutiert.

## 6 Ergebnisse der Schätzungen

Um Aufschluss über die mögliche Existenz einer Lohnkurve in Ostdeutschland zu erhalten, wird zunächst für jedes einzelne Jahr eine Querschnittsregression geschätzt, die neben der regionalen Arbeitslosenquote und den betriebsspezifischen Variablen Dummies enthält, die die Effekte der Branchenzugehörigkeit und der Betriebsgröße erfassen sollen. Die Ergebnisse dieser Schätzungen sind

beispielhaft für das Jahr 2002 in der Tabelle 5 dargestellt<sup>12</sup>.

Für jedes Jahr wurden insgesamt sechs verschiedene Spezifikationen für die Lohnkurve geschätzt. In der ersten Schätzung wurde nur die regionale Arbeitslosenquote als erklärende Variable verwendet, Gleichung 2 ergänzt die Schätzung um firmenspezifische Variablen wie Existenz eines Betriebsrates (brat), Anteil der Teilzeitarbeit (tz), Anteil der weiblichen Beschäftigten (frau) und ob der Betrieb tarifgebunden ist oder nicht (tarif). Gleichung 3 enthält zusätzlich die Betriebsgröße, gemessen durch die Anzahl der Beschäftigten (in logs) sowie durch den quadrierten Wert. Als Alternative zum Maß „Betriebsgröße“ wurden Größenklassen für die Anzahl der Beschäftigten gebildet. Gleichung 6 stellt die Ergebnisse dar. In den Gleichungen 4 und 5 wird zusätzlich für die Branchenzugehörigkeit des Betriebs kontrolliert.

Die Querschnittsergebnisse für das Jahr 2002 lassen nicht den Schluss zu, dass in den neuen Bundesländern eine Lohnkurve besteht. Zwar weisen alle Schätzungen das erwartete negative Vorzeichen für die Arbeitslosenquote auf, aber in keinem der Fälle lässt sich eine statistische Signifikanz auf den üblichen Testniveaus nachweisen<sup>13</sup>.

<sup>12</sup> Für die Jahre von 1998 bis 2001 können die Ergebnisse auf Anfrage zur Verfügung gestellt werden.

<sup>13</sup> Da sowohl die Lohnvariable als auch die Arbeitslosenquote in logarithmierter Form in die Schätzgleichungen eingehen, handelt es sich bei den geschätzten Koeffizienten um Elastizitäten. Präziser handelt es sich um die Elastizität des Lohnes bezüglich der regionalen Arbeitslosenquote. Im Folgenden wird diese Elastizität als „Lohnelastizität“ bezeichnet.

**Tabelle 5: OLS-Schätzungen<sup>1</sup> für das Jahr 2002; endogene Variable: log(wage)**

	Gleichung 1	Gleichung 2	Gleichung 3	Gleichung 4	Gleichung 5	Gleichung 6
Konstante	8.616*	8.130*	7.766*	7.484*	8.366*	8.598*
log(alq)	-0.225	-0.090	-0.099	-0.066	-0.106	-0.126
brat		0.434*	0.173*	0.183*	0.216*	0.218*
tarif		0.072***	0.060	0.063	0.078***	0.075***
frau		-0.001	-0.002**	-0.001	-0.001	-0.002**
tz		-0.011*	-0.010*	-0.009*	-0.010*	-0.010*
grln			0.190*	0.193*		
grln <sup>2</sup>			-0.011***	-0.011***		
Branchen 1 – 10				ja	ja	
Betriebsgröße 1 – 5					ja	ja
R2(adj.)	0.000	0.289	0.361	0.393	0.362	0.332
F-Stat.	0.94	51.96*	51.79*	24.96*	18.80*	32.27*
Anzahl Beobachtungen	629	629	629	629	629	629
F-Test Branchen				4.27*	3.85*	
F-Test Betriebsgröße					9.25*	9.17*

<sup>1</sup> Da der White-Test auf Heteroskedastizität in den Residuen hinweist, wurden die Standardfehler der Regressionskoeffizienten mit dem Heteroskedastizitätskonsistenten Ansatz von White (1980) geschätzt. Die Koeffizienten für die Dummy-Variablen wurden nicht einzeln ausgewiesen. Die entsprechenden F-Tests informieren über die Signifikanz der Koeffizienten insgesamt. Ein \* (\*\*), (\*\*\*) bedeutet, dass die Werte auf dem 1, (5), (10) Prozentniveau statistisch signifikant von Null verschieden sind. Die Anzahl der Beobachtungen gibt die Gesamtzahl sowie die Fallzahlen pro Welle an. Bei den Regionaldummies ist die Referenzgröße der siedlungsstrukturelle Kreistyp 9 (ländlicher Kreis geringer Dichte), bei den Branchen die Branche 1 ( ) und bei der Betriebsgröße Betriebe mit mehr als 500 Beschäftigten.

Abgesehen von der geschätzten Elastizität für die erste Gleichung aus Tabelle 5 liegen die numerischen Werte zwischen  $-0.066$  und  $-0.126$ , also in einem Bereich, der sowohl mit vorangegangenen Schätzungen für Deutschland als auch für andere Staaten kompatibel ist. Allerdings, wie bereits erwähnt, ist keine Elastizität signifikant von Null verschieden. Führt man diese Schätzungen auch für die übrigen Jahre durch (unter Verwendung von Gleichung 5) und konzentriert sich auf die geschätzten Lohnelastizitäten, dann erhält man Werte von  $-0.474$  (1998),  $-0.304$  (1999),  $-0.558$  (2000) sowie  $-0.364$  (2001). Zwar sind auch hier alle geschätzten Elastizitäten negativ, aber die Spannweite ist im Vergleich zu den Ergebnissen nur für 2002 erheblich größer.

Da in Querschnittsregressionen zeitliche Aspekte nicht berücksichtigt werden können, wird in einem weiteren Schritt die Lohnkurve mit dem gepoolten Datensatz geschätzt. Die wesentlichen Ergebnisse sind in den Tabellen 7 und 8 dargestellt.

### 6.1 Schätzergebnisse für den gepoolten Datensatz, 1998–2002

Die Daten des IAB-Betriebspanels Ost wurden sowohl für ein balanciertes als auch für ein unbalanciertes Panel für die Jahre von 1998 bis 2002 aufbereitet. Geht man davon aus, dass in beiden Panelarten die Ausfälle zufälliger Art sind, dann sollte aufgrund der im Allgemeinen größeren Datenmenge (im Querschnitt) ein nicht balanciertes

Panel verwendet werden. Wurden allerdings die Betriebe aus dem unbalancierten Panel eliminiert, für die in einem Jahr nicht alle erklärenden Variablen verfügbar waren, dann nahm die Zahl der zur Verfügung stehenden Querschnittsdaten deutlich ab. Aus diesem Grunde wurden die Schätzungen von vornherein mit einem balancierten Datensatz durchgeführt, indem nur die Betriebe berücksichtigt wurden, für die für alle Jahre die entsprechenden Informationen vorliegen. Hierbei handelt es sich um 629 Betriebe des Produzierenden Gewerbes in Ostdeutschland (ohne Berlin), sodass das Panel insgesamt 3145 Beobachtungen enthält. Um jedoch zu überprüfen, ob die Schätzergebnisse deutlich von der Zahl der Querschnittsbeobachtungen abhängen, wurde zusätzlich mit einem zeitlich verkürzten Panel die Lohnkurve überprüft. In diesen Datensatz sind die Beobachtungen von 1999 bis 2002 enthalten. Pro Jahr konnten hierdurch 972 Betriebe berücksichtigt werden, sodass das Panel aus insgesamt 3888 Beobachtungen besteht. Der „Preis“ für diese zusätzlichen Beobachtungen ist allerdings, dass die Variation über die Zeit zwischen den siedlungsstrukturellen Kreistypen etwas verringert wird. Die Ergebnisse sind auszugswise im Anhang in Tabelle A3 dargestellt.

Zunächst erfolgt wieder ein stufenweises Vorgehen, bei dem geprüft wird, welche Blöcke an Variablen zur Erklärung des regionalen Lohnniveaus herangezogen werden können. Die erste Gleichung in Tabelle 6 zeigt die Ergebnisse ohne Dummy-Variablen. Die Arbeitslosenquote weist das erwartete Vorzeichen auf und ist auf dem 1-Pro-

**Tabelle 6: OLS-Schätzungen gepoolte Daten:<sup>1</sup> 1998–2002; endogene Variable: log(wage)**

	Gleichung 1	Gleichung 2	Gleichung 3	Gleichung 4	Gleichung 5	Gleichung 6
Konstante	8.913*	8.952*	8.903	9.337*	8.388*	7.829*
log(alq)	-0.348*	-0.366*	-0.356*	-0.366*	-0.346*	0.009
brat	0.396*	0.397*	0.374*	0.184*	0.145*	0.384*
tarif	0.044**	0.038*	0.060*	0.034**	0.026	0.040**
frau	-0.003*	-0.002*	-0.003*	-0.003*	-0.003*	-0.003*
tz	-0.009*	-0.010*	-0.009*	-0.008*	-0.008*	-0.009*
grln					0.248*	
grln <sup>2</sup>					-0.017*	
Zeitdummys 99–02		ja				
Branchen 1 – 10			ja			
Betriebsgröße 1 – 5				ja		
Regionaldummys 1 – 8						ja
R2(adj.)	0.283	0.284	0.307	0.348	0.391	0.302
F-Stat.	249.68*	139.76*	93.98*	168.51*	288.88*	105.47*
Anzahl Beobachtungen	3145 / 629	3145 / 629	3145 / 629	3145 / 629	3145 / 629	3145 / 629
F-Test Zeit		1.97***				
F-Test Branchen			11.83*			
F-Test Betriebsgröße				62.98*	277.09*	
F-Test Region						11.26*

1) Da der White-Test auf Heteroskedastizität in den Residuen hinweist, wurden die Standardfehler der Regressionskoeffizienten mit dem Heteroskedastizitätskonsistenten Ansatz von White (1980) geschätzt. Die Koeffizienten für die Dummy-Variablen wurden nicht einzeln ausgewiesen. Die entsprechenden F-Tests informieren über die Signifikanz der Koeffizienten insgesamt. Ein \* (\*\*), (\*\*\*) bedeutet, dass die Werte auf dem 1, (5), (10) Prozentniveau statistisch signifikant von Null verschieden sind. Die Anzahl der Beobachtungen gibt die Gesamtzahl sowie die Fallzahlen pro Welle an. Bei den Regionaldummys ist die Referenzgröße der siedlungsstrukturelle Kreistyp 9 (ländlicher Kreis geringer Dichte), bei den Jahresdummys das Jahr 1998, bei den Branchen die Branche 7 (Recycling) und bei der Betriebsgröße Betriebe mit mehr als 500 Beschäftigte.



zent-Niveau statistisch von Null verschieden. Mit einem Wert von  $-0.348$  jedoch fällt die Schätzung deutlich höher als in vergleichbaren Studien aus. Gleichwohl erweist sich diese Koeffizientenschätzung als ziemlich stabil, mit Ausnahme für die sechste Gleichung. Die ökonomischen Variablen „Betriebsrat“, „Tarifbindung“, „Anteil Frauenbeschäftigung“ und „Teilzeit“ sind zumindest auf dem 5-Prozent-Niveau von Null verschieden und weisen die erwarteten Vorzeichen auf. Sowohl die Signifikanz dieser Variablen als auch ihre numerischen Werte bleiben erhalten, wenn alternative Spezifikationen für den Lohn gewählt werden.

Zusätzlich zu den beiden Blöcken von Dummy-Variablen für die Betriebsgröße und die Branche werden zwei weitere Gruppen von Dummy-Variablen eingeführt: ein Block soll zeitliche Effekte erfassen und ein weiterer Block soll zusätzliche regionale Effekte abbilden, die sich nicht in der regionalen Arbeitslosenquote widerspiegeln. F-Tests auf Signifikanz der einzelnen Blöcke von Dummies zeigen, dass jeder einzelne Block zur Erklärung des Lohnniveaus beiträgt. Allerdings ist der Beitrag der Zeitdummies deutlich geringer als der für die Branchen und die Betriebsgröße. Während die beiden letztgenannten Dummy-Variablen jeweils auf dem 1-Prozent-Niveau signifikant von Null verschieden sind, sind die Zeitdummies nur auf dem 10-Prozent-Niveau gesichert. Wird alternativ zu den Größenklassen die tatsächliche Beschäftigung in den einzelnen Betrieben gewählt, dann ist der Erklärungsgehalt der Gleichung hier mit einem bereinigten Bestimmtheitsmaß ( $R^2(\text{adj})$ ) von 0.391 am höchsten. Ein F-Test auf Signifikanz der beiden Betriebsgrößenvariablen ( $\text{grln}$  und der quadrierte Wert) zeigt, dass diese Variablen erwartungsgemäß besser geeignet sind, die Betriebsgröße zu erfassen als über eine Klasseneinteilung.

Wird die gepoolte Schätzung um die acht Regionaldummies erweitert, dann ist für die Arbeitslosenquote kein statistisch signifikanter Einfluss mehr feststellbar. Allerdings kann nicht a priori ausgeschlossen werden, dass sowohl die regionalen Arbeitslosenquoten als auch die regionalen Dummy-Variablen den gleichen Einfluss auf das Lohnniveau messen und zu erheblichen Multikollinearitätsproblemen führen, wenn beide in einer Gleichung als erklärende Variablen auftreten. Um diesen möglichen Schätzproblemen aus dem Wege zu gehen, wurde deshalb zunächst eine zusätzliche Regression mit dem gepoolten Datensatz geschätzt, in der die logarithmierte Arbeitslosenquote auf die Regionaldummies regressiert wurde. Die Schätzung weist ein bereinigtes Bestimmtheitsmaß von 0.853 aus und alle regionalen Dummyvariablen sind hoch signifikant. Dieses Ergebnis weist zwar nicht auf eine perfekte Multikollinearität hin, lässt aber doch vermuten, dass beide Größen – Arbeitslosenquote und Regionaldummy – für praktisch den gleichen Einfluss auf das Lohnniveau stehen.

Da für die siedlungsstrukturellen Kreistypen keine weiteren Daten zur Verfügung stehen, die unmittelbar über die ökonomische Situation Auskunft geben, werden die nachfolgenden Schätzungen ohne entsprechende Regionaldummy-Variablen durchgeführt<sup>14</sup>.

**Tabelle 7: OLS-Schätzungen gepoolte Daten:<sup>1</sup>  
1998–2002; endogene Variable:  $\log(\text{wage})$**

	Gleichung 1	Gleichung 2	Gleichung 3	Gleichung 4
Konstante	8.844*	9.129*	9.079*	8.185*
$\log(\text{alq})$	-0.366*	-0.351*	-0.329*	-0.330*
brat	0.372*	0.167*	0.166*	0.142*
tarif	0.052*	0.044**	0.051*	0.032***
frau	-0.003*	-0.003*	-0.003*	-0.003*
tz	-0.009*	-0.008*	-0.001*	-0.008*
grln				0.254*
$\text{grln}^2$				-0.019*
Zeit 99–02	ja	ja		ja
Branchen 1 – 10	ja	ja	ja	ja
Betriebsgröße 1 – 5		ja	ja	
$R^2(\text{adj.})$	0.318	0.385	0.383	0.428
F-Stat.	76.52*	82.84*	98.72*	113.17*
Anzahl Beobachtungen	3145 / 629	3145 / 629	3145 / 629	3145 / 629
F-Test Zeit		2.26***		4.61*
F-Test Branchen	14.28*	16.33*	16.54*	17.08*
F-Test Betriebsgröße		60.56*	59.80*	

<sup>1</sup> Da der White-Test auf Heteroskedastizität in den Residuen hinweist, wurden die Standardfehler der Regressionskoeffizienten mit dem Heteroskedastizitäts-konsistenten Ansatz von White (1980) geschätzt. Die Koeffizienten für die Dummy-Variablen wurden nicht einzeln ausgewiesen. Die entsprechenden F-Tests informieren über die Signifikanz der Koeffizienten insgesamt. Ein \* (\*\*), (\*\*\*) bedeutet, dass die Werte auf dem 1, (5), (10) Prozentniveau statistisch signifikant von Null verschieden sind. Die Anzahl der Beobachtungen gibt die Gesamtzahl sowie die Fallzahlen pro Welle an. Bei den Regionaldummies ist die Referenzgröße der siedlungsstrukturelle Kreistyp 9 (ländlicher Kreis geringer Dichte), bei den Jahresdummies das Jahr 1998, bei den Branchen die Branche 7 (Recycling) und bei der Betriebsgröße Betriebe mit mehr als 500 Beschäftigte.

<sup>14</sup> OLS-Schätzungen einer Lohnkurve mit einem gepoolten Datensatz sind mit einem grundsätzlichen methodischen Problem behaftet, dem so genannten „Moulton-Problem“ (Moulton 1986, 1990). Dieses kommt dadurch zustande, dass in den Schätzungen eine Mischung von Mikro- und Aggregatdaten auftritt. Die zu erklärende Variable, der Lohn, ist auf mikroökonomischer Ebene gemessen, die zentrale erklärende Variable aber, die Arbeitslosenquote, ist eine Aggregatvariable. Durch diese Vermengung kann es geschehen, dass unbeobachtete Effekte auf der Mikroebene Zusammenhänge auf der Aggregatebene überdecken können. Dies betrifft in diesem Zusammenhang insbesondere die Arbeitslosenquote. Zwar sind OLS-Schätzungen weiterhin erwartungstreu, aber sie sind nicht mehr effizient. D.h., insbesondere die Standardfehler werden verzerrt geschätzt, sodass es zu falschen Schlussfolgerungen über den Einfluss der Aggregatvariablen kommen kann. Dieses Problem kann vermieden oder abgeschwächt werden, wenn die Lohnkurve in einem Mehrebenenmodell geschätzt wird (Bellmann/Blien (1996), Blien (1996) bzw. Elhorst et al. (2003)). Eine Alternative zum Mehrebenenmodell nutzen Schwarze (1996), Pannenberg/Schwarze (1996b) sowie Rendtel/Schwarze (1996) durch die Kontrolle individueller Heterogenität, indem sie Fixed-Effects- und Random-Effects-Panelmodelle verwenden bzw. ein verallgemeinertes Varianzkomponentenmodell nutzen (Rendtel/Schwarze (1996)).

**Tabelle 8: Panel-Schätzungen mit Random effects:<sup>1</sup> 1998–2002; endogene Variable: log(wage)**

	Gleichung 1	Gleichung 2	Gleichung 3	Gleichung 4	Gleichung 5	Gleichung 6
Konstante	8.497*	8.447*	8.513*	8.814*	8.900*	8.156*
log(alq)	-0.192***	-0.183***	-0.211***	-0.194***	-0.229***	-0.209***
brat		0.245*	0.247*	0.133*	0.135*	0.108*
tarif		0.035**	0.023	0.029**	0.016	0.013
frau		-0.002*	-0.002*	-0.002*	-0.002*	-0.002*
tz		-0.007*	-0.007*	-0.006*	-0.007*	-0.007*
grln						0.121*
grln <sup>2</sup>						-0.001
Zeit 99–02			ja		ja	ja
Betriebsgröße 1 – 5				ja	ja	
R2 / R2(adj.)	0.709	0.701	0.703	0.700	0.703	0.703
Anzahl Beobachtungen	3145 / 629	3145 / 629	3145 / 629	3145 / 629	3145 / 629	3145 / 629
F-Test Zeit			25.77*		4.54*	7.15*
F-Test Betriebsgröße				37.69*	19.97*	

<sup>1</sup>) Da der White-Test auf Heteroskedastizität in den Residuen hinweist, wurden die Standardfehler der Regressionskoeffizienten mit dem Heteroskedastizitätskonsistenten Ansatz von White (1980) geschätzt. Die Koeffizienten für die Dummy-Variablen wurden nicht einzeln ausgewiesen. Die entsprechenden F-Tests informieren über die Signifikanz der Koeffizienten insgesamt. Ein \* (\*\*), (\*\*\*) bedeutet, dass die Werte auf dem 1, (5), (10) Prozentniveau statistisch signifikant von Null verschieden sind. Die Anzahl der Beobachtungen gibt die Gesamtzahl sowie die Fallzahlen pro Welle an. Bei den Jahresdummys ist die Referenzgröße das Jahr 1998 und bei der Betriebsgröße Betriebe mit mehr als 500 Beschäftigten.

Tabelle 7 weist die Ergebnisse für die Lohnkurve aus, wenn nicht jeder Dummy-Block einzeln in die Gleichung eingeht.

Hinsichtlich der geschätzten Lohnelastizitäten unterscheiden sich die Ergebnisse aus den Tabellen 6 und 7 nur wenig. Auch die weiteren betrieblichen Kontrollvariablen bleiben durchgängig statistisch signifikant von Null verschieden. Einzig die Variable „Tarifbindung bzw. Orientierung an Tarifvertrag“ ist nur noch auf dem 10-Prozent-Niveau signifikant, wenn anstelle der Betriebsgrößenklassen die Anzahl der Beschäftigten und ihr Quadrat treten. Alle Dummy-Variablen tragen deutlich zur Erklärung der Varianz des Lohnniveaus bei, wobei den zeitlichen Effekten jedoch die geringste Bedeutung zukommt. Da diese Dummy-Variablen insgesamt den Erklärungsgehalt der Gleichung erhöhen, liegen auch die adjustierten Bestimmtheitsmaße in Tabelle 7 über jenen aus Tabelle 6. Den höchsten Erklärungsgehalt liefert mit einem Bestimmtheitsmaß von 0.428 Gleichung 4 in Tabelle 7.

In einem weiteren Schritt wurde überprüft, ob die Regressionsergebnisse aus Tabelle 7 durch die Berücksichtigung von Interaktionstermen zwischen der regionalen Arbeitslosenquote und der Zeit verbessert werden können. Der Grund hierfür stammt aus den Ergebnissen für die einzelnen Jahre, in denen die Lohnelastizität doch teilweise beträchtliche Schwankungen aufwies. Alle Gleichungen aus Tabelle 6 und 7, in denen die Zeitdummys als erklärende Variablen auftreten, wurden mit den entsprechenden Interaktionstermen erneut geschätzt. F-Tests auf gemeinsame Signifikanz dieser Interaktionsterme zeigten jedoch, dass in keinem Fall die Berücksichtigung dieser Terme zu einer signifikanten Verbesserung der Schätzung führte.

## 6.2 Schätzergebnisse für Random-Effects-Modelle, 1998–2002

Panel-schätzungen für die Lohnkurve unter der Annahme fixer Effekte zeigten insgesamt keine zufrieden stellenden Ergebnisse. Wurde das Panel unter der Annahme fixer Effekte geschätzt und wurden alle Beobachtungen gleich gewichtet, dann konnte für keine Spezifikation eine statistisch signifikante Schätzung erzielt werden. Erst in einer gewichteten Schätzung<sup>15</sup> konnten teilweise Ergebnisse erzielt werden, die vom Vorzeichen her zugunsten einer Lohnkurve sprechen. Allerdings war in keiner der Schätzungen die Lohnelastizität signifikant von Null verschieden, wenn mehr Variablen als nur die regionale Arbeitslosenquote berücksichtigt wurde. Die Ergebnisse sind auszugsweise im Anhang in Tabelle A2 dargestellt. Diese Ergebnisse überraschen jedoch nicht sehr, denn die Annahme, dass sich die Betriebe systematisch in ihrem Lohnverhalten unterscheiden, scheint wenig überzeugend. Günstiger für eine Lohnkurve fallen demgegenüber Schätzungen aus, die Random-Effects unterstellen. Die Ergebnisse dieser Schätzungen sind in der Tabelle 8 dargestellt. Entsprechend der bisherigen Vorgehensweise wurden zunächst wiederum Schätzungen ohne Dummys für Betriebsgröße und die Zeit durchgeführt, Gleichungen 1 und 2 in der Tabelle 8.

Wird nur die Arbeitslosenquote als Regressor aufgenommen, dann beträgt die geschätzte Lohnelastizität - 0.192 und ist auf dem 10-Prozent-Niveau signifikant von Null

<sup>15</sup> Die Schätzungen basieren dann auf einem GLS-Ansatz, wobei die geschätzten Residualvarianzen zwischen den Querschnitteinheiten als Gewichte verwendet werden.

verschieden. Dieser Effekt bleibt im Wesentlichen erhalten, wenn weitere ökonomische Variablen zur Erklärung herangezogen werden. Die Lohnelastizitäten schwanken zwischen  $-0.183$  und  $-0.229$ . Im Gegensatz zu den Schätzungen für den gepoolten Datensatz jedoch sind diese Elastizitäten durchgängig nur noch auf dem 10-Prozent-Niveau statistisch von Null verschieden. Ein weiterer deutlicher Unterschied zu den Schätzergebnissen für den gepoolten Datensatz ist, dass für den Random-Effects-Ansatz die Lohnelastizitäten deutlich niedriger (in absoluten Werten) ausfallen.

Die übrigen ökonomischen Variablen weisen die erwarteten Vorzeichen auf und sind mit Ausnahme der Variablen „tarif“ auf dem 1%-Niveau abgesichert. Wird jedoch zusätzlich für fixe Zeiteffekte und für die Betriebsgröße kontrolliert, dann verliert insbesondere die Tarifvariable deutlich an Erklärungsgehalt und ist statistisch nicht mehr von Null verschieden. Aber auch die Variable „Existenz eines Betriebsrates“ verliert an Einfluss, wenn für die Zeit und die Betriebsgröße kontrolliert wird – ihr Koeffizient sinkt von  $0.245$  (Gleichung 1) auf nur noch  $0.108$  (Gleichung 6), jedoch bleibt sie weiterhin signifikant von Null verschieden. Als sehr stabil hingegen erweisen sich die Koeffizientenschätzungen für die Variablen „Teilzeit“ und „Frauenanteil“. Wird anstelle der Betriebsgrößenklassen die Betriebsgröße durch die Anzahl der Beschäftigten gemessen, dann ändert sich hierdurch der Erklärungsgehalt der Gleichung praktisch nicht, wie an dem Bestimmtheitsmaß abgelesen werden kann. Allerdings ist die Betriebsgröße nur für die logarithmierte Beschäftigtenzahl signifikant, nicht jedoch für ihr Quadrat. Insgesamt können die Schätzungen des Random-Effects-Modells als zufrieden stellend angesehen werden, zu bedenken ist aber, dass in den Spezifikationen keine zusätzlichen Variablen enthalten sind, die der Beschreibung der ökonomischen Situation einer Region dienen. Somit ist die einzige regionale Variable die Arbeitslosenquote.

## 7 Schlussfolgerungen

Bisher liegen nur wenige empirische Ergebnisse bezüglich der Existenz einer Lohnkurve in Ostdeutschland vor. Ein Grund hierfür ist sicherlich die besondere Situation in den neuen Bundesländern, die vermuten lässt, dass Schätzungen einer Lohnkurve hier mit zusätzlichen Problemen behaftet sein werden. Dies scheinen auch die hier vorgelegten Ergebnisse zu bestätigen: es konnte nur eine schwache Evidenz zugunsten einer Lohnkurve in Ostdeutschland gefunden werden. Ob alternative Spezifikationen und besser geeignete Schätzverfahren zu eindeutigeren Ergebnissen führen, bleibt zukünftigen Untersuchungen vorbehalten. Aber es lassen sich eine Reihe von Modifikationen nennen, die dann berücksichtigt werden sollten.

Schätzungen von Lohnkurven verlangen zwingend, dass eine räumliche Struktur in den zu schätzenden Gleichungen enthalten ist. Diese Struktur sollte neben der regionalen Arbeitslosenquote durch zusätzliche Variable erfasst werden, die in geeigneter Weise den strukturellen Besonderheiten der Regionen Rechnung tragen. Das wohl

schwierigste Problem bei der Schätzung einer Lohnkurve besteht in der räumlichen Abgrenzung. Schätzungen auf Kreisebene sind zwar grundsätzlich möglich und zu bevorzugen, jedoch sind dann geeignete räumliche Abhängigkeitsmaße zu bestimmen, die in der Lage sind, mögliche Nachbarschaftseffekte adäquat zu erfassen. Hier besteht auch zukünftig noch ein beträchtlicher Forschungsbedarf. Schätzungen auf Bundeslandebene hingegen sind zu großflächig, um regionsspezifische Effekte angemessen abbilden zu können. Gegenüber der Kreis- und der Bundeslandebene stellt die Wahl von siedlungsstrukturellen Kreistypen eine mögliche Alternative dar, jedoch sollte nicht verkannt werden, dass es sich hierbei nur um eine „second best“-Lösung handeln kann. Ob regionale Abgrenzungen auf der Ebene von Arbeitsamtsbezirken eher den ökonomischen Gegebenheiten von Regionen Rechnung tragen, bleibt zu prüfen. Ein Problem, das hier dann aber möglicherweise auftritt, besteht in der Auswahl von zusätzlichen ökonomischen Variablen, die nicht der Arbeitsmarktstatistik entnommen werden können.

Ein Grund, warum mit der hier verwendeten Datenbasis nur ein schwacher Zusammenhang zwischen der Lohnhöhe und der regionalen Arbeitslosigkeit gefunden wurde, kann darin liegen, dass für einen Zeitraum nur neun verschiedene Arbeitslosenquoten zur Verfügung stehen und somit die Varianz in dieser Variablen „zu gering“ ist, um zur Erklärung des Lohnniveaus (auf betrieblicher Ebene) beizutragen. Ob eine tiefere Untergliederung der Arbeitslosenquote nach Geschlecht, Dauer und Qualifikation zu wesentlich anderen Ergebnissen führt, bleibt zu klären. Eine mögliche Variante wäre die zusätzliche Berücksichtigung der Langzeitarbeitslosigkeit, entweder als separate erklärende Variable oder zur Ermittlung der um Langzeitarbeitslosigkeit bereinigten Arbeitslosigkeit. Versuche in diese Richtung jedoch waren nicht erfolgreich, da ab 2001 ein deutlicher Bruch in der Reihe der Langzeitarbeitslosen auf Kreisebene vorliegt, sodass diese Variable insgesamt nicht verwendet werden konnte.

Blanchflower und Oswald (1990) weisen auf ein weiteres Problem im Zusammenhang mit der Lohnkurve hin. Zwar besteht zwischen dem regionalen Lohnniveau und der regionalen Arbeitslosenquote ein (empirisch) negativer Zusammenhang, aber dieser lässt sich nur dann deutlich nachweisen, wenn die Arbeitslosenquoten „niedrig“ sind. Je höher die regionalen Arbeitslosenquoten sind, desto flacher verläuft die Lohnkurve; für Großbritannien und die USA schätzen Blanchflower und Oswald diesen „Schwellenwert“ bei einer Arbeitslosenquote zwischen 9 und 15%. Diese Werte werden in den neuen Bundesländern aber bei weitem übertroffen, sodass die Schätzergebnisse sich auf Werte stützen, die rechts von diesen Schwellenwerten liegen, also insgesamt sehr wenig Variation aufweisen und fast parallel zur Abszisse verlaufen. Ökonomisch bedeutet dies, dass bei einer bereits sehr hohen Arbeitslosigkeit von einem Anstieg der Arbeitslosenquote nur noch ein geringer Druck auf das Lohnniveau ausgehen wird.

Ein weiterer Kritikpunkt kann sich auf den gewählten Zeitraum beziehen. Zusätzliche Daten aus früheren Wellen des IAB-Betriebspanels würden zu einer größeren Va-

rianz der Arbeitslosenquote in den einzelnen Regionen über die Zeit führen. Da insbesondere in den neuen Bundesländern die Arbeitslosenquoten ein hohes Maß an Persistenz aufweisen, könnte hierüber ein zusätzlicher Erklärungsbeitrag geleistet werden. Gegen diesen Einwand spricht aber, dass mit jeder weiteren Welle die Querschnittseinheiten deutlich abnehmen, sodass die Fallzahlen im Panel – sofern ein balanced panel erhalten bleiben soll – stark sinken. Zu bedenken ist auch, dass sich über die einzelnen Jahre hinweg die Fragen im Betriebspanel geändert haben, was zusätzlich die Fallzahlen reduziert.

Eine Beschränkung dieser Untersuchung auf Betriebe des Produzierenden Gewerbes in Ostdeutschland ist eine Restriktion, die zukünftig aufgehoben werden sollte, indem zusätzlich Betriebe aus dem Dienstleistungsbereich, dem Handel und dem Transportwesen berücksichtigt werden. Für eine derartige Beschränkung spricht bzw. sprach zum Zeitpunkt der Untersuchung, dass insbesondere das Produzierende Gewerbe in den letzten Jahren einen spürbaren Aufschwung in den neuen Bundesländern genommen hat, wohingegen andere Branchen noch deutlich mit dem Transformationsprozess befasst sind.

Ein weiterer Punkt, der zukünftig zu berücksichtigen wäre, besteht in der möglichen Erweiterung der Regressionsgleichung um gesamtwirtschaftliche Effekte wie beispielsweise die gesamtwirtschaftliche Arbeitslosenquote und die Arbeitsproduktivität. Dies entspräche den Untersuchungen, die Bell, Nickel und Quintini (2002) für Großbritannien durchgeführt haben, in denen davon Abstand genommen wird, dass nur regionale Größen als erklärende Variable in den Regressionen auftauchen.

## Literatur

Baltagi, Badi H./Blien, Uwe/Wolf, Katja (2000): The East German wage curve 1993 – 1998. In: *Economics Letters* 69, 25-31.

Beck, Martin/Büttner, Thiess/Fitzenberger, Bernd (2001): Integrating Wage Bargaining into an Efficiency Wage Model. The Relationship between Wages and Unemployment Revisited. Mannheim: mimeo.

Bell, Brian/Nickell, Stephen/Quintini, Glenda (2002): Wage equations, wage curves and all that. In: *Labour Economics* 9, 341-366.

Bellmann, Lutz/Blien, Uwe (2001): Wage Curve Analyses of Establishment Data from Western Germany. In: *Industrial and Labor Relations Review* 54 (4), 851-863.

Bellmann, Lutz/Blien, Uwe (1996): Die Lohnkurve in den neunziger Jahren. Der Zusammenhang zwischen regionalen Durchschnittslöhnen und regionaler Arbeitslosigkeit in einer Mehrebenenanalyse mit dem IAB-Betriebspanel. In: *MittAB*, 29/3, 467-470.

Blanchflower, David G./Oswald, Andrew J. (1999): International Wage Curves. Revised version of the paper written for the 1992 NBER Conference on Wage Structures. mimeo.

Blanchflower, David G./Oswald, Andrew J. (1996): Effizienzentlohnung und die deutsche Lohnkurve. In: *MittAB*, 29/3, 460-466.

Blanchflower, David G./Oswald, Andrew J. (1995): An Introduction to the Wage Curve. In: *The Journal of Economic Perspectives* 9/2, 153-167.

Blanchflower, David G./Oswald, Andrew J. (1994a): Estimating a Wage Curve for Britain 1973-90. In: *The Economic Journal* 104/426, 1025-1043.

Blanchflower, David G./Oswald, Andrew J. (1994b): *The Wage Curve*. Cambridge (Mass.), London: MIT Press.

Blanchflower, David G./Oswald, Andrew J. (1990): *The Wage Curve*. In: *Scandinavian Journal of Economics* 92/2, 215-235.

Blien, Uwe (1996): Die Lohnkurve in den achtziger Jahren. Eine Mehrebenenanalyse mit der IAB-Beschäftigtenstichprobe. In: *MittAB*, 29/3: 471-474.

Blien, Uwe (1995a): Die Lohnkurve von 1989. Eine Mehrebenenanalyse zum Zusammenhang von regionalen Durchschnittslöhnen und der regionalen Arbeitslosenquote. In: *MittAB*, 28/2: 155-170.

Blien, Uwe (1995b): The Impact of Unemployment on Wage Formation. Estimating Wage Curves for Western Germany with Multilevel Linear Models. In: Gerlach/Schettkat (1995).

Büttner, Thiess (2003): Unemployment Disparities and Regional Wage Flexibility: Comparing EU Member States and Accession Countries. Mannheim (ZEW): mimeo.

Büttner, Thiess (1996): Die Lohnkurve in den westdeutschen Kreisen. Eine Analyse auf der Basis regionaler Durchschnittslöhne. In: *MittAB*, 29/3, 475-478.

Büttner, Thiess (1995): Local Wages and Sectoral Wage Bargaining in Germany. Diskussionspapier No. 30 der Universität Konstanz.

Büttner, Thiess/Fitzenberger, Bernd (2001): Central Wage Bargaining and Wage Flexibility: Evidence from the Entire Wage Distribution. Mannheim (ZEW): mimeo.

Card, David (1995): The Wage Curve: A Review. In: *Journal of Economic Literature* 33/2, 785-799.

Elhorst, J. Paul/Blien, Uwe/Wolf, Katja (2003): New Evidence on the Wage Curve – A Spatial Panel Approach. Nürnberg (IAB): mimeo.

Gerlach, Knut/Schettkat, Ronald (Hrsg.) (1995): *Determinanten der Lohnbildung. Theoretische und empirische Untersuchungen*. Berlin: Edition Sigma.

Gerlach, Knut/Wagner, Joachim (1995): Regionale Lohnunterschiede und Arbeitslosigkeit in Deutschland. Ein Beitrag zur Lohnkurven-Diskussion. In: Gerlach/Schettkat (1995).

Möller, Joachim (1996): Die Lohnkurve im Rahmen eines allgemeinen regionalen Anpassungsmodells. In: *MittAB*, 29/3, 479-483.

Möller, Joachim/Bellmann, Lutz (1996): Qualifikations- und industriespezifische Lohnunterschiede in der Bundesrepublik Deutschland. Eine Untersuchung für das Verarbeitende Gewerbe auf Basis der IAB-Beschäftigtenstichprobe. In: *Ifo Studien* 42, 235-267.

Moulton, Brent R. (1990): An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. In: *Review of Economics and Statistics* LXXII (2), 334-338.

Moulton, Brent R. (1986): Random group effects and the precision of regression estimates. In: *Journal of Econometrics* 32 (3), 385-398.

Muysken, Joan/van Veen, Tom (1996): Efficiency Wages and Local Wage Bargaining. In: Scandinavian Journal of Economics 98 (1), 119-127.

Pannenberg, Markus/Schwarze, Johannes (1996a): Unemployment, Labor Market Training Programs and Regional Wages: An Extended Wage Curve Approach. In: DIW Discussion Papers 139. Berlin.

Pannenberg, Markus/Schwarze, Johannes (1996b): Regionale Löhne und staatliche Qualifizierungsmaßnahmen: Eine erweiterte Lohnkurve für Ostdeutschland. In: MittAB, 29/3, 494-497.

Rendtel, Ulrich/Schwarze, Johannes (1995): Zum Zusammenhang zwischen Lohnhöhe und Arbeitslosigkeit: Neue Befunde auf Basis semi-parametrischer Schätzungen und eines verallgemeinerten Varianz-Komponenten-Modells. In: Steiner/Bellmann (1995).

Rendtel, Ulrich/Schwarze, Johannes (1996): Schätzungen von Lohnkurven für Westdeutschland mit einem verallgemeinerten Varianz-Komponenten-Modell. In: MittAB, 29/3, 491-493.

Schmidt, Elke Maria (1992): Intersektorale Lohndifferentiale. Eine mikroökonomische Untersuchung mit Paneldaten für die Bundesrepublik Deutschland. In: Zeitschrift für Wirtschafts- und Sozialwissenschaften 112, 201-218.

Schwarze, Johannes (1996): Arbeitslosigkeit, Langzeitarbeitslosigkeit und das regionale Lohnniveau. Eine Lohnkurven-

schätzung für Westdeutschland mit regionalen Paneldaten. In: MittAB, 29/3, 487-490.

Schwarze, Johannes (1995): Neue Befunde zur Lohnkurve in Deutschland: Eine Analyse mit Paneldaten für Raumordnungsregionen 1985 und 1989. DIW Diskussionspaper 119. Berlin.

Steiner, Viktor/Bellmann, Lutz (Hrsg.)(1995): Mikroökonomie des Arbeitsmarktes: Beiträge zur Arbeitsmarkt- und Berufsforschung 192. Nürnberg.

Topel, Robert H. (1986): Local Labor Markets. In: Journal of Political Economy 94 (3), Pt. 2: S111-S143.

Wagner, Joachim (1996a): Gibt es eine Lohnkurve in Deutschland? In: IFO Studien 42/2, 181-199.

Wagner, Joachim (1996b): Zur (Nicht-)Existenz von Lohnkurven in Deutschland: Ergebnisse aus Schätzungen mit Betriebsdaten aus der niedersächsischen Industrie. In: MittAB, 29/3, 484-486.

Wagner, Joachim (1994): German Wage Curves, 1979-1990. In: Economics Letters 44, 307-311.

White, Halbert (1980): A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity. In: Econometrica 48, 817 – 838.

Winter-Ebmer, Rudolf (1996): Wage curve, unemployment duration and compensating differentials. In: Labour Economics 3, 425-434.

**Anhang 1: Weitere Regressionsergebnisse**

**Tabelle A1: OLS-Schätzungen für den gepoolten Datensatz<sup>1</sup>; abhängige Variable: log(wage)**

	Gleichung 1	Gleichung 2	Gleichung 3	Gleichung 4
Konstante	9.196*	6.390*	-4.092*	7.519*
log(alq)	-0.430* (4.274)		6.382 *** (1.930)	
alq		0.189 (1.565)		
alq <sup>2</sup>		-0.006*** (1.760)		
(log(alq)) <sup>3</sup>			-0.265 ** (2.061)	
1/alq				7.820 (4.178)
R <sup>2</sup>	0.005	0.006	0.006	0.005
Anzahl Beobachtungen	3145	3145	3145	3145

<sup>1</sup>) In Klammern: empirische t-Werte (absolut); \* (\*\*), (\*\*\*) = Koeffizienten sind signifikant auf dem 1 % (5 %), (10 %) Niveau.

**Tabelle A2: Panel-Schätzungen mit fixed effects (Querschnittsgewichtung):<sup>1</sup> 1998–2002; endogene Variable: log(wage)**

	Gleichung 1	Gleichung 2	Gleichung 3	Gleichung 4	Gleichung 5
log(alq)	-0.109**	-0.049	0.016	-0.048	0.010
brat		-0.025*	-0.018***	-0.025*	-0.018***
tarif		0.018*	0.010***	0.015*	0.008
frau		-0.000	-0.000	-0.000	-0.001
tz		-0.004*	-0.004*	-0.004*	-0.004*
Zeit 99–02			ja		ja
Betriebsgröße 1 – 5				ja	ja
R <sup>2</sup> / R <sup>2</sup> (adj.)	0.767 / 0.709	0.772 / 0.713	0.773 / 0.715	0.773 / 0.714	0.773 / 0.715
Anzahl Beobachtungen	3145/629	3145 / 629	3145 / 629	3145 / 629	3145 / 629

<sup>1</sup>) Da der White-Test auf Heteroskedastizität in den Residuen hinweist, wurden die Standardfehler der Regressionskoeffizienten mit dem Heteroskedastizitäts-konsistenten Ansatz von White (1980) geschätzt. Die Koeffizienten für die Dummy-Variablen wurden nicht einzeln ausgewiesen. Die entsprechenden F-Tests informieren über die Signifikanz der Koeffizienten insgesamt. Ein \* (\*\*), (\*\*\*) bedeutet, dass die Werte auf dem 1, (5), (10) Prozentniveau statistisch signifikant von Null verschieden sind. Die Anzahl der Beobachtungen gibt die Gesamtzahl sowie die Fallzahlen pro Welle an. Bei den Jahresdummies ist die Referenzgröße das Jahr 1998 und bei der Betriebsgröße Betriebe mit mehr als 500 Beschäftigten.

**Tabelle A3: OLS-Schätzungen gepoolte Daten:<sup>1</sup> 1999–2002; endogene Variable: log(wage)**

	Gleichung 1	Gleichung 2	Gleichung 3	Gleichung 4	Gleichung 5
Konstante	8.716*	8.521*	8.616*	9.080*	8.921*
log(alq)	-0.277*	-0.228*	-0.266*	-0.299*	-0.294*
brat		0.377*	0.377*	0.167*	0.168*
tarif		0.089*	0.089*	0.077*	0.086*
frau		-0.002*	-0.002*	-0.002*	-0.002*
tz		-0.010*	-0.010*	-0.009*	-0.009*
grln					
grln <sup>2</sup>					
Zeitdummys 99–02		ja	ja	ja	
Branchen 1 – 10					ja
Betriebsgröße 1 – 5				ja	ja
R2(adj.)	0.002	0.262	0.263	0.335	0.353
F-Stat.	7.73*	277.35*	174.20*	151.62*	93.22*
Anzahl Beobachtungen	3888/972	3888/972	3888/972	3888/972	3888/972
F-Test Zeit			1.95	2.70**	6.56*
F-Test Branchen					11.80*
F-Test Betriebsgröße				85.24**	76.96*

<sup>1</sup>) Da der White-Test auf Heteroskedastizität in den Residuen hinweist, wurden die Standardfehler der Regressionskoeffizienten mit dem Heteroskedastizitäts-konsistenten Ansatz von White (1980) geschätzt. Die Koeffizienten für die Dummy-Variablen wurden nicht einzeln ausgewiesen. Die entsprechenden F-Tests informieren über die Signifikanz der Koeffizienten insgesamt. Ein \* (\*\*), (\*\*\*) bedeutet, dass die Werte auf dem 1, (5), (10) Prozentniveau statistisch signifikant von Null verschieden sind. Die Anzahl der Beobachtungen gibt die Gesamtzahl sowie die Fallzahlen pro Welle an. Bei den Regionaldummys ist die Referenzgröße der siedlungsstrukturelle Kreistyp 9 (ländlicher Kreis geringer Dichte), bei den Jahresdummys das Jahr 1998, bei den Branchen die Branche 7 (Recycling) und bei der Betriebsgröße Betriebe mit mehr als 500 Beschäftigten.

## Anhang 2: Daten

Die Daten aus dem IAB-Betriebspanel Ostdeutschland beziehen sich ausschließlich auf das Produzierende Gewerbe. Diesem Datensatz wurden die Arbeitslosenquoten aus „Statistik regional“ zugespielt. Aus Gründen des Datenschutzes sind die Daten des IAB-Betriebspanels zwar mit einer Identifikationsnummer versehen, jedoch nicht mit einer Information, in welchem Kreis der Betrieb angesiedelt ist. Um den Datenschutz zu wahren, wurden siedlungsstrukturelle Regionstypen und siedlungsstrukturelle Kreistypen gebildet, hierfür die entsprechenden Arbeitslosenquoten ermittelt und anschließend die Betriebe des IAB-Betriebspanels zugespielt, indem jeder Betrieb einem siedlungsstrukturellen Kreistyp zugeordnet wurde. Die hier benutzten siedlungsstrukturellen Kreistypen sind:

## Siedlungsstrukturelle Regionstypen:

- 1 Agglomerationsraum
- 2 Verstädterter Raum
- 3 Ländlicher Raum

## Siedlungsstrukturelle Kreistypen:

- 1 Agglomeration Kernstadt
- 2 Agglomeration Hochverdichteter Kreis
- 3 Agglomeration Verdichteter Kreis
- 4 Agglomeration Ländlicher Kreis
- 5 Verstädterter Kernstadt
- 6 Verstädterter Verdichteter Kreis
- 7 Verstädterter Ländlicher Kreis
- 8 Ländlicher Raum: Ländlicher Kreis höherer Dichte
- 9 Ländlicher Raum: Ländlicher Kreis geringerer Dichte

Gemäß diesen neun Kreistypen wurden die insgesamt 112 Kreise und kreisfreien Städte Ostdeutschlands (ohne Berlin) zugeordnet.

## Branchendummys

Um branchenspezifische Effekte in den Schätzungen erfassen zu können, wurden elf Dummy-Variablen erzeugt, die für insgesamt elf Branchen des Produzierenden Gewerbes in Ostdeutschland stehen. Da im Untersuchungszeitraum die Branchenstruktur geändert wurde, bezieht sich die Zuordnung zu den Branchen auf das Jahr 2002. Diese Zuordnung wurde für den gesamten Untersuchungszeitraum unterstellt. Die erfassten Branchen sind:

Bezeich- nung	Branche	Anzahl im Panel / Jahr	
BR1	Nahrung und Genuss	285 / 57	(= 9.1 %)
BR2	Textil und Bekleidung	80 / 16	(= 2.5 %)
BR3	Papier, Druck und Holz	115 / 23	(= 3.7 %)
BR4	Chemie und Gummi	125 / 25	(=4.0 %)
BR5	Glas	140 / 28	(= 4.5 %)
BR6	Metall, Stahl und Maschinen	525 / 105	(=16.7 %)
BR7	Recycling	10 / 2	(= 0.03 %)
BR8	Kraftwagen und son- stiger Fahrzeugbau	115 / 23	(= 3.7 %)
BR9	Elektrotechnik und Feinmechanik	270 / 54	(= 8.6 %)
BR10	Möbel	110 / 22	(= 3.5 %)
BR11	Bau und Bauinstallationen	1370 / 274	(= 43.6 %)

In den Schätzungen ist BR7 die Referenzgröße.

### **Unternehmensgröße**

Um für die Größe eines Unternehmens kontrollieren zu können, gemessen an der Zahl der Beschäftigten, wurden sechs Unternehmensgrößenklassen gebildet:

Anzahl Beschäftigte	Bezeichnung	Anteile in den Jahren 1998 – 2002				
		1998	1999	2000	2001	2002
weniger als 20	uz20	39,3	40,7	41,5	43,4	43,6
20 bis unter 50	uz50	22,4	20,7	20,8	18,6	20,5
50 bis unter 100	uz100	12,4	13,0	11,4	11,6	10,0
100 bis unter 200	uz200	12,2	12,1	12,7	12,9	13,0
200 bis unter 500	uz500	9,9	10,0	10,0	9,7	8,9
500 und mehr	uzl	3,8	3,5	3,5	3,7	4,0

### **Daten aus dem IAB-Betriebspanel Ostdeutschland**

Variable	Bezeichnung	Variablentyp	Mittelwert 1998	Mittelwert 1999	Mittelwert 2000	Mittelwert 2001	Mittelwert 2002
Lohn Beschäftigten je (Monat)	Wage	kardinal, DM	2944,55	3017,56	3078,37	3062,65	3162,93 <sup>1</sup>
Arbeitslosenquote	Alq	Prozent	19,42	18,78	18,57	19,03	19,47
Tarifbindung	Tarif	(0,1) Dummy	54,7 <sup>2</sup>	80,4	81,9	83,1	79,0
Betriebsrat	Brat	(0,1) Dummy	31,3	31,3	29,1	29,1	30,0
Frauenanteil	Frauen	Prozent	22,4	22,5	22,7	22,4	22,6
Anteil Teilzeit	Tz	Prozent	5,04	4,97	5,14	5,30	6,40

<sup>1</sup>) Für 2002 würden die Angaben in Euro auf DM umgerechnet; Kurs 1 Euro = 1.95582 DM.

<sup>2</sup>) 1998 wurde im IAB Betriebspanel nur die Tarifbindung abgefragt, nicht jedoch auch die Orientierung am Tarifvertrag, sodass hier der Anteil deutlich niedriger ausfällt als in den folgenden Jahren.

Für den Zeitraum von 1998 bis 2002 stehen pro Jahr insgesamt 629 Beobachtungen zur Verfügung.