

Abhandlungen

Lorenz Kähler*, Uwe Engel und Franziska Ritter

Hängen die Ergebnisse der zweiten juristischen Prüfung vom Examensort ab?

Ein Vergleich von zehn Bundesländern

Do the results of the second state examination in law depend on the place where it is taken?

A comparison between the German federal states

<https://doi.org/10.1515/zfrs-2017-0007>

Zusammenfassung: Die Examensergebnisse sind für Juristen von so zentraler Bedeutung, dass sich die Frage aufdrängt, ob und wie stark die Prüfungen in den einzelnen Bundesländern miteinander vergleichbar sind. Das ist bisher nicht systematisch untersucht worden. Statistische Anhaltspunkte für diesen Vergleich ergeben sich aber durch einen Vergleich derjenigen, die das Bundesland nach der ersten juristischen Prüfung wechseln, mit denjenigen, die dort beide Prüfungen ablegen. Der folgende Beitrag stellt die Ergebnisse einer empirischen Studie vor, nach der sich ein Wechsel des Bundeslandes auf die zu erwartende Note auswirken kann.

***Korrespondenzautor: Prof. Dr. Lorenz Kähler**, Fachbereich Rechtswissenschaft, Universität Bremen, Universitätsallee GW 1, 28359 Bremen, Telefon: +49 (0)421 218 66069, E-Mail: lkaehler@uni-bremen.de

Prof. Dr. Uwe Engel, Sozialwissenschaftliches Methodenzentrum, Universität Bremen, Mary-Somerville-Straße 9, 28359 Bremen, Telefon: +49 (0)421 218 67320, E-Mail: engel@methodenzentrum-bremen.de

Dr. Franziska Ritter, Fachbereich Rechtswissenschaft, Universität Bremen, Universitätsallee GW 1, 28359 Bremen, Telefon: +49 (0)421 218 66014, E-Mail: sekretariat.kaehler@uni-bremen.de

Abstract: Given the crucial importance for lawyers of the results in their state law examinations it needs to be asked whether, and, if so, to what extent, the examinations taken in each of the federal states (*Länder*) in Germany are comparable. Previous research has not examined this question systematically. However, a comparison of students who move from one state to another after the first state examination with students who remain in the same state provides statistical indicia for such a comparison. The paper presents the results of an empirical study, which shows that moving from one state to another can have consequences for a candidate's likely examination result.

Keywords: State law examination, comparison between federal states (*Länder*), persistence of exam grades across the first and second state law examinations

Seit 1980 schreibt das Deutsche Richtergesetz bundesweit einheitliche Bewertungen der juristischen Prüfungen vor (§ 5d Abs. 1 S. 2 DRiG).¹ Zuvor waren insbesondere bei der Notengebung der zweiten Staatsprüfung erhebliche Unterschiede aufgetreten,² die bei den Bewerbungen der Absolventen zu nicht nachvollziehbaren Ab- und Zuschlägen geführt hatten (BT-Drucks. 7/3604: 21; BT-Drucks. 8/3972: 6.). In der damals an einigen Universitäten noch praktizierten einstufigen Juristenausbildung war auf Noten teilweise sogar vollständig verzichtet worden (BT-Drucks. 7/3604: 9, 15). Inzwischen legt zwar die Bundesverordnung über eine Noten- und Punkteskala für die erste und zweite juristische Prüfung (BGBl I 1981, 1243) eine einheitliche Bewertungsskala von 0 bis 18 Punkten fest, die sich in sieben Noten von 0 (ungenügend) bis 14–18 (sehr gut) unterteilt, wobei man ab 4 Punkten (ausreichend) bestanden hat. Sie vermag dies jedoch nur mit äußerst dehnbaren Umschreibungen. Eine befriedigende Leistung etwa soll „in jeder Hinsicht durchschnittlichen Anforderungen“ entsprechen (§ 1 der Notenverordnung). Da sie von der Höchstnote abweicht, muss sie allerdings auch Mängel enthalten. Damit ist sie bei wörtlichem Verständnis kaum von einer ausreichenden Leistung unterscheidbar, die „trotz ihrer Mängel durchschnittlichen Anforderungen noch“ entsprechen soll. In beiden Fällen liegt die Leistung daher im Durchschnitt, obwohl sich die Noten unterscheiden. Die Festsetzung der Noten erfolgt denn auch nicht allein aufgrund dieser Umschreibungen, sondern ebenfalls aufgrund tradierter Bewertungsmuster. Diese lassen sich nur begrenzt verbalisieren.

¹ Ursprünglich § 5d Abs. 1 S. 1 DRiG (Fassung vom 16.8.1980).

² Es ging insbesondere um die Anrechnung von Ausbildungsnoten, Bericht des Rechtsausschusses BT-Drucks. 8/3972: 6.

Umso drängender ist die Frage, inwieweit die abstrakte Vorgabe der Einheitlichkeit auch tatsächlich zu einer einheitlichen Bewertung führt. Immerhin treten in der Häufigkeit der vergebenen Noten erhebliche Unterschiede auf. So schwankt die Häufigkeit der Note „gut“ in der ersten Staatsprüfung von 0,9 % in Mecklenburg-Vorpommern bis hin zu 12,1 % in Sachsen-Anhalt (Bundesamt für Justiz 2014, Anl. 1). Den 34,7 % der Berliner Absolventen mit der Note vollbefriedigend stehen im benachbarten Brandenburg nur 14,8 % Absolventen mit diesem Ergebnis gegenüber (ebd.). Auch in der zweiten Staatsprüfung kommt es zu eklatanten Unterschieden. Während 2014 in Hamburg 6,7 % der Kandidaten ein „gut“ erreichten, waren das in Rheinland-Pfalz nur 0,3 % (ebd.). Die Durchfallquote der zweiten Staatsprüfung schwankt ebenfalls von 5,7 % in Mecklenburg-Vorpommern bis hin zu 26,6 % in Bremen (ebd.).

Diese Unterschiede allein lassen allerdings keinen sicheren Rückschluss auf unterschiedlich strenge Bewertungsmaßstäbe und die damit einhergehenden verschiedenen Schwierigkeitsgrade zu. Nicht auszuschließen ist, dass sich unterschiedlich qualifizierte Kandidaten der Prüfung stellen und deshalb verschiedene Noten erhalten. Die Vergabe besserer Noten an einem Ort beruhte dann schlicht darauf, dass dort bessere Kandidaten zur Prüfung antreten. Plausibel ist das etwa für die Unterschiede zwischen Hamburg und Bremen in der zweiten Staatsprüfung. Die Note vollbefriedigend erhielten in Hamburg etwa dreimal so viele Kandidaten wie in Bremen (33,6 % gegenüber 11,4 %, ebd.). Da jedoch beide Stadtstaaten zusammen mit Schleswig-Holstein dieses Examen über ein Gemeinsames Prüfungsamt durchführen und dieselben Prüfer dabei tätig sind, dürften die Unterschiede eher darauf hindeuten, dass die Hamburger Referendare im Durchschnitt besser qualifiziert sind als die Bremer oder Schleswig-Holsteiner.

Sieht man von diesem Sonderfall des Gemeinsamen Prüfungsamtes ab, gibt es jedoch keine Anhaltspunkte dafür, dass die Notenunterschiede allein auf eine unterschiedliche Qualifikation der einzelnen Examenkandidatinnen und -kandidaten zurückzuführen sind. Da die Bundesländer die Aufgaben inzwischen weitgehend über einen Klausurenring austauschen,³ dürften die Unterschiede nicht an verschiedenen Aufgabenstellungen liegen. Hingegen wird vielfach vermutet, dass unterschiedlich streng bewertet wird. So hat es Böckenförde 1998 als bekannt bezeichnet, dass unterschiedliche qualitative Standards bestehen und den Ergebnissen „in Niedersachsen, Schleswig-Holstein oder Bremen ein anderes spezifisches Gewicht zugemessen wird als denen in Bayern und Baden-Württemberg“ (Böckenförde 1998: 85).

³ Nach der Darstellung des Prüfungsamtsleiters beim OLG Hamburg Michael Labe nehmen daran außer Bayern alle Prüfungsämter teil, s. Labe, Michael (2013): 9.

Derartige Unterschiede sind allerdings bisher nicht systematisch untersucht worden. In statistischen Untersuchungen der Examina standen eher die Fragen im Vordergrund, ob die Abiturnote (diesen Zusammenhang bejahen: Streng 1996: 35; Fries 2002: 48; Stephan 2002: 83 ff.; Meier 2003: 24), das Geschlecht, die Herkunft und die Zahl der Probeklausuren (zu diesen Faktoren siehe Towfigh et al. 2014: 12; Hinz & Röhl 2016: 20 ff.) sowie Vorlesungs- (Streng 1996: 37 ff.; Benke et al. 1996: 96) und Repetitorbesuche (Streng 1996: 37 ff.) die Ergebnisse der ersten juristischen Prüfung beeinflussen. Die Bedeutung des Studienorts wurde zwar untersucht, jedoch auf einen Vergleich innerhalb eines Bundeslands begrenzt (Meier 2003: 28 f. für Niedersachsen; Towfigh et al. 2014: 18 ff. für Nordrhein-Westfalen; Hinz & Röhl: 20 ff. für Baden-Württemberg). Die damit einhergehende Unsicherheit über die Vergleichbarkeit der Examina zwischen den einzelnen Bundesländern ist umso problematischer, als die Examensnoten von zentraler Bedeutung für die berufliche Perspektive sind.

Es wäre etwa misslich, wenn Bewerber nur deshalb kein Richteramt bekämen, weil sie trotz gleicher Qualifikation aufgrund ihres Studienorts eine schlechtere Examensnote haben. Ebenso problematisch wäre es, wenn im Bewerbungsprozess aufgrund einer zu Unrecht bestehenden geringeren Reputation Abschlüsse von einem Examen gemacht würden, das dem Examen anderer Länder in Wirklichkeit nicht nachsteht. Damit wäre die von § 5d Abs. 1 S. 2 DRiG vorgeschriebene Einheitlichkeit der Bewertung genau so beeinträchtigt wie der von Art. 33 Abs. 2 GG garantierte gleiche Zugang zu öffentlichen Ämtern.⁴ In einer Kommentierung des Richtergesetzes heißt es dazu ernüchternd, dass regionale Unterschiede zwar vermieden werden sollten, jedoch zweifelhaft sei, ob der Bund die Länder zu einem abgestimmten Verhalten verpflichten dürfe (Staats 2012: DRiG, § 5d, Rn. 7). Man scheint die Unterschiede hinzunehmen. Während man diese innerhalb eines Landes durch eine zufällige Zuteilung der Klausuren zu den Korrektoren mildern kann (so in Baden-Württemberg, Hinz & Röhl 2016: 21), scheidet ein entsprechendes Verfahren aus, solange die Korrektoren nur dort prüfen sollen, wo sie auch sonst tätig sind und die Klausuren jeweils in dem Land korrigiert werden, in dem sie auch geschrieben wurden. Da die Prüfer den Großteil ihrer Karriere meist im selben Bundesland verbringen, fehlt ein Ortswechsel, der Unterschiede in den regionalen Bewertungsmaßstäben abbauen könnte. Nur bei Hochschullehrern tritt ein derartiger Ortswechsel gelegentlich auf.

⁴ Bewerber dürfen zwar nicht allein aufgrund der Herkunft aus einem Bundesland ausgeschlossen werden, wohl aber darf der Dienstherr die Noten aus den einzelnen Bundesländern in „wertender Erkenntnis“ feststellen, BVerwGE 68: 109, und damit wohl auch aus sachlichen Gründen Unterscheidungen treffen.

All das spricht dafür, die Notenvergaben in den einzelnen Bundesländern einander gegenüberzustellen und auszuwerten. Dies schafft die Grundlage dafür, die Prüfungen zwischen den einzelnen Bundesländern besser zu koordinieren. Die im Folgenden vorgestellte Studie soll dafür erste Anhaltspunkte liefern.

Die Wechsler als Vergleichsgröße

Mangels eines einheitlichen bundesweiten Examens lässt sich die juristische Qualifikation sämtlicher Prüflinge nicht im direkten Vergleich feststellen. Jedoch gibt es eine Vielzahl von Kandidaten, die sich der ersten und der zweiten juristischen Staatsprüfung in jeweils unterschiedlichen Ländern stellen. Ihr Abschneiden ist von besonderem Interesse, da sich durch den Wechsel des Bundeslandes die juristische Qualifikation nicht verschlechtern dürfte. Schneiden sie in der zweiten Prüfung anders ab als etwa die Kommilitonen, die im Bundesland verbleiben, deutet das auf Unterschiede in der Notenvergabe der zweiten Prüfung hin. Verbessern sich die Kandidaten eines Landes in der zweiten Prüfung deutlich, wenn sie das Bundesland verlassen, während sich ihre Note bei einem Verbleib in diesem Land nicht ändert, ist dies ein Indiz dafür, dass die zweite Staatsprüfung dort schwieriger ist als in den übrigen Ländern.

Daten, deskriptive Kennzahlen und Methodik

Daten

Für die vorliegende Studie wurden sämtliche Justizprüfungsämter gebeten, für die Jahre 2009–2014 anonymisiert die Noten der einzelnen Kandidatinnen und Kandidaten in der ersten und zweiten juristischen Staatsprüfung zur Verfügung zu stellen und dabei zugleich anzugeben, wo die Prüfungen abgelegt wurden. Von zehn Bundesländern erhielten wir die erbetenen Informationen, und zwar von Baden-Württemberg (BAW), Bayern (BAY), Brandenburg (BRA), Bremen (HBR), Mecklenburg-Vorpommern (MVP), Nordrhein-Westfalen (NRW), Rheinland-Pfalz (RLP), dem Saarland (SLA), Sachsen (FSA) und Schleswig-Holstein (SWH).⁵ Die übrigen Bundesländer sagten wegen fehlenden Personals, mangelnder Verfüg-

⁵ In Klammern die in nachfolgenden Analysen verwendeten Länderkürzel.

barkeit der Daten oder datenschutzrechtlicher Bedenken ab. Zum Teil konnten nur Daten für die Jahre 2011–2014 zur Verfügung gestellt werden.⁶

Unsere Analysen sind entsprechend auf die genannten zehn „Notenmelde-länder“ ausgerichtet, da nur für diese Länder komplette Informationen, also Informationen von Wechslern *und* Nichtwechslern vorliegen. Die Noten von Assessorinnen und Assessoren aus anderen als diesen zehn Bundesländern konnten allerdings dann berücksichtigt werden, wenn sie von der ersten zur zweiten Staatsprüfung in eines der zehn Bundesländer gewechselt sind, die sich über ihre Notenmeldungen an der vorliegenden Studie beteiligt haben. Die von den zehn oben genannten Ländern gemeldeten Daten umfassen in diesem Sinne speziell auch Notenangaben von Personen aus Hamburg (HHA), Hessen (HES), Niedersachsen (NIE), Sachsen-Anhalt (SAH) und Thüringen (FTH).

Da uns für das gemeinsame Justizprüfungsamt für Berlin/Brandenburg nur Notenmeldungen aus Brandenburg vorliegen, kann die Studie über Berlin keine Erkenntnisse liefern. Zudem ist darauf aufmerksam zu machen, dass nicht alle Bundesländer Daten von nicht bestandenen Prüfungen zur Verfügung stellen konnten. Daher wurde die Analyse auf die Teilnehmerinnen und Teilnehmer begrenzt, die beide Staatsprüfungen bestanden haben.

Insgesamt liegen uns die Notenangaben von 18.422 Personen vor, die sich wie unten in Tabelle 1 ausgewiesen auf die zehn Bundesländer verteilen, die uns Noten gemeldet haben. Die Zahl reduziert sich um 1.091 auf 17.331 Personen, sofern die Analyse aus Gründen der Vergleichbarkeit der Notenverteilungen auf die Teilmenge derer zu begrenzen war, die auch die erste Prüfung in einem der zehn „Notenmeldeländer“ abgelegt haben mussten. Dies betrifft speziell die unten in Tabelle 5 berichtete bidirektionale Wechselanalyse.

Da die uns zur Verfügung stehenden Daten vielfach nur diejenigen aufführten, welche die zweite Prüfung bestanden hatten, wurde die Analyse durchgehend auf diese Personengruppe begrenzt. Im Vordergrund stand also die Frage, inwieweit die Ergebnisse derjenigen vergleichbar sind, die als Assessoren nach zwei bestandenen Examina in den Beruf eintreten. Das schien auch vor dem Hintergrund vertretbar, dass diejenigen, welche die zweite Prüfung endgültig nicht bestehen, eine kleine Gruppe bilden, 2014 bundesweit lediglich 3,8 % (336 von 8.789, s. Bundesamt für Justiz 2014, Anl. 2).

In den Daten ist nicht differenziert, wie sich die Noten der beiden Staatsprüfungen zusammensetzen. Das betrifft zunächst das Verhältnis mündlicher Prüfung zu schriftlichen Leistungen. So ist theoretisch denkbar, dass etwaige

⁶ Das betraf die Bundesländer Baden-Württemberg und Nordrhein-Westfalen. Bayern und Brandenburg konnten nur die Daten der Jahre 2010–2014 zur Verfügung stellen.

Diskrepanzen zu unterschiedlichem Anteil auf einer verschiedenen Benotung von Klausuren oder mündlichen Leistungen beruhen. Das müssten weitere Analysen untersuchen; es ist aber nicht von gleicher Relevanz wie die Endnote, weil sich die Assessorinnen und Assessoren mit dieser bewerben. Ebenfalls konnte mangels entsprechender Angaben nicht weiter danach differenziert werden, zu welchem Anteil sich die Note für die erste Prüfung aus Noten für den staatlichen Teil und für den Schwerpunkt zusammensetzt.⁷ Es ging zunächst um die Frage, ob überhaupt Unterschiede in den Ergebnissen zwischen den Bundesländern festzustellen sind, während die detailliertere Suche nach den Ursachen weiteren Studien überlassen werden muss.

Die Gesamtzahl der untersuchten Notenangaben unterscheidet sich von der Gesamtzahl der für die Ausbildungsstatistik gemeldeten Prüfungsergebnisse, da in den hier zugrunde gelegten Daten eine Person nur einmal mit ihrem jeweils besten Ergebnis gewertet wurde. Diejenigen, welche die zweite Prüfung bereits im ersten Versuch bestanden hatten, zur Notenverbesserung die Prüfung indes wiederholten, wurden also nicht doppelt gezählt. Zwar ist nicht auszuschließen, dass uns zu einzelnen Assessoren keine Daten übermittelt wurden, etwa weil die Prüfungsakten für ein gerichtliches Verfahren gebraucht wurden. Ein systematischer Fehler sollte damit allerdings nicht verbunden sein, weil die Auswahl der zu übermittelnden Notenangaben nach keinem anderen Kriterium erfolgte, als dass die Assessoren im jeweiligen Jahr das zweite Examen bestanden haben sollten.

Deskriptive Kennzahlen

Tabelle 1 informiert für die zehn Bundesländer über einige deskriptive Kennwerte der Notenverteilungen in der zweiten juristischen Staatsprüfung. Ausgewiesen werden einerseits das arithmetische Mittel \bar{x} und als Maß für die durchschnittliche Streuung der Noten um diesen Mittelwert die Standardabweichung s . Andererseits werden zur Beurteilung der zentralen Tendenz und Homogenität der jeweiligen Notenverteilung das 25 %-Quantil, das 50 %-Quantil, das 75 %-Quantil und der Quartilabstand⁸ ausgewiesen. Insgesamt bewegen sich die Noten im Spektrum von 4,0 bis 14,4 Punkten.

⁷ Der Schwerpunkt wird durchgehend besser bewertet als der staatliche Teil, siehe Bundesamt für Justiz (2014): 1, 4.

⁸ „Quantile“ zerlegen eine Verteilung in einen unteren und einen oberen Teilbereich. Quantilwerte sind hier die Noten, welche die jeweiligen Trennstellen darstellen. Die in den Spalten für das 25 %-/50 %-/75 %-Quantil stehenden Werte sind die Noten, für die gilt, dass 25 %/50 %/75 %

An \bar{x} gemessen, reicht das Spektrum von Durchschnittsnoten von 6,6 Punkten in Mecklenburg-Vorpommern bis zu 7,8 Punkten in Nordrhein-Westfalen. Man beachte dabei, dass dies nur diejenigen Kandidaten betrifft, die beide Examina bestanden haben, und die Durchschnittsnoten daher höher liegen, als wenn man in einem Examensdurchgang die Durchschnittsnote aller Teilnehmer berechnet und dabei auch die durchgefallenen Kandidaten einbezieht. Die zehn Bundesländer unterscheiden sich danach im Mittel um bis zu 1,2 Notenpunkte voneinander. Ungefähr⁹ zwei Drittel der Notenpunkte fallen dabei in den Bereich von $\bar{x} \pm 1 \cdot s$, liegen also beispielsweise in NRW zwischen 5,8 und 9,8 Punkten und in Mecklenburg-Vorpommern zwischen 4,9 und 8,3 Punkten. Die Homogenität der Notenverteilungen erweist sich dabei im Spiegel des Variationskoeffizienten s/\bar{x} weitgehend als ähnlich, wobei er für Brandenburg am niedrigsten ist.

Tabelle 1: Deskriptive Kennzahlen der bundeslandbezogenen Häufigkeitsverteilungen der Noten von Assessorinnen und Assessoren in der zweiten Staatsprüfung

2. Staatsprüfung abgelegt in	\bar{x}	s	s/\bar{x}	Quantil			Quantil- abstand	Zahl der Fälle, N	
				25 %	50 %	75 %			
Nordrhein-Westfalen	NRW	7,8	2,0	0,26	6,2	7,7	9,2	3,0	5.531
Schleswig-Holstein	SWH	7,5	1,8	0,24	6,2	7,3	8,7	2,5	489
Rheinland-Pfalz	RLP	7,5	1,7	0,23	6,2	7,3	8,6	2,4	1.633
Baden-Württemberg	BAW	7,4	1,9	0,26	5,9	7,1	8,7	2,8	2.259
Bremen	HBR	7,4	1,8	0,24	6,0	7,3	8,6	2,6	405
	Alle	7,4	1,9	0,26	5,9	7,1	8,6	2,7	18.422
Saarland	SLA	7,2	1,8	0,25	5,9	7,0	8,2	2,3	373
Brandenburg	BRA	7,2	1,6	0,20	6,0	7,0	8,1	2,1	689
Bayern	BAY	7,1	1,8	0,25	5,6	6,7	8,2	2,6	5.642
Sachsen	FSA	6,9	1,7	0,25	5,6	6,6	8,0	2,4	1.035
Mecklenburg-Vorpommern	MVP	6,6	1,7	0,26	5,4	6,3	7,6	2,2	366

der Personen Noten erzielt haben, die kleiner als oder gleich der jeweils ausgewiesenen Note sind. Als Anhaltspunkt dafür, wie homogen die Notenverteilung ausfällt, kann dann der „Quartilabstand“ (Differenz zwischen dem 75 %- und 25 %-Quantilwert) herangezogen werden.

⁹ Gilt nur ungefähr, da die Notenverteilungen leicht rechtsschief sind. In Tabelle 1 ist dies daran erkennbar, dass das 50 %-Quantil, also die Mediannote, geringer ausfällt als die Durchschnittsnote.

Anhaltspunkte für die Homogenität der Notenverteilungen liefern auch die Bereiche, in denen sich jeweils die mittleren 50 % der Noten bewegen. In Nordrhein-Westfalen wäre dies beispielsweise der Bereich zwischen 6,2 und 9,2 auf der Notenskala. Während der Quartilabstand in Nordrhein-Westfalen somit bei 3,0 Notenpunkten liegt, erweist sich am anderen Pol liegend die Notenverteilung in Brandenburg mit 2,1 Punkten als homogener.

Methodik

Für die in der zweiten Staatsprüfung erzielte Note spielt typischerweise die zuvor in der ersten Staatsprüfung erzielte Note, d. h. die darin ihren Ausdruck findenden *individuellen* Leistungsvoraussetzungen, eine entscheidende Rolle. Die Korrelation zwischen den beiden Notenpunkten beträgt signifikante $r=0,66$ und die Regression der Note der zweiten Staatsprüfung („Note2“) auf die Note der ersten Staatsprüfung¹⁰ („Note1“) $b=0,63$, wobei die Irrtumswahrscheinlichkeit weniger als ein Promille beträgt ($p < 0,001$). Über alle zehn Bundesländer hinweg betrachtet, kann Note2 also gut auf Basis von Note1 vorhergesagt werden. Denn zu erwarten wäre nach diesem Ergebnis, dass sich jeder zusätzliche Notenpunkt in der ersten Prüfung in einem Plus von durchschnittlich 0,63 Notenpunkten in der zweiten Prüfung niederschlägt.

Dieser Zusammenhang gilt für Nichtwechsler ($b=0,64$) und Wechsler ($b=0,61$) nahezu gleichermaßen und ist zu beachten, wenn es darum geht, die Äquivalenzfrage in der Notengebung über uni- und bidirektionale Wechselanalysen beantworten zu wollen. Zur Klärung dieser Frage ist folglich ein *operationales Kriterium* erforderlich, welches für die Erkennung etwaiger Abweichungen von der Benotungsäquivalenz als *benchmark* fungieren kann. *Deskriptive* Vergleiche von Durchschnittsnoten wären dafür nicht ideal, da in Bundeslandvergleichen gefundene Mittelwertdifferenzen in den Notenpunkten der zweiten Prüfung im Prinzip stets auf *zwei* Variationsursachen beruhen können. Die eine denkbare Ursache ist, dass vergleichsweise höhere oder weniger hohe Anforderungen gestellt werden bzw. strenger oder weniger streng benotet wird. Die andere mögliche Ursache besteht darin, dass in den Prüfungen vergleichsweise bessere oder weniger gute Prüfungsleistungen erbracht werden. Selbst wenn ein deskriptiver Vergleich signifikante Mittelwertdifferenzen aufdecken würde, bliebe unbeant-

¹⁰ Für diese Regression (mit Ergebnis: $b=0,63$ [$t=119,9$]) sowie die gleiche, separat für Nichtwechsler (mit Ergebnis: $b=0,64$ [$t=111,4$]) und Wechsler (mit Ergebnis: $0,61$ [$t=45,4$]) gerechnete Regression ist keine eigene Ergebnistabelle abgedruckt.

wortet, in welchem Ausmaß diese Differenzen der einen oder der anderen dieser beiden Variationsursachen geschuldet sind.

Unsere Analysen sind entsprechend auf *Effektzerlegung* in dem Sinne ausgerichtet, die Variation in den Punktzahlen der zweiten Prüfung *über multiple Regressionsanalysen* auf diese beiden Variationsursachen zurückzuführen. Gestützt auf die erzielten Examensnoten der (je nach Analyse) 18.422 bzw. 17.331 Assessorinnen und Assessoren (empirischer Input) wird entsprechend berechnet, unter welchen Bedingungen welche durchschnittliche Punktzahl in der zweiten Staatsprüfung zu erwarten ist. Solche über Regressionsmodelle geschätzten *Notenerwartungen* stellen bedingte Mittelwerte dar, wobei sich die Bedingungen hier aus der jeweiligen Bundeslandherkunft der ersten und zweiten Staatsprüfung, einem etwaigen Bundeslandwechsel und den in der ersten Prüfung erzielten Prüfungsnoten ableiten.

Ein erstes Regressionsmodell dient der Beschreibung der Ausgangslage (Modell 1). Es schätzt für die zehn Bundesländer, die Noten gemeldet hatten („BL2“), die erwarteten durchschnittlichen Punktzahlen in der zweiten juristischen Staatsprüfung (\hat{y}) über die Gleichung

$$\hat{y} = b_0^* + b_1^* BL2 \quad [1]$$

als Abweichungen von der mittleren Notenerwartung. Diese *mittlere Notenerwartung* findet aufgrund der zugrunde gelegten, sog. [1, 0, -1] „Effektkodierung“ in b_0^* ihren Ausdruck¹¹.

$$\hat{y} = b_0^* + b_1^* BL2 + b_{2|BL2}^* W \quad [2]$$

Modell 2 erweitert sodann die Gleichung um $b_{2|BL2}^*$ zur Schätzung konditionaler („BL2“-bezogener) Wechselwirkungen. Wie auch BL2 geht W dabei „effektkodiert“ in die Gleichung ein. $b_{2|BL2}^*$ liefert dabei Schätzungen des Wechselwirkungseffektes, die *nicht* berücksichtigen, dass *Note2* auch von *Note1* beeinflusst wird. In den nachfolgenden Analysen ist deshalb vom „Bruttoeffekt“ eines Bundeslandwechsels die Rede, wenn die Schätzung des Wechselwirkungseffektes auf *Note2* den Einfluss von *Note1* nicht berücksichtigt, und von einem „Nettoeffekt“ im gegenteiligen Fall. Solche Nettoeffekte schätzen wir entsprechend über ein um b_3^* erweitertes Modell:

¹¹ Neben der sog. [1, 0] „Dummy-Kodierung“ stellt diese Form der Kodierung kategorialer Variablen die zweite Standardvariante dafür dar, kategoriale Prädiktoren in eine Regressionsanalyse einzubeziehen.

$$\hat{y} = b_0^* + b_1^* BL2 + b_{2|BL2}^* W + b_{3|BL2}^* Note1 \quad [3]$$

Modell 3 liefert damit zum einen *bundeslandbezogene* Schätzungen des Einflusses von *Note1* auf *Note2* und zum anderen Wechseleffekte, die jeweils um genau diesen endogenen Noteneffekt korrigiert sind¹². Ein statistisch signifikanter Nettoeffekt eines Bundeslandwechsels würde somit anzeigen, dass bei statistischer Kontrolle des Einflusses der individuellen Leistungsvoraussetzungen, wie sie im Effekt von *Note1* auf *Note2* ihren Ausdruck finden, ein *genau auf diese individuellen Voraussetzungen nicht zurückzuführender* systematischer Wechseleffekt besteht, also quasi übrigbleibt. Es ist der Nachweis solcher Nettoeffekte eines Bundeslandwechsels auf die Notenerwartung in der zweiten Staatsprüfung, *aus denen wir operational auf Abweichungen von einer länderübergreifend äquivalenten Notengebung schließen wollen*.

Modelle 2 und 3 schätzen diese Wechseleffekte spezifisch für jedes (BL2-) Zielland, differenzieren dabei aber nicht, aus welchem Bundesland in das Zielland gewechselt wurde. Diese zusätzliche Differenzierung wird deshalb in den Modellen 4 und 5 vorgenommen. Modell 4 entspricht dabei Modell 2 und Modell 5 entspricht Modell 3, nur dass die Wechselvariable *W* nun „dummy-“ und nicht „effektkodiert“ in die Modelle eingeht¹³, und $b_{0|BL2}$ ziellandspezifisch berechnet wird. Die Wechseleffekte $b_{2|BL2}$ werden dabei als erwartete Unterschiede zwischen

¹² Die Unterscheidung von „Brutto-“ und „Nettoeffekten“ ist hier nur in diesem speziellen Sinne gemeint. Sofern zur Schätzung des Wechseleffektes b_2^* Gleichung [2] separat für jedes einzelne BL2 Land geschätzt und diese Gleichung dann in Gleichung [3] um b_3^* ergänzt wird, so liefe die Unterscheidung auf diejenige zwischen den b einer einfachen Regression und Partialregressionskoeffizienten hinaus. Würde Gl. [2] auf diese Weise separat für jedes BL2-Land geschätzt werden, entspräche die jeweilige Regressionskonstante der Summe: $b_0^* + b_1^* BL2$.

¹³ Bei der Dummy-Kodierung werden Effekte relativ zu einer Bezugskategorie geschätzt. Bei den Analysen in Tabelle 2 und 3 hätte dafür entsprechend eines der Bundesländer als Bezugsland ausgesucht werden müssen. Bei zehn Ländern hätte es für diese Wahl aber nicht nur zehn prinzipiell äquivalente Optionen gegeben, alle Effekte wären auch als erwartete Unterschiede relativ zu *diesem Bezugsland* geschätzt worden – und folglich auch nur relativ zu diesem Bezugsland interpretierbar gewesen. Das hätte die Frage aufgeworfen: Warum gerade dieses und warum nicht eines der anderen Länder als Bezugsgröße auswählen? Genau diese Abhängigkeit der Effektschätzungen von einem als Bezugsgröße ausgesuchten Bundesland vermeidet die Effektkodierung, da sie die Effekte relativ zu *einer gemeinsamen* Bezugsgröße, hier der mittleren Notenerwartung, schätzt. Den Modellen 1 bis 3 (Tabellen 2 und 3) liegt entsprechend diese Kodierung zugrunde. Bei den Wechselanalysen für jeweils *spezifische* Bundeslandkombinationen (Modelle 4 und 5 in Tabelle 4 und 5) ist hingegen die Dummy-Kodierung zielführender, da dort als Vergleich angestrebt ist, Wechsel *in ein Land* jeweils mit denjenigen zu vergleichen, die *in diesem Land* auch ihr erstes Staatsexamen abgelegt haben (Bezugskategorie).

denen, die in ein Zielland gewechselt sind, und denen, die dort jeweils auch ihre erste Prüfung abgelegt haben, geschätzt.

$$\hat{y} = b_{0|BL2} + b_{2|BL2} W \quad [4]$$

$$\hat{y} = b_{0|BL2} + b_{2|BL2} W + b_{3|BL2} Note1 \quad [5]$$

Ergebnisse

Modell 1 beschreibt die Ausgangslage (Tabelle 2). Es schätzt für jedes der zehn ausgewiesenen Bundesländer den jeweiligen Notendurchschnitt der zweiten Prüfung als Abweichung von der mittleren Notenerwartung und reiht die Bundesländer nach der Höhe dieser Abweichungswerte. Die landesdurchschnittlichen Notenpunkte rangieren dabei zwischen 6,6 Punkten in Mecklenburg-Vorpommern (7,245 – 0,629) und 7,8 Punkten in Nordrhein-Westfalen (7,245 + 0,516), wobei die Schätzwerte für die einzelnen Länder dicht bei den empirischen Werten aus Tabelle 1 liegen. Sieben der zehn bundeslandbezogenen Durchschnittspunkte weichen dabei statistisch signifikant¹⁴ von der über alle zehn Bundesländer hinweg errechneten Durchschnittspunktzahl, also der mittleren Notenerwartung, von 7,25 Punkten ab. Dies betrifft Nordrhein-Westfalen, Schleswig-Holstein, Rheinland-Pfalz und Baden-Württemberg mit Punktzahlen oberhalb dieses Bezugs werts sowie Bayern, Sachsen und Mecklenburg-Vorpommern mit Punkt-

14 Wenn wir in der Ergebnisdarstellung statistische Signifikanzen berichten, obwohl die Datengrundlage keine Zufallsauswahl, sondern eine Vollerhebung darstellt, so folgen wir damit einer in der Diskussion um statistische Inferenz vertretenen Empfehlung für die Analyse von Zensus- oder Aggregatdaten, um auf diese Weise z. B. zeitbezogene Variation in den Daten bzw. andere mit den Daten verbundene Unsicherheiten berücksichtigen zu können. Auch die vorliegende Vollerhebung liefert nur einen Ausschnitt aus dem Universum der Daten, die bei Berücksichtigung der Zeiträume davor und danach anfallen würden. Küchler (1979): 114 ff., spricht sich z. B. explizit für diese Vorgehensweise aus, ebenso Jones, Diskussionsbeitrag vom 23. Juni 2014 auf https://www.researchgate.net/post/Are_p-values_and_significance_tests_still_meaningful_in_census_data_analysis (abgerufen am 11.6.2017). Die zentrale Berichtsgröße ist dabei jeweils der als „t-Wert“ ausgewiesene Quotient „Effektschätzung, dividiert durch den zugehörigen Standardfehler“. Wir weisen diese Größe jedes Mal auch deshalb explizit aus (anstelle nur p-Werte für die Signifikanzen zu berichten), um darüber auch die jeweiligen Effektstärken besser einschätzen zu können. Damit folgen wir einer Praxis, wie sie z. B. Coleman (1981) konsequent verfolgt hat. Neben dem Aspekt „Signifikanz“ geht es uns also darum, die Bedeutsamkeit eines Effekts auch über dessen Standardisierung an einer einschlägigen Variationsquelle, hier: dem Standardfehler, einschätzen zu können.

zahlen darunter. Die Ausgangslage kann also nur so beschrieben werden, dass sich die durchschnittlichen Benotungen in sieben von zehn Bundesländern voneinander unterscheiden und folglich Grund zur Annahme bieten, dass die Notengebung ggf. nicht äquivalent erfolgt.

Allerdings bliebe bei dieser Schlussfolgerung unberücksichtigt, dass die Notendurchschnitte in den einzelnen Bundesländern aus zwei unterschiedlichen Quellen gespeist werden. Das sind erstens die Notenpunkte derjenigen, die für die zweite Staatsprüfung in das jeweilige Bundesland gewechselt sind, und zweitens die Notenpunkte derjenigen, die im jeweiligen Bundesland auch ihre erste Prüfung abgelegt haben. Die Anteile dieser Wechsler stellen sich aber von Bundesland zu Bundesland teilweise sehr unterschiedlich dar (Tabelle 2). Während lediglich 4,6 % der bayerischen Assessoren zuvor die erste Staatsprüfung in einem anderen Bundesland abgelegt haben, waren dies bei Schleswig-Holstein 54 %. Die Beiträge jedes Bundeslandes zur juristischen Ausbildung unterscheiden sich daher erheblich, je nachdem, ob man das Studium oder das Referendariat betrachtet. Demgemäß müssen sich die einzelnen Ausbilder in unterschiedlichem Maße auf Personen einstellen, die ihr Studium in einem anderen Bundesland absolviert haben und das in der zweiten Prüfung maßgebliche Landesrecht nicht kennen.

Ebenso variieren von Land zu Land auch die relativen Gewichte, mit denen die Notenpunkte von Wechslern und Nichtwechslern zur durchschnittlichen Notenpunktzahl eines Landes beitragen. Das Bild, das die Ausgangslage beschreibt, kann sich entsprechend ändern, wenn für jedes einzelne Bundesland ergänzend berücksichtigt wird, wie sich die landesspezifischen Binnenverteilungen der Notenpunkte im Vergleich von Wechslern und Nichtwechslern darstellen.

Tabelle 2: Für 10 Bundesländer erwartete durchschnittliche Punktzahlen in der 2. Prüfung, berechnet als mittlere Abweichungen von der globalen Notenerwartung für Assessorinnen und Assessoren

Note2		Modell 1			
		Erwartete ± Abweichungen		Wechsleranteil im jeweiligen Bundesland	
2. Prüfung abgelegt in		b_1^*	t	Anteile	N*
Nordrhein-Westfalen	NRW	+0,516	16,35	18,6 %	5.531
Schleswig-Holstein	SWH	+0,268	3,62	54,0 %	489
Rheinland-Pfalz	RLP	+0,231	5,46	37,7 %	1.633
Baden-Württemberg	BAW	+0,177	4,30	15,9 %	2.259

Tab. 2 (fortgesetzt)

Note2	Modell 1				
		Erwartete ± Abweichungen		Wechsleranteil im jeweiligen Bundesland	
Bremen	HBR	+0,153	1,83	38,3 %	405
<i>Mittlere Notenerwartung</i>		<i>7,245 Punkte ±</i>		<i>17,9 %</i>	<i>18.422</i>
Saarland	SLA	−0,074	−0,86	18,0 %	373
Brandenburg	BRA	−0,081	−1,41	38,2 %	689
Bayern	BAY	−0,193	−6,47	4,6 %	5.642
Sachsen	FSA	−0,369	−7,11	22,7 %	1.035
Mecklenburg-Vorpommern	MVP	−0,629	−7,82	12,0 %	366

*N Zahl der Fälle (Prozentuierungsbasis; Gesamt N=18.422). Lineare Regressionsanalyse (Effektkodierung); Standardfehlerschätzung: robust. Statistische Signifikanz von b ($p < 0,05$), wenn $t > |1,96|$. Berechnungen mit Stata 13.

Modell 2 erweitert das Bild entsprechend (Tabelle 3: Modell 2, „Erwartete ± Abweichungen“). Wie zuvor wird die in der zweiten Staatsprüfung im Durchschnitt erzielte Punktzahl für jedes Bundesland als Abweichung von der über alle Länder hinweg errechneten mittleren Notenerwartung geschätzt. Wie zuvor werden die Bundesländer nach der Höhe dieser Abweichungswerte sortiert. Für fünf der Länder mit den vergleichsweise stärksten Abweichungen von der mittleren Notenerwartung ändert dies nichts an ihrem Platz im Ranking nach diesen Abweichungswerten: Diese waren in Modell 1 genauso platziert und signifikant wie jetzt in Modell 2. Zudem werden sehr ähnliche Abweichungswerte geschätzt. Dies betrifft am oberen Pol der Skala Nordrhein-Westfalen und Schleswig-Holstein und am unteren Pol Bayern, Sachsen und Mecklenburg-Vorpommern. Nur unwesentlich ändern sich zudem der Platz im Ranking und die geschätzten Werte für Rheinland-Pfalz. Für alle diese Bundesländer ändert sich das Bild also nicht, wenn berücksichtigt wird, wie die Noten von Wechslerlern und Nichtwechslerlern zur Notenverteilung in der zweiten Staatsprüfung eines Bundeslandes beitragen: Wie in Modell 1 lautet die Erkenntnis auch hier, dass die bundeslandbezogenen mittleren Werte *signifikant* von der mittleren Notenerwartung abweichen.

Für zwei weitere Länder ist zudem festzustellen, dass die Werte – im Vergleich der Modelle 1 und 2 – ähnlich bleiben und somit zur selben Schlussfolgerung führen. Dies betrifft Brandenburg und das Saarland. In diesen beiden Ländern differieren die Abweichungswerte im Spiegel der Zahlen *beider* Modelle *nicht* signifikant von der mittleren Notenerwartung.

Tabelle 3: Die für 10 Bundesländer erwarteten durchschnittlichen Punktzahlen von Assessorinnen und Assessoren in der 2. Prüfung sowie damit verbundene Brutto- und Nettoeffekte eines Bundeslandwechsels zwischen der 1. und der 2. Prüfung

		Modell 2				Modell 3			
Note 2		Erwartete ± Abweichun- gen von 7,241		Wechsel → Note2		Wechsel → Note2		Note1 → Note2	
		Bruttoeffekte				Nettoeffekte			
2. Prüfung abgelegt in		b_1^*	t	$b_{2 BL2}^*$	t	$b_{2 BL2}^*$	t	$b_{3 BL2}^*$	t
Nordrhein-Westfalen	NRW	+0,489	11,75	-0,050	-1,41	-0,107	-3,49	0,567	43,20
Schleswig-Holstein	SWH	+0,304	4,07	-0,410	-5,27	-0,065	-1,02	0,669	17,57
Bremen	HBR	+0,252	2,91	+0,406	4,41	0,404	5,91	0,670	18,54
Rheinland-Pfalz	RLP	+0,233	4,93	-0,006	-0,13	0,114	3,89	0,657	46,64
Baden-Württemberg	BAW	+0,032	0,59	-0,219	-4,20	-0,147	-4,29	0,711	59,97
Brandenburg	BRA	-0,055	-0,90	+0,092	1,50	0,036	0,71	0,574	18,89
Saarland	SLA	-0,092	-0,85	-0,035	-0,30	0,118	1,54	0,671	23,87
Bayern	BAY	-0,215	-3,58	-0,029	-0,48	-0,080	-2,03	0,716	84,70
Sachsen	FSA	-0,335	-5,24	+0,055	0,84	-0,033	-0,65	0,617	29,57
Mecklenburg- Vorpommern	MVP	-0,612	-4,05	+0,018	0,11	0,189	1,55	0,658	17,19

N=18.422. Lineare Regressionsanalysen (Effektkodierung); Standardfehlerschätzung: robust. Statistische Signifikanz ($p<0,05$), wenn $t > |1,96|$. Berechnungen mit Stata13.

Bleiben die Werte für Baden-Württemberg und Bremen. Hier führt die Berücksichtigung der Binnenverteilung von Wechslern und Nichtwechslern jeweils zu Verschiebungen in Stärke und Signifikanz der Abweichungen der geschätzten mittleren Landeswerte vom generellen Mittel. Während sich die geschätzte Abweichung für Bremen von erwarteten +0,15 Notenpunkten leicht auf +0,25 Notenpunkte verstärkt und von annähernder Signifikanz ($t=1,83$) nun Signifikanz ($t=2,91$) erzielt wird, liegt der Fall bei Baden-Württemberg umgekehrt: Hier fällt die erwartete Abweichung der landesbezogenen Durchschnittsnote von der mittleren Notenerwartung von signifikanten +0,18 Notenpunkten auf insignifikante +0,03 Notenpunkte ab, wenn innerhalb des Landes die Binnenverteilung der Noten nach Wechslern und Nichtwechslern berücksichtigt wird. In der Summe bleibt es aber bei der Erkenntnis, dass bei sieben der zehn untersuchten Bundes-

länder die Notendurchschnitte in der zweiten Staatsprüfung signifikant von der mittleren Notenerwartung abweichen (NRW, SWH, HBR, RLP, BAY, FSA, MVP).

Wie aber stellt sich die Variation der Notendurchschnitte von Wechslern und Nichtwechslern *um die jeweilige geschätzte mittlere Notenpunktzahl in der zweiten Staatsprüfung eines Bundeslandes* dar? Was diese erwarteten \pm Abweichungswerte anbelangt, informiert Tabelle 3 in den Spalten „Wechsel \rightarrow Note2“ über die diesbezüglich geschätzten Brutto- und Nettoeffekte. Von „Bruttoeffekt“ ist dabei die Rede, wenn die Schätzung unberücksichtigt lässt, dass die Note in der zweiten Prüfung *unabhängig von einem etwaigen Wechsel* auch von der in der ersten Prüfung erzielten Punktzahl („Note1“) mitbestimmt wird, wohingegen diese letztgenannte Variationsursache bei der Schätzung des „Nettoeffektes“ eines Bundeslandwechsels statistisch kontrolliert wird. Die Schätzung von Nettoeffekten erfolgt dabei so, dass sich die jeweiligen, in Modell 3 für ein Bundesland ausgewiesenen Haupteffekte, „Wechsel \rightarrow Note2“ und „Note1 \rightarrow Note2“, in ihrer Wirkung addieren.

Kommen wir zunächst auf Baden-Württemberg und Bremen zurück. Für beide Länder gilt, dass sich der jeweilige Bruttoeffekt eines Wechsels in das Land auch als Nettoeffekt erhält. Nach Modell 2 ergeben sich danach für Bremen die durchschnittlichen Notenerwartungen als $7,241 + 0,252 + 0,406 = 7,9$ bei einem Wechsel nach Bremen bzw. als $7,241 + 0,252 - 0,406 = 7,1$, sofern auch die erste Prüfung in Bremen abgelegt wurde. Während sich dabei *beide* Effekte als statistisch signifikant erweisen, ist es in Baden-Württemberg *nur der Wechseleffekt*. Die durchschnittlichen Notenerwartungen liegen dort bei $7,241 + 0,032 - 0,219 = 7,1$ bei einem Wechsel nach Baden-Württemberg bzw. bei $7,241 + 0,032 + 0,219 = 7,5$, sofern auch die erste Prüfung in Baden-Württemberg abgelegt wurde. Während Brutto- und Nettoeffekt eines Wechsels nach Bremen praktisch identisch sind, schwächt sich der Bruttoeffekt eines Wechsels nach Baden-Württemberg leicht von $-0,22$ Notenpunkten auf $-0,15$ Notenpunkte ab, bleibt dabei allerdings statistisch signifikant.

Statistisch signifikante Nettoeffekte eines Wechsels des Bundeslandes zwischen der ersten und der zweiten Staatsprüfung sind insgesamt für fünf Bundesländer festzustellen. Dabei handelt es sich um Bremen (+0,40 Notenpunkte), Rheinland-Pfalz (+0,11), Bayern ($-0,08$), Nordrhein-Westfalen ($-0,11$) und Baden-Württemberg ($-0,15$). Mit Ausnahme von Baden-Württemberg erweist sich in diesen Ländern auch der „Bundeslandeffekt“, also die für ein Land erwartete Abweichung der dort im Durchschnitt erzielten Punktzahl von der mittleren Notenerwartung, als statistisch signifikant. Wenn daher für eine länderübergreifend äquivalente Notengebung in der zweiten Prüfung als operationales Kriterium herangezogen wird, dass nach Abzug des „Note1“- Effektes der in der ersten Prüfung erzielten Note kein systematischer Wechseleffekt übrig bleiben darf,

dann legt das Resultat die Schlussfolgerung nahe, dass eine länderübergreifende Äquivalenz der Benotung zumindest in Bezug auf die genannten fünf Bundesländer fraglich erscheint. Deutlich wird aber auch, dass sich die Verletzung dieser Äquivalenz bei Unterschieden von $|0,08|$ bis $|0,40|$ Notenpunkten, trotz statistischer Signifikanz in punkto *substanzieller* Größenordnung, deutlich unterhalb der Differenz von einem Notenpunkt bewegt und insofern vergleichsweise gering ist.

Nicht überraschend ist in diesem Zusammenhang der Befund, dass sich die in der ersten Staatsprüfung erzielte Note stets signifikant, positiv und stark auf die in der zweiten Staatsprüfung erzielte Note auswirkt (Tabelle 3: Modell 3, „Note1 \rightarrow Note2“). Die Effektstärken rangieren dabei zwischen 0,567 in Nordrhein-Westfalen und 0,716 in Bayern. Zu erwarten ist entsprechend, dass sich je nach Bundesland, und bei nur geringer Variation über die betrachteten Länder hinweg, sich jeder Punkt in *Note1* in einem Plus von durchschnittlich 0,6 bis 0,7 *Note2*-Punkten niederschlägt.

Die in den Modellen 2 und 3 für die zehn Zielländer berechneten Wechselwirkungen ermöglichen den Vergleich von Personen, die zur zweiten Prüfung in ein Zielland gewechselt sind, mit Personen, die dort auch ihre erste Prüfung abgelegt haben. Die Effekte werden also ziellandspezifisch geschätzt, differenzieren dabei aber nicht, *aus welchem Bundesland* heraus in das Zielland gewechselt wurde. Eine derartige Perspektive auf Wechselwirkungen ist von Vorteil, weil sie eine kompakte Schätzung des Gesamteffektes eines Wechsels in ein Zielland liefert. Sie birgt allerdings die potentielle Schwäche in sich, dass die Schätzung eines solchen Gesamteffektes auf tendenziell gegenläufigen Einzeleffekten beruhen kann. Theoretisch vorstellbar ist entsprechend die Konstellation, dass starke, aber gegenläufige Einzeleffekte nur einen schwachen Gesamteffekt produzieren. In der Gesamtbetrachtung ist dann sogar die Möglichkeit einer *scheinbaren* Effektlöslichkeit denkbar. Es bietet sich entsprechend an, auch die involvierten Einzeleffekte eines Bundeslandwechsels zu untersuchen.

Für eine solche, auf Einzeleffekte abstellende Analyse wäre die Verfügbarkeit kompletter Informationen in Bezug auf die Noten der ersten und zweiten Prüfung und darauf, in welchem Bundesland die Examina jeweils abgelegt wurden, ideal. Im Prinzip liegen genau diese Informationen auch für die zehn Bundesländer vor, die für die vorliegende Studie Noten gemeldet haben. Allerdings ist für diese Ländergruppe eine Einschränkung zu beachten, die als Folge des faktischen Wechselverhaltens auftritt: Manche Bundeslandkombinationen treten bei Wechseln zwischen der ersten und der zweiten Staatsprüfung häufiger auf, manche seltener und manche gar nicht. Dies führt dazu, dass nicht jede theoretisch mögliche Bundeslandkombination empirisch auftritt und dass die Fallzahlen von Wechs-

lern teilweise klein bis sehr klein ausfallen. Für eine belastbare statistische Wechselanalyse sind aber Mindestfallzahlen einzuhalten.

Tabelle 4: Brutto- und Nettoeffekte eines Bundeslandwechsels, für einzelne Bundeslandkombinationen sortiert nach der Größe des jeweiligen Nettoeffektes

Wechsel		Modell 4		Modell 5		Wechsel		Bruttoeffekte		Nettoeffekte	
		Bruttoeffekte		Nettoeffekte		BL 1 → BL 2		$b_{2 BL2}$	t	$b_{2 BL2}$	t
BL 1 → BL 2		$b_{2 BL2}$	t	$b_{2 BL2}$	t	BRA	FSA	0,475	1,33	-0,223	-0,62
NIE	HBR	0,882	3,26	0,834	4,18	SLA	NRW	-0,385	-1,26	-0,252	-1,05
BAW	FSA	0,914	1,96	0,661	2,12	FTH	BAY	0,185	0,54	-0,288	-1,32
BAY	FSA	0,234	0,79	0,564	2,81	HHA	SWH	-1,264	-7,69	-0,297	-2,11
NRW	SWH	0,130	0,29	0,477	1,36	MVP	SWH	-0,869	-2,05	-0,315	-0,88
HHA	NRW	1,366	3,60	0,351	1,06	MVP	NRW	-0,475	-1,36	-0,549	-1,34
BAY	SWH	0,039	0,10	0,205	0,74	MVP	BRA	-0,616	-2,68	-0,597	-3,40
BAY	RLP	0,265	1,14	0,191	1,32	SAH	FSA	-0,157	-0,56	-0,702	-3,69
FTH	FSA	0,194	1,03	0,006	0,04	HES	NRW	-0,821	-4,87	-0,908	-5,31
NIE	NRW	-0,118	-0,83	-0,078	-0,59	HES	RLP	-1,269	-5,09	-0,999	-4,70
NIE	BAY	-0,155	-0,49	-0,206	-1,03	HES	BAW	-1,239	-3,96	-1,131	-5,53

*Brutto/Netto = Ohne/Mit Berücksichtigung des Effekts von *Note 1*. Multiple Regressionsanalysen (Dummy-Kodierung). Referenzkategorie (jeweils): BL1 = BL2 = Diejenigen, die *beide* Prüfungen im jeweiligen → Zielland abgelegt haben. Standardfehlerschätzung: robust. Statistische Signifikanz ($p < 0,05$), wenn $t > |1,96|$. Berechnungen mit Stata13. Ausgewiesen werden erwartete Unterschiede in mittleren *Note 2*-Werten von Personen, die ihre erste Prüfung in dem Land vor dem → und ihre 2. Prüfung in dem Land nach dem → absolviert haben, jeweils bezogen auf die Gruppe derer, die auch ihre 1. Prüfung in dem Land nach dem → abgelegt haben.

In der vorliegenden Analyse werden die gesuchten Brutto- und Nettoeffekte eines Wechsels *explizit* daher auch nur für diejenigen Bundeslandkombinationen geschätzt, bei denen die Zahl der *Wechsler* $n \geq 20$ ist.¹⁵ Kein Fallzahlproblem ent-

15 Die fallzahlbedingt auszuschließenden Bundeslandkombinationen bleiben dabei allerdings in den zu schätzenden Regressionsgleichungen enthalten; und zwar jeweils über eine zusammenfassende Residualkategorie, die sicherstellt, dass sich die Schätzungen der Effekte der zu berücksichtigenden Bundeslandkombinationen nicht mit den Effekten der fallzahlbedingt auszuschließenden Bundeslandkombinationen vermengen („konfundieren“) können.

steht hingegen mit Blick auf die Personengruppen, welche die erste und die zweite Prüfung jeweils im selben Land abgelegt haben, da diese Konstellation stets sehr häufig auftritt. Tabelle 5 informiert vor diesem Hintergrund über die Ergebnisse zu denjenigen Bundeslandkombinationen, für die Wechsel *in beide Richtungen*, also von Land A nach B und von Land B nach A, analysiert werden konnten. Zuvor informiert Tabelle 4 über die Ergebnisse zu denjenigen Bundeslandkombinationen, in denen *datenbedingt* Wechsel *nur in eine Richtung* auswertbar waren. Dies betrifft zum einen die Beziehungen *zwischen* den „Notenmeldeländern“, bei denen fallzahlbedingt nur eine Richtung, nicht aber die Gegenrichtung auswertbar war, sowie Wechsel *in* diesen Kreis von Bundesländern, und zwar aus Bundesländern heraus, die selbst keine Noten gemeldet hatten.

Tabelle 4 informiert für diese Konstellationen über die Brutto- und Nettoeffekte eines Bundeslandwechsels sortiert nach der Größe des jeweiligen Nettoeffektes. Der stärkste *positive* Nettoeffekt ergibt sich danach für einen Wechsel von Niedersachsen nach Bremen: Verglichen mit denen, die in Bremen auch ihre erste Staatsprüfung abgelegt haben (Bezugsgröße: dortige Nichtwechsler), wird für Wechsler aus Niedersachsen nach Bremen *unabhängig von der in der ersten Staatsprüfung erzielten Leistung* („Note1“) ein signifikantes Plus von durchschnittlich 0,834, also von rd. 0,8 Punkten, auf der Notenskala der zweiten Staatsprüfung erwartet. Was diesen Wechsel nach Bremen anbelangt, hatte oben auch die in Modell 3 eingenommene Perspektive auf Wechselwirkungen einen signifikant positiven Wechselwirkungseffekt aufgedeckt (Tabelle 3).

Signifikant positive Wechselwirkungen sind zudem für zwei Wechsel nach Sachsen festzustellen. So wird bei einem Wechsel von Baden-Württemberg nach Sachsen ein Nettoeffekt von durchschnittlich 0,661, also von rd. 0,7 Notenpunkten erwartet, während der Nettoeffekt bei einem Wechsel von Bayern nach Sachsen zugleich auf ein Plus von 0,564 bzw. 0,6 Notenpunkte geschätzt wird.

Daneben werden in Tabelle 4 für Sachsen als Zielland drei weitere Nettoeffekte eines Bundeslandwechsels ausgewiesen, von denen ein Effekt statistisch signifikant ist. Er betrifft den Wechsel von Sachsen-Anhalt nach Sachsen: Für einen Wechsel von Sachsen-Anhalt nach Sachsen wird der Nettoeffekt auf ein Minus von 0,7 Notenpunkten geschätzt. Verglichen mit der Bezugsgröße derjenigen, die in Sachsen auch ihre erste Prüfung abgelegt haben, wird danach für Wechsler von Sachsen-Anhalt nach Sachsen unabhängig davon, welche Punktzahl auf der Skala der ersten Staatsprüfung erzielt wurde, eine um 0,7 Notenpunkte signifikant schlechtere Benotung in der zweiten Staatsprüfung erwartet. Zusammengefasst sind für Sachsen also zwei signifikant positive und ein signifikant negativer Nettoeffekt zu verzeichnen. Nicht überraschend ist im Spiegel dieser Erkenntnis dann auch der Sachverhalt, dass diese gegenläufigen Einzeleffekte

fekte für das Zielland Sachsen einen insignifikanten Gesamteffekt nahe Null produzieren können, wie er oben in Modell 3 für das Zielland Sachsen festzustellen war.

Signifikant negative Nettoeffekte eines Bundeslandwechsels auf die in der zweiten Staatsprüfung erzielte Punktzahl („Note2“) sind dabei für eine ganze Reihe von Bundeslandkombinationen zu beobachten (Tabelle 4). Die Größe dieses Nettoeffekts bewegt sich dabei innerhalb eines Spektrums von $-0,3$ (Hamburg \rightarrow Schleswig-Holstein), $-0,6$ (Mecklenburg-Vorpommern \rightarrow Brandenburg), $-0,7$ (Sachsen-Anhalt \rightarrow Sachsen), $-0,9$ (Hessen \rightarrow Nordrhein-Westfalen), $-1,0$ (Hessen \rightarrow Rheinland-Pfalz) und $-1,1$ Notenpunkten (Hessen \rightarrow Baden-Württemberg). Für diese Bundeslandkonstellationen ist entsprechend zu erwarten, dass Wechsler in das betreffende Zielland um den jeweiligen Betrag geringere mittlere „Note2“-Punktzahlen erzielen als Personen, die im betreffenden Zielland auch ihre erste Prüfung abgelegt haben, und zwar unabhängig von der in der ersten juristischen Staatsprüfung erzielten Note.

Die Modelle 4 und 5 stellen auch die Grundlage der in Tabelle 5 ausgewiesenen Schätzungen von Brutto- und Nettoeffekten eines Bundeslandwechsels dar. Tabelle 5 ergänzt Tabelle 4 nur dahingehend, dass sie diejenigen Bundeslandkombinationen ausweist, für welche die besagten Effekte *jeweils in beide Richtungen*, also von Land A nach B und von Land B nach A, berechnet werden konnten.

Tabelle 5: Brutto- und Nettoeffekte eines Bundeslandwechsels, für Bundeslandkombinationen, in denen Wechsel bidirektional auswertbar waren

		Modell 4: Bruttoeffekte eines Bundeslandwechsels				Modell 5: Nettoeffekte eines Bundeslandwechsels			
		Land A \rightarrow Land B		Land B \rightarrow Land A		Land A \rightarrow Land B		Land B \rightarrow Land A	
Land A	Land B	$b_{2 BL2}$	t	$b_{2 BL2}$	t	$b_{2 BL2}$	t	$b_{2 BL2}$	t
NRW	HBR	1,936	4,73	-0,842	-2,50	1,573	5,68	-0,629	-1,92
RLP	NRW	-0,284	-1,73	0,056	0,32	-0,432	-3,63	0,526	4,55
BRA	BAW	-0,689	-2,27	1,070	3,46	-0,808	-4,09	0,839	3,43
RLP	BAW	-0,138	-0,34	0,296	2,39	-0,445	-1,77	0,550	6,74
NRW	BAW	0,469	1,32	0,710	4,09	0,152	0,77	0,564	4,21
BAY	BAW	-0,445	-3,18	0,372	1,65	0,010	0,11	0,482	2,86
BRA	BAY	-0,356	-1,12	0,518	2,11	-0,339	-1,44	0,299	1,41
NRW	BAY	0,314	1,07	-0,017	-0,11	0,068	0,40	-0,152	-1,14

Tab. 5 (fortgesetzt)

		Modell 4: Bruttoeffekte eines Bundeslandwechsels				Modell 5: Nettoeffekte eines Bundeslandwechsels			
		Land A → Land B		Land B → Land A		Land A → Land B		Land B → Land A	
NRW	BRA	0,262	1,16	0,203	0,73	0,282	1,64	-0,115	-0,49
SLA	RLP	-0,366	-2,33	0,222	0,62	-0,074	-0,76	0,357	1,84

*Brutto/Netto = Ohne/Mit Berücksichtigung des Effekts von *Note1*. Multiple Regressionsanalysen. Referenzkategorie (jeweils): BL1 = BL2 = Diejenigen, die *beide* Prüfungen im jeweiligen → Zielland abgelegt haben. Ausgewiesen werden erwartete Unterschiede in mittleren *Note2*-Werten von Personen, die ihre erste Prüfung in dem Land vor dem → und ihre 2. Prüfung in dem Land nach dem → absolviert haben, jeweils bezogen auf die Gruppe derer, die auch ihre 1. Prüfung in dem Land nach dem → abgelegt haben.

Wie auch bei den in Tabelle 4 berichteten unidirektionalen Analysen ist für das Zielland Bremen auch hier ein starker *positiver* Wechseleffekt zu verzeichnen. Er betrifft den Wechsel von Nordrhein-Westfalen nach Bremen, der mit signifikanten 1,573, also rd. 1,6 Punkten auf der Notenskala der zweiten Staatsprüfung recht stark ausfällt. Verglichen mit denen, die in Bremen auch die erste Staatsprüfung abgelegt haben, ist für Wechsler aus NRW also unabhängig von ihren Noten in der ersten juristischen Staatsprüfung ein Notenplus von 1,6 Punkten zu erwarten. Ein Wechsel aus Nordrhein-Westfalen nach Bremen erweist sich so gewissermaßen als profitabel. Im Gegenzug beobachten wir sinnentsprechend einen *negativen* Wechseleffekt von -0,629, der mit $t=1,92$ das Signifikanzkriterium von $t=1,96$ nur äußerst knapp verfehlt. Für Personen, die zur zweiten Staatsprüfung von Bremen nach Nordrhein-Westfalen wechseln, ist danach zu erwarten, dass ihre durchschnittliche Punktzahl in der zweiten Staatsprüfung um rd. 0,6 Punkte unter derjenigen von Personen liegt, die auch ihre erste juristische Prüfung in NRW abgelegt haben.

Bei zwei weiteren Länderkombinationen treffen signifikante Punktgewinne bei einem Wechsel von Land A nach B auf signifikante Punktverluste bei einem Wechsel in Gegenrichtung. Dies betrifft die Kombinationen Rheinland-Pfalz und Nordrhein-Westfalen mit Nettoeffekten von -0,4 und +0,5 sowie Brandenburg und Baden-Württemberg mit Nettoeffekten von -0,8 und +0,8 (Tabelle 5, Ergebniszeilen 2 und 3). In drei Bundeslandrelationen erweist sich dann jeweils nur eine der beiden Wechselrichtungen als signifikant (Tabelle 5, Ergebniszeilen 4, 5 und 6). Dies betrifft die Wechsel von Baden-Württemberg nach Rheinland-Pfalz, von Baden-Württemberg nach Nordrhein-Westfalen und von Baden-Württemberg nach Bayern. Schließlich erweisen sich für vier Länderrelationen die Nettoeffekte

in *keine* der beiden gegenläufigen Richtungen als signifikant. Dies betrifft die Länderrelationen Brandenburg – Bayern, Nordrhein-Westfalen – Bayern, Nordrhein-Westfalen – Brandenburg und Saarland – Rheinland-Pfalz, wenngleich sich auch bei diesen Konstellationen stets das Muster zeigt, dass ein positiver Nettoeffekt in eine Richtung *tendenziell* auf einen negativen Nettoeffekt in Gegenrichtung trifft.

Zusammenfassend lässt sich konstatieren, dass bei sieben der zehn untersuchten Bundesländer die Ergebnisse in der zweiten Staatsprüfung signifikant von der mittleren Notenerwartung abweichen. Dies betrifft Nordrhein-Westfalen, Schleswig-Holstein, Bremen, Rheinland-Pfalz, Bayern, Sachsen und Mecklenburg-Vorpommern. Zudem waren statistisch signifikante Nettoeffekte eines Wechsels des Bundeslandes zwischen der ersten und zweiten Staatsprüfung für insgesamt fünf Bundesländer festzustellen. Dabei handelt es sich um Bremen, Rheinland-Pfalz, Bayern, Nordrhein-Westfalen und Baden-Württemberg. Tabelle 6 zeigt diesen Befund nochmals in der Übersicht.

Tabelle 6: Die Bundesland- und Wechseleffekte in der Übersicht

		„Bundeslandeffekt“	Wechseleffekt
Nordrhein-Westfalen	NRW	•	•
Schleswig-Holstein	SWH	•	
Bremen	HBR	•	•
Rheinland-Pfalz	RLP	•	•
Baden-Württemberg	BAW		•
Bayern	BAY	•	•
Sachsen	FSA	•	
Mecklenburg-Vorpommern	MVP	•	

Die in diesen Analysen eingenommene Perspektive auf Wechseleffekte war derart, dass die diversen Wechsel in ein Zielland in der Schätzung des Wechseleffekts zusammengefasst wurden, also nicht danach differenziert wurde, aus welchen Bundesländern heraus in das jeweilige Zielland gewechselt wurde. Soweit möglich, wurden daher in weiteren Analysen die den Wechseln zugrundeliegenden spezifischen Bundeslandkombinationen genauer unter die Lupe genommen.

Tabelle 7: Nettoeffekte eines Bundeslandwechsels, für Bundeslandkombinationen mit signifi-
kanten Nettoeffekten (Zusammenfassung)

Wechsel		Modell 5	
Nettoeffekte			
BL 1 → BL 2		b	t
NRW	HBR	1,573	5,68
BAW	BRA	0,839	3,43
NIE	HBR	0,834	4,18
BAW	FSA	0,661	2,12
BAW	NRW	0,564	4,21
BAY	FSA	0,564	2,81
BAW	RLP	0,550	6,74
NRW	RLP	0,526	4,55
BAW	BAY	0,482	2,86
HHA	SWH	−0,297	−2,11
RLP	NRW	−0,432	−3,63
MVP	BRA	−0,597	−3,40
SAH	FSA	−0,702	−3,69
BRA	BAW	−0,808	−4,09
HES	NRW	−0,908	−5,31
HES	RLP	−0,999	−4,70
HES	BAW	−1,131	−5,53

Tabelle 7 fasst die diesbezüglichen Ergebnisse aus den Tabellen 4 und 5 in der Weise zusammen, dass sie die dort jeweils ausgewiesenen statistisch signifi-
kanten Nettoeffekte eines Bundeslandwechsels nach ihrer Größe sortiert auflistet.
Deutlich wird dadurch noch einmal, dass sich die erwarteten wechselbedingten
Notenunterschiede im Spiegel der so entstehenden Reihe von Bundeslandrelatio-
nen zwischen −1,1 und +1,6 Notenpunkten bewegen. Je nachdem, ob der jeweilige
Effekt positiv oder negativ ausfällt, erweist sich ein Wechsel entsprechend ganz
unabhängig von der in der ersten Prüfung erzielten Punktzahl für die wechselnde
Person als vorteilhaft oder nachteilig, und zwar jeweils verglichen mit denjeni-
gen, die im jeweiligen Zielland auch ihre erste juristische Prüfung abgelegt haben.

Durch die Analyse der spezifischen Wechsel von Land A nach B wird zudem sichtbar, dass Wechsel in ein und dasselbe Zielland sowohl positiv als auch negativ ausfallen können, und zwar je nachdem, aus welchem Bundesland heraus gewechselt wurde. Dies betrifft etwa die Zielländer Rheinland-Pfalz und Nordrhein-Westfalen.

Schlussfolgerungen

Welche Schlüsse lassen sich aus diesen Ergebnissen ziehen? Deutlich ist zunächst, dass über alle untersuchten Bundesländer hinweg eine Korrelation zwischen der ersten und der zweiten juristischen Staatsprüfung besteht. Vermutungen, dass eine eher an der Theorie orientierte erste Prüfung nichts über die eher an der Praxis ausgerichtete zweite Prüfung aussagt, fehlt damit die Grundlage. Der Zusammenhang zwischen den beiden juristischen Staatsprüfungen ist für alle Bundesländer ähnlich eng (Tabelle 3).

Des Weiteren deuten die Ergebnisse darauf hin, dass es im Schwierigkeitsgrad der zweiten juristischen Staatsprüfungen bundesweite nennenswerte Unterschiede gibt. Dies wäre jedenfalls eine Erklärung für das gefundene systematisch unterschiedliche Abschneiden von Wechslern und Nichtwechslern mit jeweils gleichen Noten in der ersten juristischen Staatsprüfung, also bei Betrachtung der Nettoeffekte eines Bundeslandwechsels.

Dies wäre zudem eine Erklärung dafür, dass zwischen den zehn untersuchten Ländern auch bei Berücksichtigung von Wechseleffekten systematische Notenunterschiede bestehen bleiben, und zwar innerhalb eines Rahmens von ca. 1,1 Notenpunkten Differenz (Tabelle 3, Modell 2). Diese Unterschiede zeigen in der Tendenz ein Nord/West-Süd/Ost-Gefälle auf. Auf der einen Seite sehen wir Schleswig-Holstein, Bremen, Nordrhein-Westfalen und Rheinland-Pfalz mit Durchschnittsnoten oberhalb der mittleren Notenerwartung der untersuchten Länder. Dies deutet für diese Länder auf einen vergleichsweise geringeren Schwierigkeitsgrad der zweiten juristischen Staatsprüfung hin. Auf der anderen Seite stehen Bayern, Sachsen und Mecklenburg-Vorpommern mit Durchschnittsnoten, die unterhalb dieser mittleren Notenerwartung liegen (Tabelle 3). Dies deutet einen höheren Schwierigkeitsgrad der Prüfungen an.

Auch die Analyse von Einzeleffekten liegt insoweit auf dieser Linie (Tabelle 7), als die beiden Länder, für welche die stärksten Wechseleffekte zu beobachten sind, das Muster unterstützen. Auf der einen Seite ist dies Bremen, für das als eines von zwei Ländern nur positive Wechseleffekte zu berichten sind. Auf der

anderen Seite ist Baden-Württemberg als eines von zwei Ländern anzuführen, für das die Tabelle nur negative Wechseleffekte ausweist.

Allerdings zeigt die Wechselanalyse ebenso deutlich auf, wie wichtig und notwendig es ist, nicht nur zu betrachten, in welches Land gewechselt wird, sondern auch, aus welchem Land heraus gewechselt wird. So zeigen sich bei einem Wechsel nach Brandenburg, Nordrhein-Westfalen und Sachsen sowohl positive als auch negative Wechseleffekte, und zwar je nachdem, aus welchem Bundesland die Kandidaten jeweils kommen (Tabelle 7).

Wechseleffekte treten dabei auch im Binnenverhältnis von Ländern auf, die mit Blick auf das oben skizzierte Nord/West-Süd/Ost-Gefälle entweder der einen oder der anderen der beiden „Ländergruppen“ zuzurechnen wären. Das betrifft auf der einen Seite den mit einem erwarteten Plus von 1,6 Notenpunkten verknüpften Wechsel von Nordrhein-Westfalen nach Bremen, den Wechsel von Nordrhein-Westfalen nach Rheinland-Pfalz (+0,5 Notenpunkte) und in Gegenrichtung den Wechsel von Rheinland-Pfalz nach Nordrhein-Westfalen (–0,4 Notenpunkte). Auf der anderen Seite ist ein Wechsel von Bayern nach Sachsen mit einem erwarteten Notenplus von 0,6 Punkten verbunden (Tabelle 7).

Darüber hinaus ergibt sich im Spiegel der Zahlen von Tabelle 7 auch für Baden-Württemberg ein interessanter Befund. Einerseits sind Wechsel in das Land in beiden analysierten Konstellationen mit einem erwarteten Notenminus verbunden, und zwar von 0,8 Punkten bei einem Wechsel aus Brandenburg, und von 1,1 Punkten bei einem Wechsel aus Hessen. Andererseits sind Wechsel aus Baden-Württemberg heraus gleich in verschiedene andere Länder mit einem erwarteten Notenplus verbunden; so ein Wechsel nach Brandenburg (+0,8), nach Sachsen (+0,7), nach Nordrhein-Westfalen (+0,6), nach Rheinland-Pfalz (+0,6) und nach Bayern (+0,5 Notenpunkte).

Je nach Herkunfts- und Zielort sind für die in Tabelle 7 ausgewiesenen Konstellationen Notenunterschiede in der Größenordnung von –1,1 bis 1,6 Notenpunkte allein durch den Wechsel des Ortes zu erklären. Wechsel können also für die wechselnden Personen sowohl von Vorteil als auch von Nachteil sein. Zugleich stellen diese Effekte aber auch die länderübergreifende Äquivalenz der Notengebung in Frage, da sie als Nettoeffekte um den Einfluss individueller Leistungsvoraussetzungen korrigiert sind, und somit auszuschließen ist, dass sich ihnen ggf. systematisch nach Ländern variierende Unterschiede in diesen Leistungsvoraussetzungen widerspiegeln.

Damit sind die Unterschiede angesichts der Bedeutung der Staatsprüfungen als erheblich anzusehen. Die Wahl des Examensortes wirkt sich auf die Höhe des zu erwartenden Examens aus. Berücksichtigt man aber, dass die Punkteskala von 0 bis 18 reicht und sich Notenverteilungen nicht auf Durchschnittswerte reduzieren lassen, so wird deutlich, dass trotz dieser systematischen Unterschiede viel

Spielraum in den individuellen Noten bleibt. Systematische Unterschiede in zentralen Tendenzen von bundeslandbezogenen Notenverteilungen ist eine Sache, die Variation der Examensnoten von Person zu Person eine andere. Auch diese Variation ist nicht unerheblich, und zwar per se wie auch im Bundeslandvergleich betrachtet. So erstrecken sich beispielsweise die mittleren 50 Prozent der Examensnoten je nach Bundesland über eine Spanne von 2,1 bis 3 Skalenpunkten (Tabelle 1). Spielraum existiert also nicht nur nominell. Die jedem Prüfling über das Notenspektrum eröffnete prinzipielle Chance, auch gute Examensnoten erzielen zu können, besteht trotz systematischer Unterschiede in den zentralen Tendenzen der Notenverteilungen und wird de facto auch genutzt. Allerdings sind es genau diese systematischen Unterschiede, die es geraten erscheinen lassen, die Vergleichbarkeit der Prüfungen zu verbessern.

Unterschiede zwischen den Bundesländern sind angesichts des Föderalismus nicht untypisch¹⁶ und regen womöglich einen Wettbewerb an. Denjenigen Bundesländern, für die der Wechseleffekt besonders groß ist, sollte die Statistik allerdings ein Anlass zur Prüfung sein, inwieweit die Bewertungen der Staatsprüfungen verändert werden sollten, um eine bessere Vergleichbarkeit herzustellen.

Die vorliegende Studie ist auf zehn der sechzehn Bundesländer beschränkt und konnte aufgrund von teilweise zu geringen Fallzahlen auch einige Einzeleffekte nicht schätzen. Wenn aber schon die Auswertung von zehn Bundesländern erhebliche Unterschiede in den Staatsprüfungen aufzeigen kann, dürfte es künftig sinnvoll sein, die von § 5d Abs. 1 S. 2 DRiG geforderte Herstellung einheitlicher Prüfungsanforderungen durch eine Wechselstatistik zu erleichtern, die sämtliche Bundesländer erfasst. Denkbar wäre eine Abstimmung etwa dadurch, dass das Bundesjustizministerium für die schon derzeit erstellte Ausbildungsstatistik künftig auch die Examensorte der einzelnen Kandidaten abfragt. Eine auf dieser Grundlage erstellte Statistik zeigte, ob ein Bundesland strenger oder nachsichtiger als die übrigen Bundesländer bewertet.

Trotz aller Erwartungswerte und statistischen Korrelationen sind individuelle Verbesserungen zwischen der ersten und zweiten Staatsprüfung nach wie vor möglich und anstrebenswert. Ohnehin sollte die Wahl des Studien- und Ausbildungsorts von anderen Faktoren abhängen als von der Leichtigkeit der dortigen Prüfung.

¹⁶ *BVerfG*, NJW 1994, 1577: 1583 verlangt nur (aber auch) einen im wesentlichen einheitlichen Gesetzesvollzug.

Literatur

- Benke, Michael, Hermann, Dieter & Tag, Brigitte (1996) Determinanten der Studienleistung, Ergebnisse einer Befragung von Examenskandidaten, S. 88–114 in D. Hermann & B. Tag (Hrsg.), *Die universitäre Juristenausbildung*. Bonn: Dt. Hochschulverband.
- Böckenförde, Ernst-Wolfgang (1998) Juristenausbildung – auf dem Weg ins Abseits?, S. 63–88 in D. Stempel (Hrsg.), *Juristenausbildung zwischen Internationalität und Individualität*. Baden-Baden: Nomos.
- Bundesamt für Justiz (2014) *Ausbildungsstatistik für 2014*, <https://www.bundesjustizamt.de/DE/Themen/Buergerdienste/Justizstatistik/Juristen/Ausbildung.html> (abgerufen am 20.3.2017).
- Coleman, James S. (1981) *Longitudinal Data Analysis*. New York: Basic Books.
- Fries, Marlene (2002) Abitur und Studienerfolg. Welchen „Wert“ hat das Abitur für ein erfolgreiches Studium? *Beiträge zur Hochschulforschung* 24 (1): 30–51.
- Hinz, Thomas & Röhl, Hans Christian (2016) Juristische Fakultäten in Baden-Württemberg – Wo studiert man am besten? *Verwaltungsblätter für Baden-Württemberg* 37: 20–23.
- Küchler, Manfred (1979) *Multivariate Analyseverfahren*. Stuttgart: Teubner Studienskripten.
- Labe, Michael (2013) Examensnoten. *Kammerreport Hanseatische Rechtsanwaltskammer Hamburg* 4/2013: 9–10, http://rak-hamburg.de/h/2013_215_de.php (abgerufen am 5.6.2017).
- Meier, Bernd-Dieter (2003) Ist der Erfolg im Jurastudium vorhersagbar? Empirische Befunde zum Zusammenhang zwischen Schulnoten und Abschneiden im Ersten Juristischen Staatsexamen. *Beiträge zur Hochschulforschung* 25 (4): 18–35.
- Staats, Johann Friedrich (2012) *Deutsches Richtergesetz*. Baden-Baden: Nomos.
- Stephan, Hermann (2002) Lässt sich aus der Abiturgesamtnote die Studieneignung erkennen? *Zeitschrift für Rechtspolitik* 35: 83–85.
- Streng, Franz (1996) Determinanten und Indikatoren von Examenserfolg und Studiendauer im Jurastudium – Ergebnisse der Absolventenbefragung der Juristischen Fakultät Erlangen, S. 88–114 in D. Hermann & B. Tag (Hrsg.), *Die universitäre Juristenausbildung*. Bonn: Dt. Hochschulverband.
- Towfigh, Emanuel, Traxler, Christian & Glöckner, Andreas (2014) Zur Benotung in der Examensvorbereitung und im ersten Examen. Eine empirische Analyse. *Zeitschrift für Didaktik der Rechtswissenschaft* 1: 8–27.