

PROJEKTBERICHT

Empirische Untersuchung zur Benotung in der staatlichen Pflichtfachprüfung und in der zweiten juristischen Staatsprüfung in Nordrhein-Westfalen von 2006 bis 2016

vorgelegt von

Prof. Dr. Andreas Glöckner^{1,2}
Prof. Dr. Emanuel V. Towfigh^{3,2}
Prof. Dr. Christian Traxler^{4,2}

¹ FernUniversität in Hagen

² Max-Planck-Institut zur Erforschung von Gemeinschaftsgütern, Bonn

³ EBS Universität – Law School, Wiesbaden

⁴ Hertie School of Governance, Berlin

7. Dezember 2017

Inhalt

Zusammenfassung	1
1 Einleitung.....	2
2 Bisherige Befunde.....	3
2.1 Geschlecht und Migrationshintergrund im ersten juristischen Examen	3
2.2 Differentielle Geschlechterunterschiede in Staatsprüfungen für Lehrer in Frankreich	4
2.3 Zusammensetzung der Prüfungskommission	4
3 Zielstellung und allgemeine Vorgehensweise	5
4 Forschungsfragen.....	6
5 Beschreibung des Datensatzes	7
6 Analysen zu Effekten des Geschlechts	10
6.1 Effekte des Geschlechts im ersten Examen	10
6.1.1 Unterschiede in der Gesamtnote	10
6.1.2 Unterschiede in der Wahrscheinlichkeit, relevante Notenschwellen zu erreichen	11
6.2 Effekte des Geschlechts im zweiten Staatsexamen	13
6.2.1 Unterschiede in der Gesamtnote	13
6.2.2 Unterschiede in der Wahrscheinlichkeit, relevante Notenschwellen zu erreichen	14
6.2.3 Differenzierte Betrachtung der Geschlechterunterschiede in mündlicher und schriftlicher Note im zweiten Examen.....	16
6.2.4 Vornotenorientierte Notenvergabe in mündlichen Prüfungen	17
6.2.5 Detailanalyse der Geschlechterunterschiede in den mündlichen Teilen	18
6.2.6 Entwicklung der Geschlechterunterschiede zwischen dem ersten und zweiten Examen sowie über Abschlussjahrgänge hinweg.....	21
7 Analysen zu Effekten eines Migrationshintergrunds	22
7.1 Effekte des Migrationshintergrunds im ersten Examen.....	23
7.1.1 Unterschiede in der Gesamtnote	23
7.1.2 Wahrscheinlichkeiten der Erreichung von Notenstufen im ersten Examen	23
7.2. Effekte des Migrationshintergrunds im zweiten Examen.....	25
7.2.1 Unterschiede in der Gesamtnote	25
7.2.2 Wahrscheinlichkeit, relevante Notenschwellen zu erreichen.....	26
7.2.3 Differenzierte Analyse der Migrationseffekte auf schriftliche und mündliche Noten	26
7.2.4 Differenzierte Analyse der einzelnen Rechtsgebiete.....	27
7.2.5 Entwicklung zwischen beiden Abschlussprüfungen sowie über Abschlussjahrgänge	28
8 Überblick über die Ergebnisse	29
8.1 Analysen zu Effekten des Geschlechts.....	29
8.2 Analysen zu Effekten eines Migrationshintergrunds.....	29
Appendix A: Geschlechtseffekte.....	31
Appendix B: Migrationseffekte.....	41

ZUSAMMENFASSUNG

Anhand der Analysen von bis zu 18.000 Ergebnissen der ersten und zweiten juristischen Staatsprüfungen in Nordrhein-Westfalen (NRW) aus den Jahren 2006 bis 2016 wurde untersucht, inwieweit Unterschiede in der Benotung von (a) Männern und Frauen und (b) von Personen mit bzw. ohne Indikatoren für einen Migrationshintergrund bestehen.

Frauen erzielen sowohl in der ersten juristischen Prüfung als auch im zweiten Staatsexamen schlechtere Noten als Männer. Ein klarer zeitlicher Trend in diesen Geschlechterunterschieden lässt sich nicht finden. Die Unterschiede in beiden Examen sind ähnlich stark ausgeprägt. Die Unterschiede im zweiten Examen bleiben nach Kontrolle für die Note aus dem 1. Examen – auf einem niedrigen Niveau – statistisch signifikant. Frauen erreichen sowohl im ersten als auch im zweiten Examen mit deutlich geringerer Wahrscheinlichkeit als Männer eine Prädikatsnote. In den mündlichen Noten sind größere Geschlechterunterschiede zu beobachten als in den schriftlichen Noten (zweites Examen). Diese Unterschiede bleiben auch bestehen, wenn für die Note aus dem schriftlichen Teil und/oder aus dem ersten Examen kontrolliert wird. Die Wahrscheinlichkeit, mit einer guten mündlichen Note die relevante Schwelle für eine Gesamtnote über der nächsten Noten-Stufe (z.B. über 9,0 Punkte) zu erreichen, ist – selbst bei identischen schriftlichen Noten – bei Männern deutlich höher als bei Frauen. Dieser Geschlechterunterschied verschwindet jedoch, wenn mindestens eine Frau Teil der Prüfungskommission in der mündlichen Prüfung ist. Die Zusammensetzung der Kommission hat damit einen ausgleichenden Effekt auf das Erreichen der nächsten Noten-Stufe. Der Effekt ist jedoch lokal, d.h. auf die mündliche Note rund um die jeweiligen Schwellenwerte begrenzt; auf die Geschlechtsunterschiede in der durchschnittlichen Gesamtnote hat die Zusammensetzung der Kommission der mündlichen Prüfung keinen statistisch signifikanten Effekt.

Differenziert man im Datensatz anhand der Indikatoren für Migrationshintergrund, so ergeben sich die folgenden zentralen Ergebnisse: Alle Indikatoren eines Migrationshintergrunds (d.h. Geburtsort im Ausland, keine deutsche Staatsangehörigkeit, Namensursprung im Ausland) weisen einen deutlichen negativen Zusammenhang mit den Noten im ersten und zweiten Examen auf. Die Effektgrößen sind in beiden Examen ähnlich stark ausgeprägt und bestehen im zweiten Examen tendenziell auch nach Kontrolle für die Note aus dem ersten Examen. Es ist wiederum kein klarer zeitlicher Trend in den Notenunterschieden zu beobachten. Der Einfluss der Migrationsindikatoren ist dabei deutlich größer als der Geschlechtereffekt. Ähnlich wie bei den Geschlechtsunterschieden zeigen sich stärkere negative Effekte bei den mündlichen als bei den schriftlichen Noten. Die Unterschiede in den mündlichen Noten bleiben auch bei Kontrolle für die Noten aus dem schriftlichen Teil und/oder dem ersten Examen bestehen. Personen mit Migrationshintergrund erreichen sowohl im ersten als auch im zweiten Examen mit deutlich geringerer Wahrscheinlichkeit ein Prädikat. Sie weisen ferner eine deutlich geringere Wahrscheinlichkeit auf, mit einer guten mündlichen Note die Schwelle für eine Gesamtnote in der nächsthöheren Noten-Stufe zu erreichen. Vergleiche zwischen Migranten der ersten und zweiten (bzw. höheren) Generation liefern vor allem für das erste Examen schwache Evidenz für positive Integrationseffekte. Gleichzeitig deuten einige Resultate auf strukturelle, gruppenspezifische Effekte (Herkunft) hin. Aufgrund der eingeschränkten Fallzahlen (zwar haben rd. 14% der beobachteten Studierenden einen Migrationshintergrund, eine Differenzierung innerhalb dieser Gruppe führt jedoch zu sehr kleinen Subgruppen) sind diesbezüglich aber keine verlässlichen Aussagen möglich.

1 EINLEITUNG

Die Verfasser des vorliegenden Berichts haben 2014 eine Studie zum staatlichen Teil der ersten juristischen Prüfung veröffentlicht, bei der sich u.a. zeigte, dass Frauen und Kandidaten, deren Name auf einen Migrationshintergrund schließen lässt, im Examen systematisch schlechter abschneiden.¹ Konfrontiert mit den Ergebnissen und den Thesen zu deren Ursachen, hat das für die Juristenausbildung und die staatlichen Prüfungen zuständige Ministerium für Justiz in NRW diese Studie in dem Bestreben in Auftrag gegeben, größere Klarheit über die Lage zu erlangen und etwaige Missstände festzustellen, damit ihnen ggf. aktiv entgegengewirkt werden kann. Diese Studie soll (1) beide juristischen Examina (2) in ganz NRW (3) in so weitem Umfang wie möglich in den Blick nehmen.

Die akademische Ausbildung an juristischen Fakultäten, die Prüfungsvorbereitung sowie die Rahmenbedingungen und Bewertungsprozesse in den Prüfungen sind auch deshalb kontinuierlich dahingehend zu überprüfen, ob sie den Ansprüchen an eine objektive und faire Ausbildung und Benotung entsprechen, weil wie in kaum einer anderen akademischen Disziplin Unterschiede in der Benotung in juristischen Examen von zentraler Bedeutung für die Karrieremöglichkeiten und den Karriereerfolg von angehenden Juristinnen und Juristen sind. So haben *Freier* und Kollegen 2015 in einer empirischen Studie nachgewiesen, dass bereits kurz nach dem Studium ein erheblicher Gehaltsunterschied (14%) zwischen Personen mit einem Prädikatsexamen und solchen, die kein Prädikat erreichen konnten, besteht.² Die Autoren führen dies — unter Rückgriff auf verschiedene Kontrollfaktoren — sowohl auf die Signalwirkung des Prädikats wie auch auf die damit verbundenen Karrieremöglichkeiten zurück.

Vor diesem Hintergrund ist insbesondere darauf zu achten, dass alle Teilnehmer mit gleichen Voraussetzungen auch die gleichen Chancen auf gute Noten erhalten, und Personen weder aufgrund ihres Geschlechts noch aufgrund ihrer Herkunft benachteiligt werden.

Das Ministerium der Justiz in NRW stellt sich dieser Herausforderung proaktiv, u.a. indem die Durchführung eines unabhängigen Forschungsprojekts finanziell und organisatorisch unterstützt wurde, dessen Ergebnisse Gegenstand dieses Berichts sind. In diesem von uns im Zeitraum 2015 – 2017 durchgeführten Projekt wurden alle elektronisch verfügbaren Examensnoten der ersten und zweiten juristischen Staatsprüfung aus NRW von 2006 bis 2016 in anonymisierter Form statistisch analysiert. Die Ziele, Vorgehensweise und Ergebnisse dieses Projekts werden nach einem kurzen Überblick über vorangegangene Befunde im vorliegenden Bericht dargestellt.

¹ *Emanuel V. Towfigh/Andreas Glöckner/Christian Traxler*, Zur Benotung in der Examensvorbereitung und im ersten Examen: Eine empirische Analyse, ZDRW 2014, 8-27.

² *Ronny Freier/Mathias Schumann/Thomas Siedler*, The earnings returns to graduating with honors — Evidence from law graduates, Labour Economics 34(C) (2015), 39–50.

2 BISHERIGE BEFUNDE

2.1 Geschlecht und Migrationshintergrund im ersten juristischen Examen

In der erwähnten vorangegangenen Studie konnten wir unter Nutzung von Daten aus der staatlichen Pflichtfachprüfung des ersten juristischen Examens im Bezirk des Oberlandesgerichtsbezirks Hamm für den Zeitraum 2007 – 2010 (N= 2.217) nachweisen, dass Frauen – gemessen an der Abiturnote – mit besseren Voraussetzungen in das Studium starten, am Ende aber mit einer um 0,29 Punkte schlechteren Examensnote abschließen.³ Bei statistischer Kontrolle u.a. für die Abiturnote betrug der Unterschied ca. 10% der Gesamtnote (0,7 Punkte). Dieser Unterschied ist dabei in der mündlichen Prüfung ausgeprägter als in der anonymisierten schriftlichen Prüfung. So bleibt ein Unterschied von 0,24 Punkten in der mündlichen Prüfung bestehen, selbst wenn man zusätzlich für die schriftliche Note kontrolliert; anders ausgedrückt: vergliche man zwei Examenskandidaten, die sich nur im Geschlecht unterscheiden, so schnitten Frauen mit gleichem Abitur und gleichen schriftlichen Noten im Vergleich zu ihren männlichen Kollegen 0,24 Punkte schlechter ab.

Ähnliche Unterschiede zeigten sich zwischen Personen, deren Name aufgrund einer onomastischen Analyse auf einen Migrationshintergrund schließen ließ, im Vergleich zu Personen mit „traditionell deutschen“ Namen. Trotz vergleichbarer Abiturleistungen zeigte sich ein Notenunterschied von 0,73 Punkten. Der Unterschied in der mündlichen Note war wiederum stärker als in der schriftlichen, und es blieb ein Unterschied von bis zu 0,43 Punkten bestehen, wenn zusätzlich für die in der schriftlichen Note bereits anonym abgefragten juristischen Kenntnisse und Fähigkeiten kontrolliert wurde.

In einer Studie von *Hinz und Röhl*, die im Laufe des vorliegenden Forschungsprojekts erschienen ist, wurde der Unterschied zwischen Personen mit einem Namen, der auf einen Migrationshintergrund hindeutet, für eine Stichprobe aus Baden-Württemberg zunächst bestätigt.⁴ Allerdings verschwand der Unterschied in der mündlichen Note bei Kontrolle für die schriftliche Note und unter Berücksichtigung weiterer Kontrollvariablen wie sozioökonomischer Status. Da sozioökonomischer Status und Migrationshintergrund üblicherweise korreliert sind, könnte dieses positivere Ergebnis allerdings auf einer Unterschätzung des reinen Effekts eines Migrationshintergrunds zurückzuführen sein. Auch sind Unterschiede zwischen Bundesländern, wie eine weitere aktuelle Studie darlegt, nicht auszuschließen.⁵ In jedem Fall wäre eine Entwarnung auf Basis der baden-württembergischen Befunde verfrüht.

Eine Überprüfung der Befunde auf Basis einer breiteren Datengrundlage, besserer Kontrollvariablen sowie unter Berücksichtigung des ersten und zweiten Examens wird mit dem Projekt, dessen Ergebnisse hier präsentiert werden, angestrebt.

³ *Towfigh/Glückner/Traxler*, Zur Benotung in der Examensvorbereitung und im ersten Examen: Eine empirische Analyse, ZDRW 2014, 8-27.

⁴ *Thomas Hinz/Hans-Christian Röhl*, Juristische Fakultäten in Baden-Württemberg: Wo studiert man am besten?, VBIBW 2016, 20-23.

⁵ *Lorenz Kähler/Uwe Engel/Franziska Ritter*, Hängen die Ergebnisse der zweiten juristischen Prüfung vom Examensort ab? Ein Vergleich von zehn Bundesländern, ZfRS 2017, 133-159.

2.2 Differentielle Geschlechterunterschiede in Staatsprüfungen für Lehrer in Frankreich

In einer Analyse von über 100.000 Benotungen einer staatlichen Prüfung, auf die bei der Auswahl fast aller französischen Lehrer der Sekundar- und Oberstufe sowie von Professoren zurückgegriffen wird, konnten *Breda* und *Hillion* Effekte einer differenziellen Geschlechterdiskriminierung nachweisen.⁶ Der Vergleich der anonymen schriftlichen Noten mit den naturgemäß nicht anonymen mündlichen Bewertungen ergab, dass die Personengruppe, die in einem Fach unterrepräsentiert war, bei Bekanntheit des Geschlechts jeweils tendenziell bessere Noten erhielt als Personen des anderen Geschlechts. So zeigt die Analyse über 11 Fächer u.a. dass Frauen Mathematik, Physik und Philosophie in der mündlichen Prüfung durchschnittlich 10 Prozent Rangplätze besser eingeschätzt wurden als in der anonymen schriftlichen Prüfung. Dies indiziert eine *positive* Diskriminierung von Frauen in Fächern, in denen weniger Frauen als Männer tätig sind. Der umgekehrte Effekt ließ sich für die Fächer Literatur und Fremdsprachen nachweisen, in denen Männer in den mündlichen Prüfungen um 3 bis 5 Prozent besser abschnitten als in den anonymen schriftlichen Prüfungen.

Es ist natürlich unklar, ob diese Befunde über Fächergrenzen, Länder und Prüfungsformate übertragbar sind; auch der Umstand, dass es sich um eine staatliche Prüfung handelt, ist hier nicht von weiterer Relevanz. Die Studie untersucht aber ebenfalls Geschlechtereffekt in Prüfungsverfahren und ist insofern thematisch einschlägig. Sie ist auch in verschiedener Hinsicht bemerkenswert: So zeigt sie (a), dass auch unter bestmöglicher statistischer Kontrolle eine Ungleichbehandlung von Geschlechtern in staatlichen Examina nachgewiesen werden kann; dass es sich (b) nicht immer um eine Benachteiligung von Frauen handeln muss, sondern dass (c) das übergeordnete Ziel der Gleichbehandlung paradoxerweise vielleicht gerade dadurch verletzt wird, dass Prüfer — wahrscheinlich unbewusst und wohl vor allem bei der Benotung in mündlichen Prüfungen — von ihnen wahrgenommene Nachteile auszugleichen versuchen.⁷

2.3 Zusammensetzung der Prüfungskommission

Ein Ansatzpunkt zur Verringerung einer potenziellen Ungleichbehandlung von Prüflingen kann darin bestehen, die Prüfungskommissionen ausgeglichener bzw. repräsentativer zu besetzen. Speziell sind Frauen in Prüfungskommissionen üblicherweise unterrepräsentiert, und selbiges gilt in einem noch deutlicheren Ausmaß für Personen mit Migrationshintergrund. Eine diversere Besetzung der Prüfungskommissionen könnte zum einen dazu beitragen, dass einer potenziellen Diskriminierung durch diese Prüfer direkt entgegengewirkt wird. Darüber hinaus — und vermutlich praktisch bedeutsamer — ist ein positiver Effekt auf die Prüflinge zu erwarten. Durch das Vorhandensein positiver Rollenmodelle ist es wahrscheinlich, dass die empirisch nachgewiesenen negativen Effekte von Stereotypen — speziell der wahrgenommenen Bedrohung durch Stereotype⁸ — reduziert werden. Eine aktuelle Studie, in der die Beteiligung von Frauen in Berufungskommissionen an einer deutschen Hochschule untersucht wurde, konnte allerdings keinen generellen positiven Effekt von einem höheren Frauenanteil auf den Frauenanteil unter den Neuberufungen bzw. Berufungsvorschlägen finden.⁹

⁶ *Thomas Breda/Melina Hillion*, Teaching Accreditation Exams Reveal Grading Biases Favor Women in Male-Dominated Disciplines in France, *Science* 353 (2016), 474-478.

⁷ Vgl. dazu auch *Andreas Glöckner/Emanuel V. Towfigh*, Messgenauigkeit und Fairness in Staatsprüfungen: Aktuelle Studien zeigen Gruppen-Unterschiede in juristischen Examina auf, *AnwBl* 2016, 706-709.

⁸ Bspw. *Steven J. Spencer/Claude M Steele/Diane M Quinn*, Stereotype Threat and Women's Math Performance, *Journal of Experimental Social Psychology* 35 (1999), 4-28.

⁹ *Katrin Auspurg/Thomas Hinz/Andreas Schneck* Wishful Thinking — Verbessern mehr Frauen in Berufungskommissionen die Berufungschancen?, *Forschung & Lehre* 2017, 770-774.

3 ZIELSTELLUNG UND ALLGEMEINE VORGEHENSWEISE

Zusammenfassend bleibt auf Basis der oben beschriebenen Befunde festzuhalten, dass die potenziellen Unterschiede zwischen Männern und Frauen sowie Personen mit bzw. ohne Migrationshintergrund in staatlichen Examina durch sehr unterschiedliche Ursachen hervorgerufen werden könnten, die im Ausbildungssystem (bspw. unzureichende Motivation von Frauen in Vorlesungen zur aktiven Teilnahme), im Prüfling selbst (bspw. empfundene Angst / Bedrohung durch Stereotype), in der Organisation der Prüfung (bspw. Zusammensetzung der Kommission) oder auch den Prüfern (bspw. bewusste oder unbewusste Diskriminierung) zu verorten sind. Im aktuellen Projekt wird das Vorhandensein der Unterschiede mit besseren und umfangreicheren Daten untersucht.

Speziell werden in diesem Projekt Examensnoten im ersten und zweiten juristischen Examen auf

- Unterschiede zwischen Männern und Frauen
- Unterschiede zwischen Personen mit bzw. ohne Migrationshintergrund
- eine kontinuierliche Verteilung von Noten über die Notenskala sowie
- Unterschiede zwischen Rechtsgebieten und Universitäten

untersucht.

Die Analysen beleuchten nur teilweise mögliche Ursachen für Unterschiede. Eine weiterführende Untersuchung von unterschiedlichen, kausalen Prozessen, die schließlich zu systematischen Notenunterschieden zwischen verschiedenen Gruppen führen, würde eine andere Datenlage erfordern und den Rahmen dieses Berichts sprengen. Die Analysen erfolgen vorwiegend unter Nutzung statistischer Regressionsmodelle, in denen überprüft wird, ob unter Berücksichtigung verschiedener Kontrollvariablen der Effekt eines Faktors wie bspw. Geschlecht oder Migrationshintergrund auf die Bewertung „statistisch bedeutsam“ (d.h. statistisch signifikant unterschiedlich von einem Null-Effekt) bleibt. Das Signifikanzniveau gibt dabei an, wie (un)wahrscheinlich es ist, dass ein solch starker Einfluss beobachtet werden sollte, obwohl tatsächlich kein Unterschied besteht. Nach einer verbreiteten Konvention spricht man bei Wahrscheinlichkeiten von unter 1%, unter 5% bzw. unter 10% von einem hoch-signifikanten, einem signifikanten respektive einem schwach-signifikanten Effekt.

Die Nutzung von Kontrollvariablen erlaubt es dabei konzeptuell „statistische Zwillinge“ zu erzeugen — also Personen die in gemessenen Dimensionen (wie bspw. Alter oder Abiturnote) vergleichbar sind, weil der partielle Einfluss dieser Variablen statistisch „kontrolliert“ wird.

4 FORSCHUNGSFRAGEN

Die folgenden Forschungsfragen (F1 – F7) werden in der aktuellen Studie bearbeitet:

Bleiben die in der oben berichteten Vorgängerstudie (s.o. bei Fn. 1) beobachteten Unterschiede für Geschlecht und Migrationshintergrund bestehen...

- F1: ... bei einer differenzierteren Betrachtung des ersten Examen?
- F2: ... im zweiten Examen...?
 - F2a: ... bei Kontrolle für erstes Examen?
 - F2b: ... für die mündliche Note bei Kontrolle für die schriftliche Note?

Von speziellem Interesse ist die Analyse von Diskontinuitäten in der Notenverteilung rund um die Noten-Stufen (mangelhaft/ausreichend, ausreichend/befriedigend, befriedigend/vollbefriedigend, vollbefriedigend/gut, gut/sehr gut) und speziell die Analyse von Unterschieden in der Wahrscheinlichkeit, die nächste Stufe zu erreichen, zwischen Gruppen sowohl im ersten und im zweiten Examen:

- F3: Besteht für Personen unterschiedlicher Personengruppen (insbes. Frauen und Migranten) mit sonst gleichen Voraussetzungen die gleiche Wahrscheinlichkeit, die nächste Noten-Stufe zu erreichen?

Darüber hinaus wird analysiert:

- F4: Verringern sich potenzielle Gruppen-Unterschiede über die Abschlussjahrgänge von 2006 bis 2016?
- F5: Verändern sich potenzielle gruppenbezogene Leistungsunterschiede (insbes. Frauen und Migranten) zwischen dem ersten und zweiten Examen?

Eine zentrale Analyse betrifft den Effekt der Zusammensetzung von Prüfungskommissionen:

- F6: Verringert sich ein potenzieller Geschlechtereffekte bei Kommissionen mit Beteiligung von Frauen?

Aufgrund der Datenverfügbarkeit sowie der geringen Anzahl von Personen mit Migrationshintergrund in den Kommissionen kann keine entsprechende Analyse für diesen Faktor durchgeführt werden.

5 BESCHREIBUNG DES DATENSATZES

Die Forschungsfragen wurden beantwortet durch Analyse eines vom Landesjustizprüfungsamt NRW bereitgestellten Datensatzes aller elektronisch erfassten Noten des zweiten juristischen Staatsexamens der Abschlussjahrgänge 2006 bis 2016 in NRW. Der Datensatz deckt dabei mehrere (auch erfolglose) Prüfungsversuche ab und beinhaltet – zumindest teilweise, nämlich dann, wenn beide Examina in NRW abgelegt wurden – die Noten des ersten Staatsexamens.

Zur Gewährleistung eines maximalen Schutzes der Daten wurden die Daten im Landesjustizprüfungsamt zusammengeführt, aufbereitet und danach in anonymisierter Form übergeben. Die Namen wurden (getrennt von den übrigen Daten) vor der Anonymisierung im Landesjustizprüfungsamt einer onomastischen Analyse durch die Firma Humpert & Schneiderheinze Sozial- und Umfrageforschung unterzogen, mit deren Hilfe eine differenzierte Kodierung der Vor- und Nachnamen nach deren Herkunftsregionen erfolgte.

Insgesamt umfasst der Datensatz Noten von 19.883 Personen, die 26.342 Prüfungsversuche im zweiten Examen durchliefen. 5.208 Prüfungsversuche wurden nicht bestanden. Die Versuche enthalten reguläre Versuche, Wiederholungsversuche sowie Notenverbesserungsversuche.

Die zentralen empirischen Analysen für das zweite Examen fokussieren sich im Folgenden auf den für die Forschungsfrage „relevanten“ Versuch, d.h. entweder den direkt bestandenen Versuch oder – falls dadurch eine Notenverbesserung erzielt wurde – den relevanten Wiederholungs- bzw. Verbesserungsversuch. Dadurch wird sichergestellt, dass für jede Person nur ein Versuch analysiert und eine Doppelbetrachtung bzw. eine höhere Gewichtung von Wiederholungs-/Verbesserungsversuchen vermieden wird. Die Effekte sind jedoch robust (und tendenziell stärker ausgeprägt), wenn Wiederholungen und freiwillige Verbesserungsversuche ausgeschlossen werden, d.h., wenn nur die jeweils ersten Versuche untersucht werden.

Die zentralen Analysen basieren damit auf maximal 17.971 beobachteten Gesamtnoten (von ebenso vielen Personen; siehe *Tabelle 1 Panel B*). Für die schriftlichen Teilnoten liegen maximal 18.958 Beobachtungen vor.¹⁰ Teilnoten für die Rechtsgebiete Öffentliches Recht, Strafrecht und Zivilrecht (im zweiten Examen) liegen für maximal 9.696 Fälle vor.

Für die Analysen der Noten aus dem ersten Examen stehen für 10.042 Kandidatinnen und Kandidaten – nahezu allen Personen, die zwischen dem Wintersemester 2010 und dem Sommersemester 2016 einen Prüfungsversuch im zweiten Examen unternommen haben – die Gesamtnoten der ersten Prüfung zur Verfügung (*Tabelle 1 Panel A*). Separate Daten für den staatlichen und universitären Teil des ersten Examens liegen für 4.390 bzw. 4.354 Personen vor.

Die basalen Informationen zu den wichtigsten Kontroll- und Vorhersagevariablen sind in Panel *Tabelle 1 Panel C* zusammengefasst. Für nahezu alle berücksichtigten Personen liegen Informationen zu Geschlecht und Alter der Prüflinge vor. 52,9% der Geprüften waren weiblich, das mittlere Alter betrug 30 Jahre. Die Abiturnote liegt nur für einen Teil der Beobachtungen vor (was die statistische Aussagefähigkeit hingegen kaum beeinträchtigt); ihr Durchschnitt beträgt 2,26.

¹⁰ Die im Vergleich zu den Gesamtnoten höhere Beobachtungszahl ist damit zu erklären, dass hier einigen Fälle (vor allem Fälle mit Noten unterhalb der 3,5-Punkte-Schwelle) im Beobachtungszeitraum nicht zur mündlichen Prüfung angetreten sind, so dass keine Gesamtnote vorliegt.

Für alle Beobachtungen, bei denen Information über die mündlichen Prüfungsteile im zweiten Examen vorliegen wird auch verzeichnet, ob zumindest eine Frau Teil der Prüfungskommission war; in 35,4% der mündlichen Prüfungen war dies der Fall.

Für 9.784 Beobachtungen liegen zusätzlich Informationen zur Staatsangehörigkeit (2,8% haben eine nicht-deutsche Staatsangehörigkeit), zum Geburtsort (9,1% wurden im Ausland geboren) sowie der aus der onomastischen Analyse abgeleitete Indikator für einen Migrationshintergrund vor. 13,8% der Prüflinge haben demnach einen Namen der auf eine nicht-deutschen Ursprung schließen lässt. Der Vergleich mit den zwei „härteren“ Indikatoren für einen Migrationshintergrund (Geburtsort und Staatsangehörigkeit) zeigt dabei, dass der Onomastik-Indikator eine sehr hohe Treffgenauigkeit aufweist.

Tabelle 1: Deskriptive Statistiken.

Variable	Mittelwert	St.Abw.	Min	Max	N
A. Erstes Examen					
Gesamtnote erste Prüfung	7,772	1,985	4,000	15,940	10.042
• staatliche Pflichtfachprüfung (70%)	7,315	1,946	3,700	14,200	4.390
• universitäre Prüfung (30%)	9,146	2,355	0,400	16,833	4.351
B. Zweites Examen					
Gesamtnote zweites Staatsexamen	7,495	1,993	3,270	15,020	17.971
Note schriftliche Prüfungsteile	6,169	2,105	0,250	14,625	18.958
Note mündliche Prüfungsteile	9,200	2,439	2,000	17,750	17.970
Note Zivilrecht	6,164	2,389	0,500	15,250	9.696
Note Strafrecht	5,811	2,465	0,000	15,000	9.696
Note öffentliches Recht	6,077	2,652	0,000	16,500	9.696
C. Kontrollvariablen					
Abiturnote	2,263	0,605	0,900	3,900	4.592
Alter	30,026	2,914	21,052	67,584	19.865
Frauen	0,529	0,499	0	1	19.865
Staatsangehörigkeit (nicht-deutsch)	0,028	0,166	0	1	9.784
Geburtsort (nicht-deutsch)	0,091	0,290	0	1	9.784
Onomastik (nicht-deutsch)	0,138	0,345	0	1	9.784
Migrationsindikatoren: 1. Generation	0,061	0,240	0	1	9.784
Migrationsindikatoren: \geq 2. Generation	0,079	0,270	0	1	9.784
Frau in Prüfungskommission	0,354	0,478	0	1	17.970

Bei Berücksichtigung aller drei Indikatoren kann innerhalb der Gruppe von Personen mit Migrationshintergrund weiter differenziert werden: Die Analyse trennt hier zwischen jenen *Migranten* (erster Generation), die im Ausland geboren wurden und über keine deutsche Staatsbürgerschaft verfügen (etwa 6%) und jenen, die zwar in Deutschland geboren wurden, aber entweder einen Namen ausländischen Ursprungs oder eine nicht-deutsche Staatsbürgerschaft haben. Etwa 8% verfügen über einen solchen Migrationshintergrund der zweiten oder späteren Generation.¹¹

¹¹ Die Summe aus den so gebildeten Indikatoren für Migrationshintergrund der ersten oder höheren Generationen ergibt *nicht* den ausschließlich Onomastik-basierten Indikator. Diese liegt daran, dass hier bei der Definition der generationspezifischen Migrationsindikatoren neben den Onomastik Werten auch weitere Variablen genutzt werden, sodass hier eine marginal breitere Definition entsteht.

6 ANALYSEN ZU EFFEKTEN DES GESCHLECHTS

6.1 Effekte des Geschlechts im ersten Examen

6.1.1 Unterschiede in der Gesamtnote

Die Analyse der Abschlussnote für die erste Prüfung liefert ein klares Ergebnis: Frauen erzielen durchschnittlich um 0,29 Notenpunkte (entspricht 3,6%) schlechtere Noten als Männer ($\text{Note}_{\text{Männer}} = 7,93$; $\text{Note}_{\text{Frauen}} = 7,64$; hoch-signifikanter Unterschied mit $p < 0,001$; siehe Tabelle 2, Modell 1). Der Unterschied steigt auf 0,38 Punkte, wenn für die Abiturnote kontrolliert wird (Modell 2).

Ein ähnlicher Geschlechtseffekt tritt auch im staatlichen Teil der ersten Prüfung („staatliche Pflichtfachprüfung“) auf: Basierend auf einer kleineren Fallzahl zeigt sich hier, dass Frauen um 0,23 Notenpunkte (3,0%) schlechtere Noten erhalten als Männer ($\text{Note}_{\text{Männer}} = 7,44$; $\text{Note}_{\text{Frauen}} = 7,21$; $p < 0,001$; Modell 4). Wie zuvor steigt dieser Unterschied, wenn für die Abiturnote kontrolliert wird (Unterschied von 0,51 Punkten; Modell 5).

Im universitären Teil der Prüfungen fallen die Geschlechtsunterschiede qualitativ anders aus. Kontrolliert man nicht für das Abiturergebnis, so erhalten Frauen hier sogar um 0,15 Punkte *bessere* Noten als Männer ($p < 0,05$; Modell 6). Dieser Unterschied kehrt sich jedoch wieder um, wenn für die Abiturnote kontrolliert wird: Nach Berücksichtigung dieser Variable schneiden Frauen im universitären Prüfungsteil durchschnittlich 0,16 Punkte schlechter ab als Männer ($p < 0,05$; Modell 7). Unter Berücksichtigung der Abiturnote finden wir also auch im universitären Teil überzufällig schlechtere Noten von Frauen – wenn auch in deutlich geringerem Ausmaß – als in der staatlichen Pflichtfachprüfung.

Basierend auf diesen ersten Befunden kehren wir wieder zur Analyse der Gesamtnote im ersten Examen zurück. In Modell (3) zeigt sich nun, dass – bei Kontrolle für die Note im staatlichen Teil sowie für die Abiturnote – Frauen sogar marginal (0,03 Notenpunkte, $p < 0,10$) bessere Gesamtnoten als Männer erzielen. Das heißt, dass die Geschlechterunterschiede in der Gesamtnote vorwiegend in der staatlichen Pflichtfachprüfung entstehen und im universitären Teil nicht weiter verstärkt, sondern sogar marginal abgeschwächt werden.

Ergebnis F1 (Geschlecht im ersten Examen): Auch bei einer umfangreichen Datengrundlage und bei differenzierter Betrachtung zeigt sich im ersten Examen ein substantieller negativer Effekt in der Gesamtbenotung von Frauen im Vergleich zu Männern. Der Noten-Unterschied wird hauptsächlich durch den staatlichen Teil getrieben.

Tabelle 2: Analyse von Geschlechterunterschieden in der ersten Prüfung.

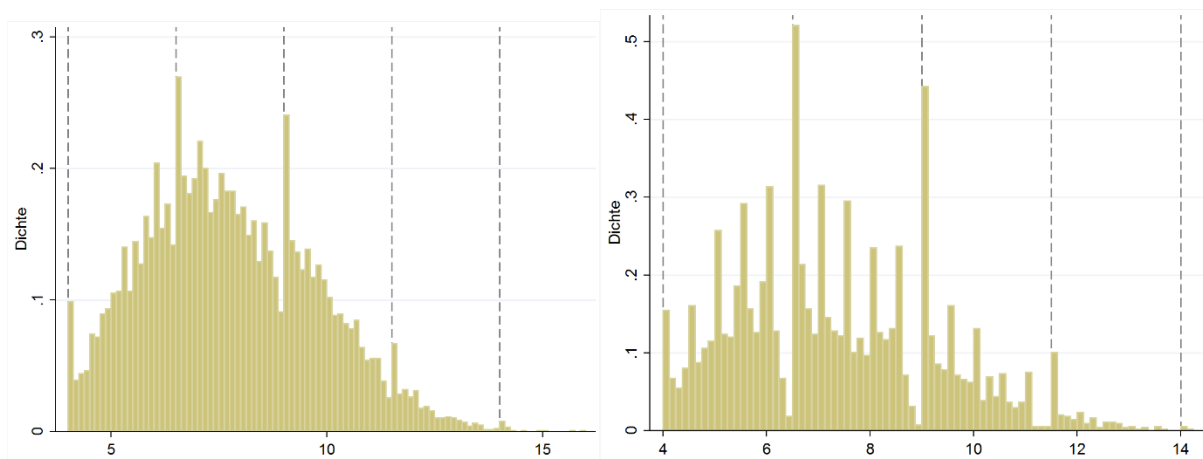
Variable	(1) Gesamtnote (erste Prüfung)	(2)	(3)	(4) staatliche Pflichtfachprüfung	(5)	(6) universitäre Prüfung	(7)
Frauen	-0,287*** [0,000]	-0,377*** [0,000]	0,031* [0,090]	-0,225*** [0,000]	-0,510*** [0,000]	0,147** [0,044]	-0,161** [0,018]
Abiturnote		-1,568*** [0,000]	-0,259*** [0,000]		-1,554*** [0,000]		-1,675*** [0,000]
Note staatl. Pflichtfachpr			0,857*** [0,000]				
Konstante	7,927*** [0,000]	11,590*** [0,000]	2,164*** [0,000]	7,438*** [0,000]	11,112*** [0,000]	9,058*** [0,000]	13,017*** [0,000]
N	10.042	4.596	4.252	4.390	4.254	4.354	4.252
R ²	0,005	0,261	0,898	0,003	0,231	0,001	0,179

Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1

6.1.2 Unterschiede in der Wahrscheinlichkeit, relevante Notenschwellen zu erreichen

Zur Analyse des auf das erste Examen bezogenen Teils der vierten Forschungsfrage betrachten wir Diskontinuitäten in der Notenverteilung. Abbildung 1 zeigt die Notenverteilung für die Gesamtnoten (links) bzw. für den staatlichen Prüfungsteil (rechts). Die Abbildungen zeigen deutlich, dass die Notenverteilungen keinem kontinuierlichen Muster folgen. Ähnlich wie in unserer vorangegangenen Studie beobachteten wir eine übermäßige Häufung von Noten an bzw. knapp oberhalb der durch vertikale Linien markierten Notenstufen. Diese „zusätzliche“ Verteilungsmasse fehlt in den Bereichen knapp unter den Notenstufen. Dieses Muster kann – unter der Annahme eines eigentlich kontinuierlichen Leistungsspektrums – nur damit erklärt werden, dass das Wissen über die (schriftlichen) Vornoten von den Prüferinnen und Prüfern in der mündlichen Prüfung genutzt wird, um „eindeutigere“ Gesamtnoten herzustellen, die nicht knapp unter Notenstufen liegen. Einige Prüflinge scheinen also so bewertet zu werden, dass sie die Notenstufe gerade erreichen.

Abbildung 1: Notenverteilung in der ersten Prüfung (links: Gesamtnote; rechts: staatlicher Teil).



Diesen vornotenorientierten Benotungsprozess, der systematisch die jeweiligen Vornoten berücksichtigt, werden wir im Rahmen unserer Analyse der Notenverteilung im zweiten Examen – bei der wir ein ähnliches Verteilungsmuster feststellen werden – näher untersuchen. Nachdem uns im untersuchten Datensatz hinsichtlich der ersten Prüfung keine Differenzierung der Note nach den mündlichen und schriftlichen Teilen vorliegt, betrachten wir im Folgenden nur, ob es einen grundsätzlichen Geschlechterunterschied im Erreichen der relevanten Notenschwellen gibt.

Die Analyse der Abschlussnoten der ersten Prüfung zeigt, dass Frauen eine um 5,3%-Punkte niedrigere Rate an „besser als 9,0“-Noten haben (M: 31,5% F: 26,2%, $p < 0,001$; siehe Tabelle 3, Modell 1). Relativ zu Männern haben Frauen also eine etwa 17% geringere Rate an vollbefriedigenden oder besseren Noten. Bei Kontrolle für die Abiturnote steigt der Unterschied auf 7,3%-Punkte bzw. 23% an (Modell 2).¹²

Wird für die Note aus dem staatlichen Prüfungsteil kontrolliert, verschwindet jedoch der Geschlechterunterschied (siehe Modell 3). Das bedeutet, dass der Geschlechtsunterschied v.a. vom staatlichen Teil des ersten Examins getrieben wird. Diese Aussage wird durch die Resultate in Modellen 4 bzw. 5 unterstützt, die aufzeigen, dass Frauen eine um 2,7 bzw. 7,3%-Punkte geringere Wahrscheinlichkeit haben, in diesem Teil der Prüfung eine Note von 9,0 (oder mehr) zu erreichen.

Tabelle 3. Wahrscheinlichkeit, die Notenschwelle von 9,0 zu erreichen (erstes Examen).

	(1)	(2)	(3)	(4)	(5)
Variable:	Indikator: Note \geq 9,0 (Gesamtnote)			Note \geq 9,0 (staatl. Pflichtfachpr.)	
Frauen	-0,053*** [0,000]	-0,070*** [0,000]	0,013 [0,156]	-0,027** [0,036]	-0,073*** [0,000]
Abiturnote		-0,310*** [0,000]	-0,054*** [0,000]		-0,269*** [0,000]
Note staatl. Pflichtfachpr.			0,168*** [0,000]		
Konstante	0,315*** [0,000]	1,014*** [0,000]	-0,835*** [0,000]	0,253*** [0,000]	0,887*** [0,000]
N	10.042	4.596	4.252	4.390	4.254
R ²	0,003	0,174	0,589	0,001	0,144

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Ähnliche Effekte zeigen sich bei den Notenstufen 6,5 (befriedigend) und 11,5 (gut). Frauen haben eine um 4,5%-Punkte niedrigere Rate an „befriedigenden“ (oder besseren) Noten (M: 74.0% F: 69.6%; siehe Appendix A, Tabelle A1, Modell 1) und eine um 2,0%-Punkte niedrigere Rate an Noten besser als 11,5 (M: 5.2% F: 3.2%; Appendix A, Tabelle A2, Modell 1).

¹² Die hier und im Folgenden präsentierten Resultate für binäre abhängige Variablen wurden mit dem linearen Wahrscheinlichkeitsmodell (*linear probability model, LPM*) geschätzt. Dies vereinfacht die Interpretation der Koeffizienten für Dummy-Indikatoren und Interaktionsterme. Schätzungen basierend auf nicht-linearen Modellen (*Probit*) liefern qualitativ und quantitativ vergleichbare Ergebnisse.

6.2 Effekte des Geschlechts im zweiten Staatsexamen

6.2.1 Unterschiede in der Gesamtnote

Die Analyse der Gesamtnoten im zweiten Staatsexamen ergibt, dass Frauen um 0,14 Notenpunkte schlechtere Noten (1,9%) erzielen als Männer (N = 17.971; $Note_{\text{Männer}} = 7,57$; $Note_{\text{Frauen}} = 7,43$; $p < 0,001$; siehe Tabelle 4, Modell 1). Dieser Notenunterschied erhöht sich nach Kontrolle für Alter und Abschlusszeitpunkt-spezifische Effekte¹³ auf 0,44 Notenpunkte (Modell 2). Bei zusätzlicher Kontrolle für die Note im ersten Examen (Modell 3) sowie die Abiturnote (Modell 4) sinken die Unterschiede deutlich, bleiben aber auf niedrigerem Signifikanzniveau bestehen. Bei der Einschätzung der reduzierten Signifikanzlevels ist zu berücksichtigen, dass sich auch die Größe der Stichprobe in den beiden letzten Modellen deutlich reduziert.

Ergebnis F2 und F2a (Geschlecht): Es zeigen sich signifikante Geschlechterunterschiede in der Gesamtnote im zweiten Examen. Diese Unterschiede bleiben – in abgeschwächter Form – auch nach Kontrolle für die Note aus dem ersten Examen sowie der Abiturnote qualitativ bestehen.

Tabelle 4: Geschlechterunterschiede im zweiten Staatsexamen: Gesamtnote.

	(1)	(2)	(3)	(4)
Variable:	Gesamtnote (zweites Examen)			
Frauen	-0,143*** [0,000]	-0,446*** [0,000]	-0,060** [0,033]	-0,070* [0,074]
Gesamtnote erstes Examen			0,729*** [0,000]	0,695*** [0,000]
Abiturnote				-0,257*** [0,000]
Konstante	7,571*** [0,000]	23,569*** [0,000]	9,289*** [0,000]	7,515*** [0,000]
Weitere Kontrollvariablen	Nein	Ja	Ja	Ja
N	17.971	17.971	9.086	4.251
R ²	0,001	0,095	0,582	0,597

Modelle (2) – (4) kontrollieren für das Alter zum Prüfungszeitpunkt (linearer und quadratischer Term) sowie für Abschlussmonat-spezifische Effekte. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

¹³ Konkret kontrollieren die Modelle 2 bis 4 für das Alter im Zeitpunkt der Prüfung (um nicht-lineare Effekte abzubilden, wird hier neben einem linearen auch ein quadratischer Term berücksichtigt) sowie (über einen Vektor an Dummy-Variablen) für die Durchschnittswerte aller relevanten Prüfungsmonate. Das bedeutet, dass zeitlich-spezifische Effekte eliminiert werden (z.B. eine besonders schwierige Klausuraufgabe).

Eine Analyse getrennt nach Herkunftsuniversität (siehe Appendix A, Tabelle A3) zeigt, dass die Geschlechtsunterschiede in den Noten zwischen Referendarinnen und Referendaren aus unterschiedlichen Universitäten deutlich variieren. Deutlich stärker (und auch statistisch signifikant stärker als die Referenzgruppe Köln) sind die Geschlechtsunterschiede an der Universität Bonn ausgeprägt; deutlich weniger stark (aber statistisch nicht signifikant unterschiedlich zur Referenzgruppe) sind die Geschlechtsunterschiede in Münster.

6.2.2 Unterschiede in der Wahrscheinlichkeit, relevante Notenschwellen zu erreichen

Auch im zweiten Examen zeigen sich deutliche Häufungen von Noten über den Notenschwellen und fehlender Beobachtungsmasse unter den jeweiligen Schwellenwerten (Abbildung 2). Im Vergleich zur Notenverteilung im ersten Examen ist diese Tendenz jedoch deutlich stärker ausgeprägt.

Eine Analyse der Wahrscheinlichkeit der Erreichung der Notenstufe 9,0 zeigt dass Frauen eine 3,6%-Punkte geringere Rate an Prädikatsexamen als Männer aufweisen (M: 29,3%; F: 25,7%, $p < 0,001$; siehe Tabelle 5, Modell 1). Verwendet man die Rate an Prädikatsexamen bei Männern als Basis, so entspricht das einem Geschlechterunterschied von über 12% ($= 3,6/29,3$). Dieser Unterschied steigt bei Kontrolle für Alter und Abschlusszeitpunkt deutlich an (Modell 2). Wird jedoch für die Gesamtnote im ersten Examen kontrolliert – in der es ja ebenfalls einen deutlichen Geschlechterunterschied gibt – so zeigt sich kein *zusätzlicher* Unterschied in der hier untersuchten Wahrscheinlichkeit (Modell 3). Dieser Befund ändert sich nicht, wenn zusätzlich auch für die Abiturnote kontrolliert wird (Modell 4).

Abbildung 2: Notenverteilung im zweiten Examen (Gesamtnote).

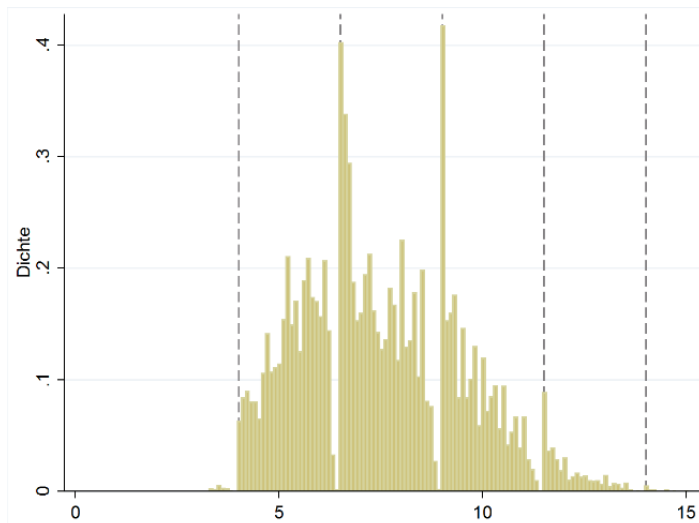


Tabelle 5. Wahrscheinlichkeit zur Erreichung der Notenschwellen von 9,0 (zweites Examen)

	(1)	(2)	(3)	(4)
Variable:	Indikator: Gesamtnote \geq 9,0 (2. Examen)			
Frauen	-0,036*** [0,000]	-0,083*** [0,000]	-0,010 [0,194]	-0,018 [0,113]
Gesamtnote erstes Examen			0,141*** [0,000]	0,136*** [0,000]
Abiturnote				-0,060*** [0,000]
Konstante	0,293*** [0,000]	2,871*** [0,000]	0,561** [0,017]	1,057** [0,017]
Weitere Kontroll- variablen	Nein	Ja	Ja	Ja
N	17.971	17.971	9.086	4.251
R ²	0,002	0,051	0,390	0,401

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Modelle (2) – (4) kontrollieren für das Alter zum Prüfungszeitpunkt (linearer und quadratischer Term) sowie für Abschlussmonats-spezifische Effekte. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Werden also ein durchschnittlicher Jurist und eine durchschnittliche Juristin gleichen Alters und mit gleichen Noten im ersten Examen miteinander verglichen, so ist *kein* statistisch signifikanter Unterschied hinsichtlich der Wahrscheinlichkeit festzustellen, mit der beide die Notenschwelle 9,0 erreichen. Das Resultat legt also nahe, dass der Geschlechtsunterschied – der in der Rate der Prädikatsexamen im zweiten Examen klar zu beobachten ist (siehe Modell 1 und 2) – im Wesentlichen über die Notenunterschiede aus dem ersten Examen erklärt wird, und darüber hinaus *durchschnittlich* kein *weiterer* Unterschied zu beobachten ist (Modell 3 und 4). In weiteren Analyseschritten werden wir über diese Betrachtung von durchschnittlichen Unterschieden hinausgehen (siehe Abschnitt 6.2.5).

Bei Betrachtung der Wahrscheinlichkeiten, Noten über den Schwellen 6,5 Punkten (Appendix A, Tabelle A4) bzw. 11,5 Punkten (Tabelle A5) zu erzielen zeigt sich, dass die Geschlechtsunterschiede vorwiegend am oberen Ende der Notenverteilung auftreten. Den Schwellenwert 11,5 Notenpunkte überspringen etwa 1,5%-Punkte weniger Frauen als die 4,7% der Männer, die diese Hürde nehmen. Das entspricht einem relativen Unterschied von nahezu 32%. Ähnlich wie in Tabelle 5 zeigt sich aber auch hier, dass dieser Geschlechtsunterschied durch die unterschiedlichen Vornoten erklärt werden kann.

Ergänzung zu Ergebnissen F2 und F2a (Geschlecht): Der abgeschwächte negative Effekt für Frauen im Mittelwert der Gesamtnote und das Verschwinden des Effekts bei Betrachtung der Notenschwellen bei Kontrolle für die Note aus dem ersten Examen deuten darauf hin, dass sich der Geschlechterunterschied zwischen den Examen kaum verstärkt.

6.2.3 Differenzierte Betrachtung der Geschlechtsunterschiede in mündlicher und schriftlicher Note im zweiten Examen

Separate Analysen der (durchschnittlichen) mündlichen und schriftlichen Noten zeigen, dass Geschlechterunterschiede in beiden Prüfungsteilen auftreten. Der Unterschied beträgt 0,116 Notenpunkte (1,9%) in den schriftlichen Noten (Appendix A, Tabelle A6, Modell 1) und 0,225 Notenpunkte (2,4%) in den mündlichen Noten (Appendix A, Tabelle A7, Modell 1). Bei Kontrolle für das Alter der Prüflinge bzw. für zeitspezifische Faktoren verstärken sich diese Unterschiede deutlich (Tabellen A6 und A7, Modell 2). Diese Befunde sind insofern bemerkenswert, als das Geschlecht der Referendarinnen und Referendare in der schriftlichen Prüfung nicht direkt beobachtbar sein sollte.

Wird nun weiter für Vornoten (die Abiturnote und/oder die Noten aus dem ersten Examen) kontrolliert, so ist – konsistent mit den zuvor besprochenen Ergebnissen – kein signifikanter Geschlechtsunterschied bei den schriftlichen Noten festzustellen (Tabelle A6, Modelle 3 und 4). Dies gilt jedoch nicht für die mündlichen Teilnoten.

In den mündlichen Prüfungsteilen finden wir – selbst nach Kontrolle für die Vornoten aus den schriftlichen Teilen, aus dem ersten Examen oder aus dem Abitur – einen (gleichsam zusätzlichen) hochsignifikanten Geschlechtsunterschied (Tabelle A7, Modelle 4 bis 6): Vergleicht man also zwei „gleich gute“ Rechtsreferendarinnen und Rechtsreferendare – gemessen anhand einer identischen Note im schriftlichen Prüfungsteil des zweiten Examens oder im ersten Examen – so erzielt eine Frau durchschnittlich eine um etwa 0,2 Notenpunkte schlechtere Note in den mündlichen Prüfungsteilen.

Ergebnis F2b (Geschlecht): Frauen erzielen in den mündlichen Prüfungen im zweiten Examen schlechtere Noten als Männer auch bei Kontrolle für die anonyme schriftliche Note und auch bei zusätzlicher Kontrolle für Abitur und Note im ersten Examen. Generell sind die Geschlechterunterschiede in den mündlichen Noten größer und robuster als in den schriftlichen Noten.

Eine vertiefende Analyse der schriftlichen Teilbereiche zeigt weiterhin, dass die Geschlechtsunterschiede in den schriftlichen Prüfungsteilen vor allem im Bereich Zivilrecht konzentriert sind (Appendix A, Tabelle A8, Modell 1). In dieser Teildisziplin bleibt der Geschlechtsunterschied selbst nach Kontrolle für die Gesamtnote im ersten Examen (in der Frauen signifikant schlechter abschneiden als Männer) bestehen (Tabelle A8, Modell 2).¹⁴ In keinem der weiteren beiden Rechtsgebiete – d.h. weder in den Prüfungen im Strafrecht noch in denen zum öffentlichen Recht – zeigt sich ein Geschlechtsunterschied. Dieses Null-Resultat ist unabhängig davon, ob man für Vornoten kontrolliert (Tabelle A8, Modell 4 – 9).

¹⁴ Der Geschlechterunterschied in den zivilrechtlichen Teilnoten verschwindet zwar nach Kontrolle für die Abiturnoten, dieses Resultat ist aber vorwiegend dadurch getrieben, dass das Analyse-Sample (durch die eingeschränkte Verfügbarkeit der Abiturnote) um etwa 50% eingeschränkt wird (Tabelle A8, Modell 3).

Ein weiteres Ergebnis zeigt sich hinsichtlich der Beteiligung von Frauen in der Prüfungskommission: Ob zumindest eine weibliche Prüferin Teil der Kommission ist, hat *keinen* Einfluss auf die durchschnittliche Note im mündlichen Prüfungsteil – weder für die Note von weiblichen noch für jene von männlichen Studierenden (Tabelle A7 – Modelle 3 bis 6). Das bedeutet, dass der Geschlechtsunterschied in der *durchschnittlichen* mündlichen Note nicht durch die Zusammensetzung der Prüfungskommission beeinflusst zu werden scheint. In dem in Abschnitt 6.2.5 dargestellten Analyseschritt werden wir jedoch sehen, dass der Fokus auf die *durchschnittlichen* mündlichen Noten einen systematischen Effekt verbirgt, der erst bei einer differenzierteren Betrachtung sichtbar wird.

6.2.4 Vornotenorientierte Notenvergabe in mündlichen Prüfungen

In Abschnitt 6.2.2 haben wir gezeigt, dass es deutliche Diskontinuitäten in der Verteilung der Gesamtnoten des zweiten Examens gibt (siehe Abbildung 2). Wodurch entsteht dieses ungewöhnliche Verteilungsbild? Ein Blick auf die Verteilung der *schriftlichen* Noten zeigt, dass es weder zu „unnatürlichen“ Häufungen über, noch zu einer fehlender Beobachtungsmasse unter den relevanten Schwellenwerten kommt (Abbildung 3). Dies legt nahe, dass die Anomalien in der Verteilung durch eine vornotenorientierte Notengebung in den mündlichen Teilen des zweiten Examens produziert werden. Wie bereits zuvor angesprochen, scheint also das Wissen über die schriftlichen Vornoten von den Mitgliedern der Prüfungskommissionen berücksichtigt zu werden, um über Anpassungen der mündlichen Noten „eindeutige“ Gesamtnoten herzustellen, die nicht knapp unter Notenstufen liegen. Diese Interpretation wird durch einen genaueren Blick auf die Verteilung der mündlichen Noten bestätigt.

Um den Punkt der vornotenorientierten Notengebung grafisch zu untersuchen, berechnen wir für alle Referendarinnen und Referendare jenen fiktiven Schwellenwert, den sie – gegeben durch ihre schriftlichen Noten – in der mündlichen Prüfung erzielen müssten, um die nächste Notenstufe zu erreichen.¹⁵ In einem zweiten Schritt stellen wir dann die Notenverteilung aus der mündlichen Prüfung, normiert um den jeweiligen, Referendar-spezifischen Schwellenwert, dar. Der Nullpunkt stellt dabei genau jene mündliche Note dar, die für das Erreichen der Notenstufe notwendig ist. Wenden wir diesen Ansatz auf alle Studierenden mit einer schriftlichen Note von maximal 1,5 Punkten unter oder 0,5 Punkten über einer der vier Notenschwellen (4,0 / 6,5 / 9,0 / 11,5 oder 14,0) an, so erhalten wir die in Abbildung 4 dargestellte Verteilung.

¹⁵ Um diesen Ansatz zu illustrieren, betrachten wir das Beispiel einer hypothetischen Studierenden, die in den schriftlichen Teilen eine durchschnittliche Note von 6,0 erzielt hat. Die Studentin müsste in den mündlichen Teilen einen Notenschnitt von 7,25 schaffen, um mindestens die Schwelle von 6,5 zu erreichen ($0,6 \times 6,0 + 0,4 \times 7,25 = 6,5$). Für einen Studenten mit einer schriftlichen Durchschnitt von 8,5 wäre der fiktive Schwellenwert in der mündlichen Prüfung gleich 9,75, um die Notenstufe von 9,0 zu erreichen ($0,6 \times 8,5 + 0,4 \times 9,75 = 9,0$).

Abbildung 3: Notenverteilung im zweiten Examen: Durchschnitt aus schriftlichen Teilnoten.

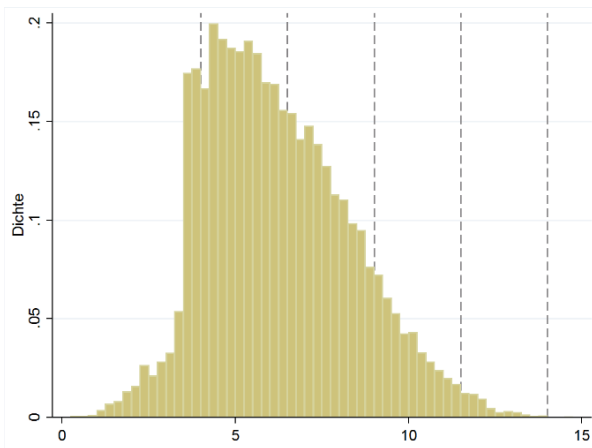
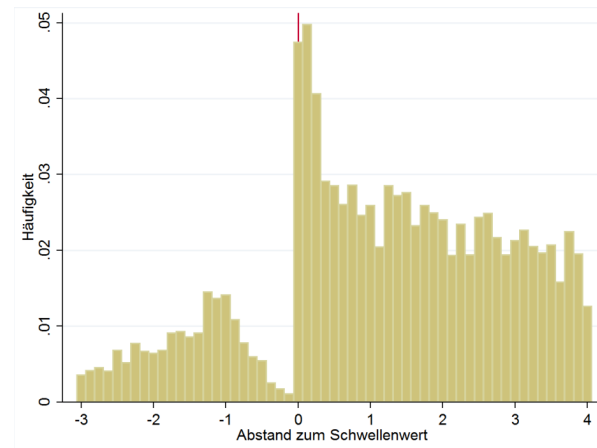


Abbildung 4: Verteilung der mündlichen Noten relativ zum Schwellenwert.



Anm. Die Abbildung rechts bildet die Häufigkeit der Noten aus den mündlichen Prüfungsteilen, relativ zu einem virtuellen Schwellenwert, ab. Der Schwellenwert (der auf Null normiert wird) gibt jene mündliche Note an, die mindestens erreicht werden muss, um eine relevante Notenstufe (d.h., Gesamtnote von 4,0, 7,5, 9,0, oder 11,5) zu erreichen. Noten „über [unter]“ Null geben an, dass die Notenstufe [nicht] erreicht wird.

Im Fall einer von der Vornote unabhängigen Notenvergabe in der mündlichen Prüfung sollten wir eine Verteilung beobachten, die um den Null-Punkt stetig streut. Abbildung 4 zeigt indessen ein deutlich anderes Bild auf: Fälle, in denen Studierende knapp (genauer: zwischen 0,1 und 0,5 Notenpunkte) unter ihren jeweiligen Zielwert liegen, treten überzufällig selten auf. Gleichzeitig sehen wir eine deutliche Häufung von Fällen, in denen der jeweilige Schwellenwert relativ genau erreicht wird. Dieses Muster zeigt sehr deutlich die vornotenorientierte Vergabe von mündlichen Noten auf: Die Häufung der Werte oberhalb der Schwelle in Kombination mit den fehlenden Werten unterhalb der Schwelle spricht für „strategisches“ Anheben. Abbildung A1 (Appendix) zeigt zusätzlich, dass dieses Muster an allen vier Schwellenwerten auftritt.

6.2.5 Detailanalyse der Geschlechtsunterschiede in den mündlichen Teilen

Die in Abschnitt 6.2.2 dargestellten Analysen der Schwellen-Erreichung auf Basis der Gesamtnoten für alle Prüflinge haben gezeigt, dass generelle Geschlechtsunterschiede bei der Erreichung der Notenschwellen bestehen. Die im letzten Abschnitt dargestellten Analysen zeigen zudem, dass Geschlechtsunterschiede besonders stark bei den nicht anonymen mündlichen Prüfungen festzustellen sind (und *zusätzlich* zu den Unterschieden in schriftlichen Noten und Vornoten auftreten). Vor dem Hintergrund dieser klaren Evidenz (die in Abbildung 4 besonders anschaulich gezeigt wird und die eine systematische Notenverschiebung durch die Prüfungskommissionen dokumentiert) stellt sich nun die Frage, inwiefern es hier zu einer Gleichbehandlung von Prüflingen mit gleichen Voraussetzungen oder Vornoten kommt, bzw. ob alle Referendarinnen und Referendare die gleiche Wahrscheinlichkeit haben, die nächste Notenstufe zu erreichen.

Zur Beantwortung der Forschungsfrage 3 – ob Frauen und Männer mit gleichen Voraussetzungen auch die gleichen Chancen haben, eine Notenstufe zu erreichen – schätzen wir ferner die Wahrscheinlichkeit der Schwellenwert-Erreichung in der mündlichen Prüfung. Um Forschungsfrage 6 zu beantworten,

fragen wir darüber hinaus, ob die Zusammensetzung der Prüfungskommission einen Einfluss auf diese Wahrscheinlichkeiten hat (und wenn ja: welchen).

Wie für Abbildung 4 betrachten wir dazu alle Kandidatinnen und Kandidaten, die nach den schriftlichen Noten in einem hinreichend engen Bereich um eine der relevanten Notenstufen (4,0; 6,5; 9,0; 11,5) liegen.¹⁶ Mit dieser Stichproben-Definition wird dann (über ein lineares Wahrscheinlichkeitsmodell; LPM) geschätzt, wie hoch die Wahrscheinlichkeit ist, dass Referendarinnen und Referendare mit bestimmten Vornoten einen Schwellenwert überspringen und die nächste Notenstufe erreichen.

Unsere Analyse kontrolliert dabei für den Abstand zum jeweiligen Schwellenwert – der sich aus der schriftlichen Note ergibt –, da die Wahrscheinlichkeit der Schwellenerreichung maßgeblich von diesem Abstand getrieben wird. Dies bedeutet, dass eine mögliche „Quelle“ für Geschlechtsunterschiede in der Wahrscheinlichkeit einen Schwellenwert zu erreichen – nämlich die im Vergleich zu Männern schlechteren schriftlichen Noten von Frauen – bereits absorbiert wird. Anders ausgedrückt zeigt die Analyse mögliche Geschlechtsunterschiede zwischen Studierenden auf, die *mit gleicher schriftlicher Durchschnittsnote* in die mündliche Prüfung gehen.

Die Resultate der Schätzungen zeigen, dass – bei Betrachtung aller vier Schwellenwerte – Frauen durchschnittlich eine um 1,3%-Punkte niedrigere Wahrscheinlichkeit haben, die nächste Notenstufe zu erreichen als Männer (M: 84,5% F: 83,3%; Appendix A, Tabelle A9, Modell 1). Durchschnittlich bedeutet hier auch, dass dieses Ergebnis für eine „durchschnittliche Prüfungskommission“ gilt. Differenziert man zwischen ausschließlich mit männlichen Prüfern besetzten Kommissionen und solchen, an denen zumindest eine Prüferin beteiligt ist, so ergibt sich ein deutlich anderer Befund: Sind alle Prüfer männlich, so finden wir einen größeren Geschlechtsunterschied von 2,3%-Punkten (Tabelle A9, Modell 2). Ist jedoch zumindest eine Frau Teil der Prüfungskommission, so verschwindet der Geschlechtsunterschied: Relativ zu einer rein männlich besetzten Kommission erhöht sich die Wahrscheinlichkeit von Frauen, die nächste Notenschwelle zu erreichen, marginal (+0,9%-Punkte), während die Wahrscheinlichkeit von Männern signifikant fällt (-1,8%-Punkte). In „weiblichen“ Prüfungskommissionen verbleiben damit keine quantitativ relevanten oder statistisch signifikanten Geschlechtsunterschiede.

Wenn wir (basierend auf der zugrundeliegenden Schätzung, siehe Modell 2) testen, ob – aus epistemischen Gründen etwas umständlich ausgedrückt – die Präsenz einer weiblichen Prüferin *keinen* differenzialen Effekt auf männliche und weibliche Studierende hat, so verwerfen wir diese Null-Hypothese (mit $p=0,012$). Die empirische Evidenz zeigt also auf, dass die Zusammensetzung der Prüfungskommission einen klaren Einfluss auf die Geschlechtsunterschiede im Erreichen der Notenschwellen hat. Diese Befunde bleiben sowohl qualitativ wie auch quantitativ weitgehend unverändert, wenn wir für zusätzliche Faktoren (etwa die Note aus dem ersten Examen) kontrollieren (Appendix A, Tabelle A9, Modelle 3 und 4).

Eine differenzierte Analyse dieser Effekte an unterschiedlichen Notenschwellen zeigt zusätzlich, dass diese Effekte vor allem in höheren Notenbereichen – genauer: an den Schwellenwerten 9,0 und 11,5 – auftreten (Tabelle A9, Modell 5-8). Für diese beiden Schwellenwerte besteht in ausschließlich männlich besetzten Kommissionen ein substantieller Geschlechtsunterschied von 6,1%-Punkten: Für Frauen besteht also – selbst bei gleichen schriftlichen Noten – eine deutlich niedrigere Wahrscheinlichkeit, ein

¹⁶ Genauer betrachten wir dazu wieder Fälle mit schriftlichen Noten, die -1,5 bis +0,5 um die Schwellenwerte liegen. Weder die grafische Evidenz aus Abbildung 4 noch die im Folgenden dargestellten Regressionsergebnisse sind sensitiv hinsichtlich der exakten Definition dieses Bereichs rund um die Schwellenwerte. Die Regressionsergebnisse sind qualitativ auch robust zwischen linearen (LPM) und nicht-linearen Schätzern (Probit, Logit).

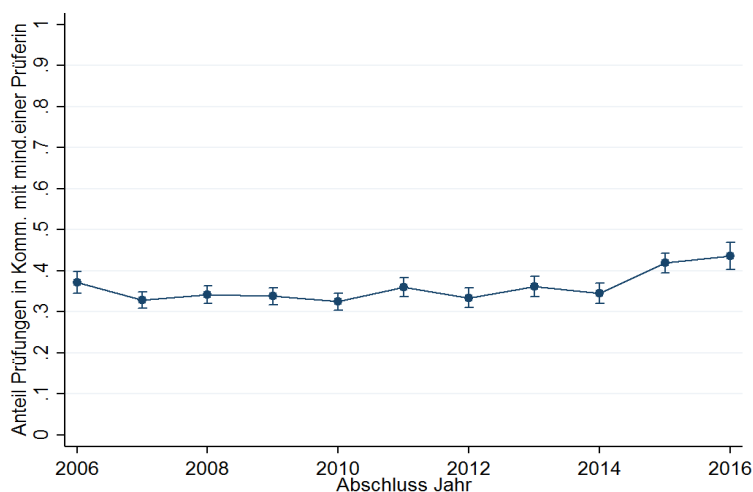
Prädikatsexamen oder sogar ein „gut“ zu erreichen (Tabelle A9, Modell 6). Ist zumindest eine Prüferin Mitglied der Kommission, so erhöht sich (im Vergleich zum Ergebnis in einer männlich besetzten Kommission) die Wahrscheinlichkeit von Frauen, die Schwelle von 9,0 oder 11,5 Punkten zu erreichen, genauso signifikant (nämlich um +3,7%-Punkte), wie die Wahrscheinlichkeit von Männern zurückgeht (-3,1%-Punkte). Wie zuvor heben diese Effekte in Summe den Nachteil von weiblichen Studierenden in rein männlich besetzten Kommissionen (6,1%) auf, sodass in Kommissionen mit mindestens einer Frau keine signifikanten Geschlechtsunterschiede mehr zu beobachten sind.

Ergebnis F3 (Geschlecht): Frauen haben in der mündlichen Prüfung – selbst nach Kontrolle für die durchschnittliche Vornote im schriftlichen Teil – eine signifikant geringere Wahrscheinlichkeit, die nächste Notenschwelle zu erreichen als Männer. Diese Geschlechtsunterschiede sind bei den Schwellwerten 9,0 und 11,5 stärker ausgeprägt.

Ergebnis F6: In einer rein männlich besetzten Prüfungskommission sind diese Geschlechtsunterschiede deutlich stärker ausgeprägt. Ist mindestens eine Frau Teil der Prüfungskommission, so verschwinden die Geschlechtsunterschiede vollständig. Dieser nivellierende Effekt ist bei den Schwellwerten von 9,0 und 11,5 Punkten stärker ausgeprägt.

Weitere Analysen deuten darauf hin, dass der „Nivellierungs-Effekt“ nicht mit der Zahl der Frauen in der Kommission steigt. Die Aussagekraft dieser Analyse ist jedoch stark eingeschränkt, da kaum Beobachtungen mit mehr als einer Frau pro Kommission vorliegen: 65% der Prüfungen werden vor rein männlich besetzten, 30% von Kommissionen mit genau einer Frau, 5% von Kommissionen mit zwei und nur 0,3% vor Kommissionen mit drei Prüferinnen abgenommen. Bei einer zeitlichen Betrachtung ist allerdings zu beobachten, dass der Anteil der Prüfungen mit mindestens einer Frau in der Kommission in den letzten Jahren – insbesondere in den Jahren 2015 und 2016 – angestiegen ist (siehe Abbildung 5). Vor dem Hintergrund der oben berichteten Befunde erscheint dieser Trend höchst begrüßenswert.

Abbildung 5: Anteil von Kommissionen im zweiten Examen mit mindestens einer Prüferin (Mittelwerte und 95%-Konfidenzintervall).

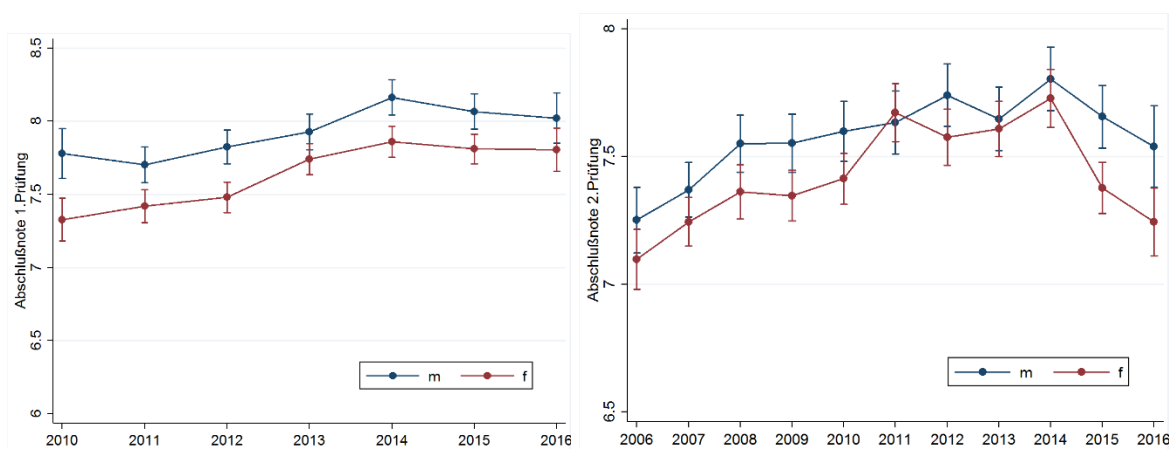


6.2.6 Entwicklung der Geschlechterunterschiede zwischen dem ersten und zweiten Examen sowie über Abschlussjahrgänge hinweg

Ein Vergleich der absoluten Größe der beobachteten Geschlechtsunterschiede im ersten und zweiten Examen zeigt, dass sich diese nicht substantiell unterscheiden, im ersten Examen aber tendenziell etwas größer sind. Bei Kontrolle für die Note aus der ersten Prüfung verringern sich die Geschlechterunterschiede in der zweiten Prüfung, in vielen Fällen verschwinden diese. Punktuell – etwa bei den mündlichen Noten oder bei den schriftlichen Noten im Zivilrecht – deutet einiges auf zusätzliche Effekte im zweiten Examen hin. Zusammengefasst liegt somit keine eindeutige Evidenz darüber vor, wie sich die Geschlechtsunterschiede zwischen den Prüfungen verändern.

Bei einer Analyse der zeitlichen Trends in den durchschnittlichen Abschlussnoten zeigt sich kein eindeutiger Effekt. Für das erste Examen (Abbildung 6, links) scheint sich der Notenunterschied zwischen den Geschlechtern über die Zeit etwas zu reduzieren. Dieser Rückgang ist jedoch weder statistisch signifikant noch robust hinsichtlich der Kontrolle für kohortenspezifische Eigenschaften. Gleiches gilt für die Variation über unterschiedliche Abschlussjahrgänge, die für das zweite Examen beobachtet werden (Abbildung 6, rechts). Insbesondere der Vergleich mit den mündlichen Noten (siehe Appendix A, Abbildung A2) legt nahe, dass Schwankungen der Gesamtnoten eher durch nicht von uns erfasste (und/oder Zufalls-)Faktoren bedingt sind.

Abbildung 6: Entwicklung von Geschlechterunterschieden in Abschlussnoten über Abschlussjahrgänge im ersten (links) und zweiten (rechts) Examen.



Anm. Die Abbildung zeigt Mittelwerte gemeinsam mit den 95%-Konfidenzintervallen.

Ergebnis F4 (Geschlecht): Es zeigt sich kein eindeutiger Trend, dass sich die Geschlechterunterschiede in den Noten über die Abschlussjahrgänge verringern würden.

7 ANALYSEN ZU EFFEKTEN EINES MIGRATIONSHINTERGRUNDS

Zur Analyse der Effekte eines potenziellen Migrationshintergrunds auf die Notengebung verwenden wir als direkte und indirekte Indikatoren – wie bereits oben eingeführt – den Geburtsort (nicht-deutsch: 9,1%), das Nicht-Vorhandensein einer deutschen Staatsangehörigkeit (nicht-deutsch: 2,8%) sowie die Herkunftszuordnung der Namen nach Länder/Regionen auf Basis eines Onomastik-Verfahren (nicht-deutsch: 14%). Unter Berücksichtigung aller drei Indikatoren differenzieren wir – wie oben besprochen – auch Personen mit Migrationserfahrung aus der ersten bzw. der zweiten oder späteren Generationen.

Basierend auf diesen Indikatoren werden zwei Arten von Analysen durchgeführt. **Analyse A** verwendet stets die drei eigenständigen Indikatoren, und zwar Indikatoren für eine nicht-deutsche Namensherkunft (I1), für einen Geburtsort außerhalb Deutschlands (I2) und für keine deutsche Staatsangehörigkeit (I3). Die Effektgrößen für diese drei Indikatoren sind dabei als kumulierende Werte zu verstehen, d.h., die Effekte summieren sich bei Personen mit mehreren dieser Indikatoren (linear additiv).

Analyse B verwendet Analysen mit den zwei abgeleiteten Indikatoren (MI) für einen Migrationshintergrund der 1. bzw. der ≥ 2 . Generation. Nachdem diese beiden Indikatoren ausschließend sind, bilden die geschätzten Werte den jeweiligen durchschnittlichen Unterschied zur Referenzgruppe (Dt. Name/in Dt. geboren) ab.¹⁷

In beiden Analyse-Ansätzen werden auch Spezifikationen berücksichtigt, die für die aus dem Onomastik-Verfahren abgeleiteten Herkunftsregionen differenzieren. Es wird hier (gemäß dem UN-Klassifikationsprotokoll) differenziert zwischen folgenden MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien (Mittlerer Osten), (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft. Die Referenzkategorie ist dabei Zentral- und Nord-Europa, sodass neben Deutschen auch Studierende mit nicht-deutscher Herkunft (etwa aus Frankreich, Österreich oder Dänemark) diese Referenzgruppe bilden. In diesen Spezifikationen gilt es weiterhin zu berücksichtigen, dass die jeweilig geschätzten Effektgrößen aus den primären Migrationsvariablen (I1-3, bzw. MI) mit dem Wert der jeweiligen Herkunftsregion addiert werden müssen, um „Gesamteffekte“ einschätzen zu können.

Der Fokus der Analysen liegt wiederum auf der „finalen Note“, ggf. nach Wiederholungs- und Verbesserungsversuch. Da die primären Indikatoren für die Migrationsanalyse nicht für alle Personen vorhanden sind (siehe oben) reduziert sich die Größe der verfügbaren Stichprobe für die Hauptanalysen des zweiten Examens auf ca. 8.700 relevante Beobachtungen.

Nachdem durch diese (und auch andere) Spezifikationen der Umfang der Schätztabellen deutlich größer ist als bei jenen in Abschnitt 6 (Geschlechterunterschiede) besprochenen Tabellen, werden nun alle Tabellen im Anhang aufgeführt. Dies soll jedoch keinesfalls als „Abwertung“ der hier präsentierten Ergebnisse verstanden werden.

¹⁷ Zusätzlich zu den zwei Migrationsindikatoren kontrollieren wir auch noch für eine dritte Gruppe von Studierenden, die zwar im Ausland geboren wurden, jedoch einen deutschen Namen und eine deutsche Staatsbürgerschaft haben. Nachdem es sich dabei um eine Gruppe mit sehr kleiner Fallzahl handelt, werden die geschätzten Werte für diese Gruppe (die in den Tabellen in der Variable „MI: Dt./Geb.Ausland“ abgebildet werden) im Folgenden nicht getrennt diskutiert.

7.1 Effekte des Migrationshintergrunds im ersten Examen

7.1.1 Unterschiede in der Gesamtnote

Die Analyse der Gesamtnote im ersten Examen zeigt signifikante negative Effekte für alle drei Indikatoren (Appendix B, Tabelle B1, Modell 1), die auch bei Kontrolle für die Abiturnote und Geschlecht bestehen bleiben (Tabelle B1, Modell 2). Auch für die Noten in der staatlichen Pflichtfachprüfung gilt dieser Befund (Tabelle B1, Modelle 5-6). Bei den Noten für die universitären Prüfungsteile sind, je nach Spezifikation, die Effekte für den ersten (I1) bzw. den dritten Indikator (I3) statistisch insignifikant (Modell 8 bzw. 9).

Die Größe der geschätzten Effekte ist dabei erheblich. Für die Gesamtnote zeigt die Basisspezifikation (Tabelle B1, Modell 1) folgende Notenunterschiede auf: Bei einer/m nicht-deutschen...

- Namensherkunft (I1) ist die Gesamtnote um 0,79 Notenpunkte schlechter;
- Geburtsort (I2) ist die Gesamtnote um 0,37 Notenpunkte schlechter;
- Staatsangehörigkeit (I3) ist die Gesamtnote um 0,26 Notenpunkte schlechter.

Oder anschaulicher: Ein durchschnittlicher „deutscher“ Prüfling (d.h. mit Null-Werten in allen drei Indikatoren) erzielt im untersuchten Datensatz im Durchschnitt eine Gesamtnote von 7,93 Punkten. Studierende mit einer nicht-deutschen Namensherkunft, einem nicht-deutschen Geburtsort und ohne deutsche Staatsangehörigkeit erreichen im Durchschnitt lediglich eine Gesamtnote von 6,51. D.h., letztere Gruppe hat im ersten Examen eine um etwa 18% (1,42 Punkte) schlechtere Abschlussnote.

Interessant zu beobachten ist dabei auch, dass der Einfluss der drei Indikatoren auf die Gesamtnote nach Kontrolle für die Note aus der staatlichen Pflichtfachprüfung nahezu vollständig verschwindet (Tabelle B1, Modelle 3).¹⁸ Dies legt nahe, dass im universitären Prüfungsteil (im Vergleich zum staatlichen Teil) keine weiteren, „zusätzlichen“ Notenunterschiede entstehen, die in diesen drei Indikatoren aufgefangen werden.

Die Analysen (B) der abgeleiteten Indikatoren für einen Migrationshintergrund liefern ein ähnliches Bild (Appendix B, Tabelle B2). In nahezu allen Spezifikationen ist jedoch der Notenunterschied für Migranten der ersten Generation stärker ausgeprägt als für Studierende aus der zweiten oder höheren Generation. Dieser Unterschied innerhalb der Gruppe von Studierenden mit Migrationshintergrund ist, für die Gesamtnote, nur in Modell 2 signifikant (siehe p-Wert des entsprechenden F-Tests in Modell 2). Bei getrennter Betrachtung der universitären und der staatlichen Pflichtfachprüfung ist der „geringere Nachteil“ der 2+ Generation nahezu durchgehend signifikant unterschiedlich zum deutlich stärkeren Notenunterschied der ersten Generation (siehe die F-Tests für Modelle 5-7 sowie für Modelle 9 und 10). Die Evidenz deutet daher auf einen positiven integrativen Effekt hin, der zu einer Verminderung der Notennachteile zu führen scheint.

7.1.2 Wahrscheinlichkeiten der Erreichung von Notenstufen im ersten Examen

Motiviert durch die Diskontinuitäten in der Notenverteilung (vergleiche Abbildung 1) wurde der Einfluss der unterschiedlichen Migrationsindikatoren auf die Wahrscheinlichkeit untersucht, bestimmte Notenschwellen im der Gesamtnote des ersten Examens zu erreichen. Analyse A für die Notenschwelle

¹⁸ Der Effekt für I2 (den Geburtsort) bleibt jedoch auf einem schwachen Signifikanzniveau bestehen.

9,0 zeigt dabei erneut einen starken Einfluss der drei Indikatoren (Tabelle B3, Modell 1): die Wahrscheinlichkeit, eine Note von 9,0 oder mehr zu erreichen ist bei einer/m nicht-deutschen...

- Namensherkunft (I1) um 10,5 %-Punkte geringer;
- Geburtsort (I2) um 7,9 %-Punkte geringer;
- Staatsangehörigkeit (I3) um 6,7 %-Punkte geringer.

In Summe ist damit die Wahrscheinlichkeit auf eine Gesamtnote von 9,0 oder mehr um insgesamt 25%-Punkte geringer. Bei jenen „deutschen“ Studierenden, bei denen alle drei Indikatoren auf Null stehen, liegt diese Wahrscheinlichkeit im Übrigen bei etwa 31%.

Es zeigt sich jedoch, dass diese substantiellen Effekte teilweise an statistischer Signifikanz verlieren, wenn für die Abiturnote bzw. die Vornote aus dem staatlichen Teil des ersten Examens kontrolliert wird (Tabelle B3, Modelle 2 und 3); andere Effekte werden dabei jedoch etwas größer, sodass Analyse A keine allzu klare Aussage hinsichtlich der abschwächenden Rolle der Vornoten ermöglicht.

Analyse B liefert hier ein klareres Bild (Tabelle B6, Modelle 1). Wie zuvor finden wir – hier nun auch signifikant – stärkere, negative Effekte für den Migrationsindikator der ersten Generation (-19,2 %-Punkte) als für die den Indikator der zweiten/höheren Generation (-12,6 %-Punkte). Nach Kontrolle für die Abiturnote (Modell 2) verliert letzterer Migrationsindikator sogar an statistischer Signifikanz; der erste Indikator bleibt jedoch hoch-signifikant und bildet einen starken negativen Unterschied (von weiterhin 16,7 %-Punkte) ab. Die Ergebnisse legen nahe, dass bei einer gleich-guten (oder schlechten) Abiturnote, Studierende mit einem Migrationshintergrund, insofern sie bereits in Deutschland geboren wurden, eine gleich hohe Chance haben, ein Prädikatsexamen zu erreichen. Für Migrantinnen und Migranten der ersten Generation gilt dies jedoch nicht. Selbst bei gleich-guten Abiturnoten werden gute Noten unterdurchschnittlich häufig erreicht.¹⁹

In zusätzlichen Analyseschritten wurden diese Effekte auch für die staatliche Pflichtfachprüfung des ersten Examens sowie für weitere Notenschwellen (6,5 aber auch über 11,5 Notenpunkte) repliziert (Tabellen B3-B8). Die Ergebnisse zeigen hier vor allem für die *höheren* Notenwerte einen *größeren* Nachteil von Studierenden mit Migrationshintergrund. Der Befund, dass ein struktureller Unterschied zwischen erster und zweiter/höherer Migranten- und Migrantinnen-Generationen bestehen, verfestigt sich dabei.

Ergebnis F1 (Migration): Es zeigt sich ein sehr starker Einfluss aller drei direkten und indirekten Migrations-Indikatoren auf die Note im ersten Examen. Die Effekte der 'harten' direkten Migrations-Indikatoren (I2 und I3) bleiben bei Kontrolle für Abiturnote und Geschlecht tendenziell bestehen. Für Migranten der ersten Generation (Geburt im Ausland) sind die Unterschiede deutlich stärker und robuster. Für Studierende mit Migrationshintergrund, die bereits in Deutschland geboren wurden, sind die Notenunterschiede geringer.

¹⁹ Unklar bleibt hier, inwiefern die An- und Über-Rechnung von in Drittstaaten erzielten Abiturnoten zu einer Verzerrung der Befunde führt.

7.2. Effekte des Migrationshintergrunds im zweiten Examen

7.2.1 Unterschiede in der Gesamtnote

Analysen der Gesamtnoten im 2. Examen zeigen die folgenden Effekte für die direkten Indikatoren (N = 8.757; Tabelle B9, Modell 1): die durchschnittliche Gesamtnote im zweiten Examen ist ...

- bei nicht-deutscher Namensherkunft (I1) um 0,82 Notenpunkte schlechter;
- bei einem nicht-deutschen Geburtsort (I2) um 0,49 Notenpunkte schlechter.

Bezüglich der Staatsangehörigkeit (I3) zeigt sich kein statistisch signifikanter Unterschied für die Gesamtnote. In Summe ergibt sich aus den beiden hoch-signifikanten Indikatoren eine – im Vergleich zu einem „deutschen Referendar“ (Durchschnittsnote 7,74) – um bis zu 17% (1,31 Punkte) schlechtere Note.

Die Effektgrößen gehen nur marginal zurück, wenn für Alter (nicht-linear), Geschlecht und Abschlusszeitpunkt-spezifische Effekte kontrolliert wird (Tabelle B9, Modell 2). Wird auch für die Abschlussnote aus dem ersten Examen kontrolliert, so schrumpfen die Effekte erheblich (auf 0,32 (I1) bzw. 0,13 Notenpunkte (I2); Modell 3), bleiben jedoch auf dem 1%- bzw. einem hohen 5%-Niveau statistisch signifikant. Vergleicht man also Referendarinnen und Referendare mit exakt gleichen Noten im ersten Examen, so erzielen jene mit einem nicht-deutschen Namen bzw. einem nicht-deutschen Geburtsort im zweiten Examen weiterhin systematisch schlechtere Noten.

Trennt man hier nach unterschiedlichen Herkunftsregionen (Tabelle B9, Modell 4), so zeigt sich ein hohes Maß an Heterogenität innerhalb der Gruppe der „nicht-deutschen“ Referendarinnen und Referendare. Für Referendarinnen und Referendare mit osteuropäischen Wurzeln (MI: Region 4) zeigt sich kein oder nur ein kleiner negativer Effekt, während für Studierende aus dem nahen- und mittleren Osten (MI: Region 3) die größten Unterschiede festzustellen sind.²⁰ Bei der Interpretation dieser regionalen Herkunfts-Indikatoren sollte jedoch vorsichtig argumentiert werden, zumal auch innerhalb der einzelnen regionalen Gruppen teilweise ein hohes Maß an Heterogenität vorliegt. Die weiterführende Analyse diese Unterschiede würden den Rahmen dieses Berichts deutlich sprengen.

Abschließend wird auch noch für die Abiturnote kontrolliert (Tabelle B9, Modell 5). Damit verliert man zwar die Hälfte der Beobachtungen, die relevanten Effektgrößen – im Unterschied zu den p-Werten (bzw. den darunter liegenden Standardfehlern) – ändern sich aber kaum. Nicht unerwähnt bleiben soll ein unerwartet positiver Effekt für den Indikator I3: zumal nur sehr wenige in diesem Sample eine nicht-deutsche Staatsbürgerschaft haben, sollte auch dieser Effekt (dessen statistische Signifikanz auf dem 10%-Niveau liegt) nicht überinterpretiert werden.

Basierend auf den abgeleiteten Indikatoren liefert Analyse B sehr ähnliche Resultate (Tabelle B10). Der durchschnittliche Notenunterschied für Studierende mit einem Migrationshintergrund liegt hier bei etwa 1,0 bis 1,1 Notenpunkte (Tabelle B10, Modell 1). Es liegen keine statistisch unterschiedlichen Effekte für die beiden Indikatoren der ersten bzw. zweiten/höheren Generation vor (ein F-Test hinsichtlich der Generationenunterschiede liefert einen p-Wert von 0,43).

Wie zuvor betrachtet sinken die Notenunterschiede, wenn für Alter, Geschlecht, Abschlusszeitpunkt (Modell 2) oder auch die Note aus dem ersten Examen (Modell 3) kontrolliert wird. Beide Migration-indikatoren weisen aber weiterhin einen hoch-signifikanten Effekt auf. Das Signifikanzniveau verrin-

²⁰ Wie oben erwähnt muss in dieser Spezifikation der Effekt aus den jeweiligen Indikatoren mit dem jeweiligen Herkunftsland-Indikator summiert werden, um einen gruppenspezifischen Effekt bewerten zu können.

gert sich jedoch drastisch, wenn auch für die Abiturnote kontrolliert wird (Modell 5): das in der Schätzung verwendete Sample wird substantiell kleiner und nur mehr der Effekt für die 1. Generations-Migranten bleibt auf einem schwachen 10%-Niveau statistisch signifikant.

Ergebnis F2 und F2a (Migration): Es zeigt sich ein sehr starker Einfluss der abgeleiteten und teilweise auch der direkten Indikatoren für Migrationshintergrund (insbesondere Namensherkunft I1 und Geburtsort nicht Deutsch I2) auf die Gesamtnote im zweiten Examen. Diese Effekte bleiben in abgeschwächter Form auch nach Kontrolle für die Note im ersten Examen und teilweise auch nach Kontrolle für die Abiturnote bestehen.

7.2.2 Wahrscheinlichkeit, relevante Notenschwellen zu erreichen

Konsistent mit den Ergebnissen für die Gesamtnoten zeigen die beiden Analyseansätze A und B auch hinsichtlich der Wahrscheinlichkeit eine Note von 9,0 oder höher zu erreichen, einen starken Einfluss der unterschiedlichen Migrationsindikatoren auf (vergleiche Tabellen B11 und B14). Der Einfluss der Indikatoren sinkt dabei nach Kontrolle für Alter, Geschlecht, Abschlusszeitpunkt und für die Note in der ersten Prüfung bzw. die Abiturnote; signifikante Notenunterschiede sind typischerweise aber auch nach Kontrollen für diese Noten festzustellen. Zwischen den beiden Gruppen der ersten und zweiten/höheren Generationen gibt es keinen statistisch signifikanten Unterschied. Abschließend kann auch festgehalten werden, dass die *relativen* Unterschiede für „deutsche“ Referendarinnen und Referendaren und solche mit Migrationshintergrund am oberen Ende der Notenverteilung etwas stärker ausgeprägt sind (vergleiche Tabellen B12 und B13 bzw. B15 und B16).²¹

7.2.3 Differenzierte Analyse der Migrationseffekte auf schriftliche und mündliche Noten

In einem nächsten Analyseschritt wird nun betrachtet, ob sich die Migrationseffekte zwischen den schriftlichen und mündlichen Prüfungsteilen unterscheiden. Die Untersuchung der schriftlichen Noten im zweiten Examen zeigt dabei konsistente und starke Effekte der beiden ersten Indikatoren (Tabelle B17, Modell 1). Die durchschnittliche schriftliche Note ist bei Studierenden mit...

- nicht-deutscher Namensherkunft (I1) um 1,11 Notenpunkte schlechter;
- nicht-deutschem Geburtsort (I2) um 0,52 Notenpunkte schlechter.

Selbst ohne Berücksichtigung des negativen (aber insignifikanten) Effektes des dritten Indikators ergibt sich damit ein Notenunterschied von bis zu 1,63 Notenpunkten, was – relativ zur schriftlichen Durchschnittsnote von 6,25 Punkten bei stereotypen deutschen Referendarinnen und Referendaren – einen Unterschied von 26% ausmacht.

Unter Berücksichtigung von Kontrollvariablen gehen diese Unterschiede teilweise recht deutlich zurück, bleiben jedoch über alle Spezifikationen hinweg statistisch signifikant (Tabelle B17, Modelle 2-5). Analyseansatz B bestätigt diese Ergebnisse, liefert teilweise jedoch statistisch noch robustere und präzisere geschätzte Effektgrößen (Tabelle B20).

²¹ Um die relativen Effektgrößen zu bestimmen, müssen in den unterschiedlichen Tabellen die Koeffizienten jeweils mit der Konstante aus Modell 1 verglichen werden. (Die Konstante bildet die Wahrscheinlichkeit ab, mit der die Referenzgruppe – „deutsche“ Studierende – die jeweilige Notenschwelle erreichen.)

Auch für die durchschnittliche Note aus den mündlichen Prüfungsteilen zeigt sich ein vergleichbares Bild. Auch hier weisen die beiden ersten Indikatoren wieder quantitativ starke, statistisch hoch-signifikante Notenunterschiede auf (Tabelle B18, Modell 1), die sich auf 1,45 Notenpunkte addieren. In absoluten Zahlen ist diese Effektgröße damit ähnlich zu jenem Notenunterschied, der für die schriftlichen Noten festgestellt wurde. Relativ zu der besseren mündlichen Durchschnittsnote (9,52 Punkte), beläuft sich der Unterschied jedoch auf „nur“ 15% - und damit deutlich geringer, als die 26% die für die schriftlichen Noten gemessen wurden. Relativ betrachtet, scheinen die Migrationseffekte eine stärkere Rolle in der schriftlichen als in der mündlichen Prüfung zu spielen.

Nach Kontrolle für Alter, Geschlecht und Abschlusszeitpunkts-spezifischen Effekten gehen die Effekte etwas zurück, bleiben dabei aber hoch-signifikant unterschiedlich von Null (Tabelle B18, Modell 2). Wird auch für die schriftlichen Teilnoten kontrolliert (Modell 3), so sinken die Effekte weiter; der zweite Indikator (I2) verliert dabei auch an statistischer Signifikanz. Ein recht ähnliches Bild ergibt sich, wenn anstatt der schriftlichen Teilnoten für die Note aus dem ersten Examen kontrolliert wird (Modell 4): in beiden Fällen behält der erste Indikator einen hoch-signifikanten negativen Effekt.

Vergleichbare, aber statistisch etwas präzisere Resultate liefert Analyseansatz B (Tabelle B21). Unter Verwendung der beiden abgeleiteten MI Indikatoren findet man sehr stark ausgeprägte Effekte (Tabelle B21, Modell 1), die auch nach Kontrolle für die schriftlichen Noten im 2. Examen oder für die Gesamtnote in der ersten Prüfung auf hohem Signifikanzniveau bestehen bleiben (Modelle 2-5). Beachtenswert erscheint auch, dass in diesen Schätzungen *keine* Vorteile der zweiten/höheren gegenüber der ersten Migrantengeneration festgestellt werden kann.²² Sollten sich positive integrative Effekte auf einen sichereren mündlichen Sprachgebrauch niederschlagen – so schlagen sich diese zumindest nicht in besseren mündlichen Noten im zweiten Examen nieder.

Eine Reihe weiterer spannender, wenn auch nur suggestiver Beobachtungen können aus dem Vergleich von Modellen 4 und 5 zwischen Tabelle B20 und Tabelle B21 hergeleitet werden. Bei den schriftlichen Noten scheinen die Studierenden aus Ost-Europa (MI Region 4; aber auch Afrika, MI Region 1) die „geringsten Probleme“ zu haben; die größten Notenunterschiede bestehen hier für Referendarinnen und Referendare mit Wurzeln im Nahen und Mittleren Osten (MI Region 3). Bei den mündlichen Noten schneiden Referendarinnen und Referendare mit Wurzeln in Ost-Europa ähnlich wie andere Gruppen ab; die schlechtesten mündlichen Noten erzielen wieder Studierende die aus dem Nahen und Mittleren Osten aber auch jene, die aus Süd-Europa stammen.

Ergebnis F2b (Migration): Personen mit Indikatoren für einen Migrationshintergrund erzielen in den mündlichen Prüfungen im zweiten Examen schlechtere Noten als andere Prüflinge. Dies gilt auch bei Kontrolle für die anonyme schriftliche Note bzw. für die Note im ersten Examen.

7.2.4 Differenzierte Analyse der einzelnen Rechtsgebiete

Um die Migrationseffekte zwischen den einzelnen Rechtsgebieten zu differenzieren, wurden die schriftlichen Teilnoten des zweiten Examens getrennt nach Rechtsgebiet untersucht. Die unterschiedlichen Analyseansätze liefern dazu *keine* Evidenz für differenzierte Migrationseffekte, die nach Rechts-

²² Es besteht sogar schwache Evidenz für den kontraintuitiven, umgekehrten Fall (Tabelle B21, F-Tests für Modelle 3 und 4).

gebieten unterschiedlich ausgeprägt wären (Tabelle B19 und B22). Ganz im Gegenteil, die Resultate dokumentieren sehr ähnliche Effektgrößen für die unterschiedlichen Teilgebiete.

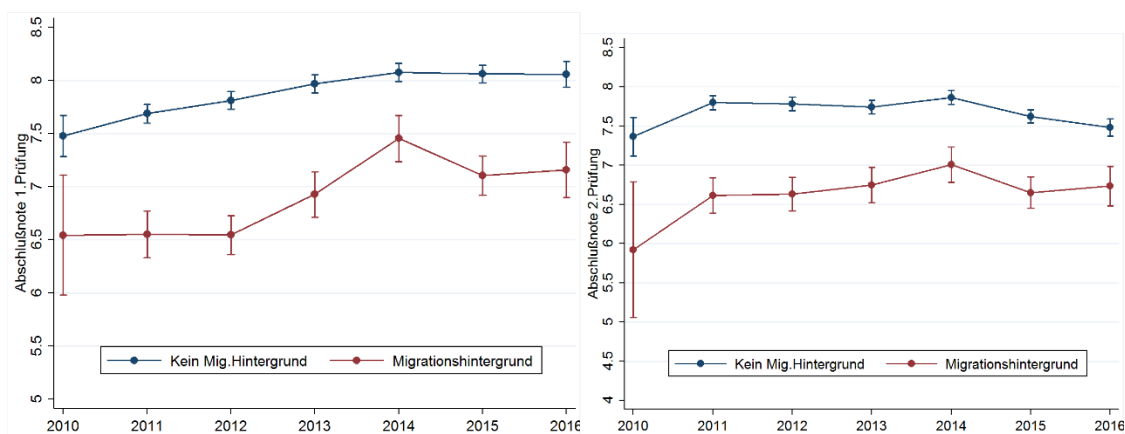
7.2.5 Entwicklung zwischen beiden Abschlussprüfungen sowie über Abschlussjahrgänge

Generell sind die Größen der Effekte in der ersten Prüfung und der zweiten juristischen Staatsprüfung vergleichbar. Teilweise bleiben Effekte im 2. Examen bestehen bei Kontrolle für die Note in der ersten Prüfung.

Der Unterschied zwischen Personen mit Migrationsindikatoren und anderen Referendarinnen und Referendaren bleibt relativ stabil über Abschlussjahre mit einem leichten deskriptiven Trend zur Reduktion der Unterschiede (Abbildung 6).

Ergebnis F4 und F5 (Migration): Es zeigt sich kein statistisch belastbarer, klarer Trend, dass sich die Migrationsunterschiede in den Noten über die Abschlussjahrgänge oder zwischen erstem und zweitem Examen verringern würden.

Abbildung 6: Entwicklung von Migrationsunterschieden in Abschlussnoten über Abschlussjahrgänge im ersten (links) und zweiten (rechts) Examen.



Anm. Die Abbildung zeigt Mittelwerte gemeinsam mit den 95%-Konfidenzintervallen.

8 ÜBERBLICK ÜBER DIE ERGEBNISSE

8.1 Analysen zu Effekten des Geschlechts

- Ergebnis F1 (Geschlecht im ersten Examen): Auch bei einer umfangreichen Datengrundlage und bei differenzierter Betrachtung zeigt sich im ersten Examen ein substantieller negativer Effekt in der Gesamtbenotung von Frauen im Vergleich zu Männern. Der Noten-Unterschied wird hauptsächlich durch den staatlichen Teil getrieben.10
- Ergebnis F2 und F2a (Geschlecht): Es zeigen sich signifikante Geschlechterunterschiede in der Gesamtnote im zweiten Examen. Diese Unterschiede bleiben – in abgeschwächter Form – auch nach Kontrolle für die Note aus dem ersten Examen sowie der Abiturnote qualitativ bestehen.13
- Ergänzung zu Ergebnissen F2 und F2a (Geschlecht): Der abgeschwächte negative Effekt für Frauen im Mittelwert der Gesamtnote und das Verschwinden des Effekts bei Betrachtung der Notenschwellen bei Kontrolle für die Note aus dem ersten Examen deuten darauf hin, dass sich der Geschlechterunterschied zwischen den Examen kaum verstärkt.....15
- Ergebnis F2b (Geschlecht): Frauen erzielen in den mündlichen Prüfungen im zweiten Examen schlechtere Noten als Männer auch bei Kontrolle für die anonyme schriftliche Note und auch bei zusätzlicher Kontrolle für Abitur und Note im ersten Examen. Generell sind die Geschlechterunterschiede in den mündlichen Noten größer und robuster als in den schriftlichen Noten.16
- Ergebnis F3 (Geschlecht): Frauen haben in der mündlichen Prüfung – selbst nach Kontrolle für die durchschnittliche Vornote im schriftlichen Teil – eine signifikant geringere Wahrscheinlichkeit, die nächste Notenschwelle zu erreichen als Männer. Diese Geschlechtsunterschiede sind bei den Schwellwerten 9,0 und 11,5 stärker ausgeprägt.20
- Ergebnis F6: In einer rein männlich besetzten Prüfungskommission sind diese Geschlechtsunterschiede deutlich stärker ausgeprägt. Ist mindestens eine Frau Teil der Prüfungskommission, so verschwinden die Geschlechtsunterschiede vollständig. Dieser nivellierende Effekt ist bei den Schwellwerten von 9,0 und 11,5 Punkten stärker ausgeprägt.....20
- Ergebnis F4 (Geschlecht): Es zeigt sich kein eindeutiger Trend, dass sich die Geschlechterunterschiede in den Noten über die Abschlussjahrgänge verringern würden.....21

8.2 Analysen zu Effekten eines Migrationshintergrunds

- Ergebnis F1 (Migration): Es zeigt sich ein sehr starker Einfluss aller drei direkten und indirekten Migrations-Indikatoren auf die Note im ersten Examen. Die Effekte der 'harten' direkten Migrations-Indikatoren (I2 und I3) bleiben bei Kontrolle für Abiturnote und Geschlecht tendenziell bestehen. Für Migranten der ersten Generation (Geburt im Ausland) sind die Unterschiede deutlich stärker und robuster. Für Studierende mit Migrationshintergrund, die bereits in Deutschland geboren wurden, sind die Notenunterschiede geringer.24
- Ergebnis F2 und F2a (Migration): Es zeigt sich ein sehr starker Einfluss der abgeleiteten und teilweise auch der direkten Indikatoren für Migrationshintergrund (insbesondere Namensherkunft I1 und Geburtsort nicht Deutsch I2) auf die Gesamtnote im zweiten Examen. Diese Effekte bleiben in abgeschwächter Form auch nach Kontrolle für die Note im ersten Examen und teilweise auch nach Kontrolle für die Abiturnote bestehen.....26

Ergebnis F2b (Migration): Personen mit Indikatoren für einen Migrationshintergrund erzielen in den mündlichen Prüfungen im zweiten Examen schlechtere Noten als andere Prüflinge. Dies gilt auch bei Kontrolle für die anonyme schriftliche Note bzw. für die Note im ersten Examen.....27

Ergebnis F4 und F5 (Migration): Es zeigt sich kein statistisch belastbarer, klarer Trend, dass sich die Migrationsunterschiede in den Noten über die Abschlussjahrgänge oder zwischen erstem und zweitem Examen verringern würden.....28

APPENDIX A: GESCHLECHTSEFFEKTE

Table A1: Wahrscheinlichkeit Note von 6,5 Punkte oder mehr zu erreichen (1. Examen)

	(1)	(2)	(3)	(4)	(5)
Frauen	-0,045*** [0,000]	-0,057*** [0,000]	0,008 [0,464]	-0,038*** [0,008]	-0,093*** [0,000]
Abiturnote		-0,257*** [0,000]	-0,047*** [0,000]		-0,285*** [0,000]
Note staatl. Prüfungsteil			0,136*** [0,000]		
Konstante	0,740*** [0,000]	1,356*** [0,000]	-0,145*** [0,001]	0,670*** [0,000]	1,345*** [0,000]
N	10.042	4.596	4.252	4.390	4.254
R ²	0,002	0,124	0,412	0,002	0,130

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau:
 *** p<0,01; ** p<0,05; * p<0,1.

Table A2: Wahrscheinlichkeit Note von 11,5 Punkte oder mehr zu erreichen (1. Examen).

	(1)	(2)	(3)	(4)	(5)
Frauen	-0,020*** [0,000]	-0,033*** [0,000]	-0,014** [0,011]	-0,022*** [0,000]	-0,033*** [0,000]
Abiturnote		-0,075*** [0,000]	-0,012*** [0,007]		-0,063*** [0,000]
Note staatl. Prüfungsteil			0,041*** [0,000]		
Konstante	0,052*** [0,000]	0,223*** [0,000]	-0,230*** [0,000]	0,045*** [0,000]	0,194*** [0,000]
N	10.042	4.596	4.252	4.390	4.254
R ²	0,003	0,059	0,200	0,004	0,047

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau:
 *** p<0,01; ** p<0,05; * p<0,1.

Tabelle A3: Differenzierte Geschlechtereffekte auf Gesamtnote (2. Examen).

	(1)	(2)	(3)	(4)
Frauen	0,043 [0,687]	-0,257** [0,013]	-0,395*** [0,000]	0,009 [0,898]
Frauen x Bielefeld	0,200 [0,462]	0,058 [0,821]	0,146 [0,544]	0,104 [0,544]
Frauen x Bochum	0,029 [0,871]	-0,027 [0,882]	-0,037 [0,828]	-0,053 [0,689]
Frauen x Bonn	-0,387** [0,015]	-0,316** [0,039]	-0,263* [0,057]	-0,174* [0,088]
Frauen x Düsseldorf	0,030 [0,897]	0,011 [0,962]	0,030 [0,888]	-0,070 [0,638]
Frauen x Münster	0,202 [0,202]	0,187 [0,225]	0,179 [0,200]	-0,069 [0,503]
Bielefeld	-0,678*** [0,001]	-0,508*** [0,007]	-0,180 [0,325]	0,155 [0,253]
Bochum	-0,737*** [0,000]	-0,706*** [0,000]	-0,311** [0,015]	0,283*** [0,006]
Bonn	0,285** [0,019]	0,235** [0,043]	0,243** [0,019]	0,317*** [0,000]
Düsseldorf	-0,491*** [0,009]	-0,656*** [0,000]	-0,669*** [0,000]	-0,170 [0,152]
Münster	0,034 [0,779]	-0,133 [0,277]	-0,137 [0,217]	0,410*** [0,000]
Abiturnote			-1,248*** [0,000]	-0,257*** [0,000]
Gesamtnote 1. Examen				0,702*** [0,000]
Zusätzliche Kontrollvariablen	Nein	Ja	Ja	Ja
Konstante	7,672*** [0,000]	28,814*** [0,000]	21,963*** [0,000]	6,433*** [0,000]
N	4.255	4.255	4.254	4.251
R2	0,029	0,128	0,270	0,606

Analyse von Geschlechterunterschieden getrennt nach Universität der Studierenden. Referenzgruppe: Universität Köln; Modelle (2) – (4) kontrollieren für das Alter zum Prüfungszeitpunkt (linearer und quadratischer Term) sowie für Abschlussmonats-spezifische Effekte. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Table A4: Wahrscheinlichkeit Note von 6,5 Punkte oder mehr zu erreichen (2. Examen).

	(1)	(2)	(3)	(4)
Frauen	-0,005 [0,442]	-0,072*** [0,000]	-0,007 [0,408]	0,005 [0,683]
Note aus 1. Examen Abiturnote			0,116*** [0,000]	0,115*** [0,000] -0,036*** [0,002]
Konstante	0,689*** [0,000]	4,095*** [0,000]	1,662*** [0,000]	1,350** [0,011]
Weitere Kontrollvariablen	Nein	Ja	Ja	Ja
N	17.971	17.971	9.086	4.251
R ²	0,000	0,084	0,314	0,312

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Modelle (2) – (4) kontrollieren für das Alter zum Prüfungszeitpunkt (linearer und quadratischer Term) sowie für Abschlussmonats-spezifische Effekte. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Table A5: Wahrscheinlichkeit Note von 11,5 Punkte oder mehr zu erreichen (2. Examen).

	(1)	(2)	(3)	(4)
Frauen	-0,015*** [0,000]	-0,022*** [0,000]	-0,002 [0,633]	-0,006 [0,266]
Note			0,036*** [0,000]	0,029*** [0,000]
1. Examen				
Abiturnote				-0,014*** [0,001]
Konstante	0,047*** [0,000]	0,415*** [0,000]	-0,205** [0,040]	-0,205 [0,295]
Weitere Kontrollvariablen	Nein	Ja	Ja	Ja
N	17.971	17.971	9.086	4.251
R ²	0,001	0,014	0,131	0,126

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Modelle (2) – (4) kontrollieren für das Alter zum Prüfungszeitpunkt (linearer und quadratischer Term) sowie für Abschlussmonats-spezifische Effekte. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle A6: Durchschnittliche schriftliche Note im zweiten Examen.

	(1)	(2)	(3)	(4)
Frauen	-0,116*** [0,000]	-0,441*** [0,000]	-0,045 [0,124]	-0,024 [0,574]
Note 1. Examen			0,750*** [0,000]	0,718*** [0,000]
Abiturnote				-0,215*** [0,000]
Weitere Kontrollvariablen	Nein	Ja	Ja	Ja
Konstante	6,230*** [0,000]	22,260*** [0,000]	9,148*** [0,000]	6,274*** [0,000]
N	18.958	18.958	9.969	4.579
R ²	0,001	0,126	0,581	0,575

Modelle (2) – (4) kontrollieren für das Alter zum Prüfungszeitpunkt (linearer und quadratischer Term) sowie für Abschlussmonats-spezifische Effekte. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$.

Tabelle A7: Durchschnittliche mündliche Note im zweiten Examen.

	(1)	(2)	(3)	(4)	(5)	(6)
Frauen	-0,225*** [0,000]	-0,538*** [0,000]	-0,546*** [0,000]	-0,238*** [0,000]	-0,132*** [0,006]	-0,205*** [0,003]
Frauen x FPK			-0,006 [0,902]	0,024 [0,518]	0,027 [0,614]	0,097 [0,185]
Männer x FPK			-0,028 [0,613]	-0,023 [0,579]	0,064 [0,270]	0,021 [0,801]
Note schriftl. Prüfungsteile				0,823*** [0,000]		
Note 1. Examen					0,778*** [0,000]	0,733*** [0,000]
Abiturnote						-0,294*** [0,000]
Konstante	9,319*** [0,000]	25,741*** [0,000]	25,753*** [0,000]	7,561*** [0,000]	9,298*** [0,000]	7,983*** [0,001]
Weitere Kontrollvariablen	Nein	Ja	Ja	Ja	Ja	Ja
N	17.970	17.970	17.969	17.969	9.084	4.251
R ²	0,002	0,078	0,078	0,487	0,449	0,450
F-test FPK (p-Werte)			0,767	0,395	0,634	0,482

Die Variable „FPK“ ist ein binärer Indikator für die Beteiligung von mindestens einer Frau in der Prüfungskommission. In der Zeile F-test werden p-Werte für Tests der Null-Hypothese „Frauen X FPK = Männer x FPK“ (also der Hypothese, dass die Präsenz einer Frau in der Prüfungskommission den gleichen Effekt auf Männer und Frauen hat) berichtet. Modelle (2) – (6) kontrollieren für das Alter zum Prüfungszeitpunkt (linearer und quadratischer Term) sowie für Abschlussmonats-spezifische Effekte. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle A8: Schriftliche Teilnoten getrennt nach Rechtsgebieten (2. Examen).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variablen:	Zivilrecht (Modelle 1-3)			Strafrecht (Modelle 4-6)			Öffentliches Recht (Modelle 7-9)		
Frauen	-0,174*** [0,000]	-0,132*** [0,000]	-0,078 [0,126]	0,008 [0,877]	0,043 [0,290]	0,016 [0,790]	-0,028 [0,601]	0,054 [0,209]	0,044 [0,476]
Gesamtnote		0,769*** [0,000]	0,731*** [0,000]		0,702*** [0,000]	0,679*** [0,000]		0,779*** [0,000]	0,731*** [0,000]
1. Examen									
Abiturnote			-0,228*** [0,000]			-0,091* [0,090]			-0,311*** [0,000]
Konstante	6,258*** [0,000]	9,337*** [0,000]	6,142*** [0,000]	5,807*** [0,000]	9,556*** [0,000]	7,560*** [0,000]	6,092*** [0,000]	7,582*** [0,000]	5,251*** [0,000]
Weitere									
Kontrollvariablen	Nein	Ja	Ja	Nein	Ja	Ja	Nein	Ja	Ja
N	9.696	9.364	4.576	9.696	9.364	4.576	9.696	9.364	4.576
R ²	0,001	0,509	0,499	0,000	0,406	0,386	0,000	0,426	0,407

Abhängige Variable ist die durchschnittliche, schriftliche Note für die Rechtsgebiete Zivil- (Modelle 1-3), Straf- (4-6) und Öffentliches-Recht (7-9). Modelle (2), (3), (5), (6), (8) und (9) kontrollieren jeweils für das Alter zum Prüfungszeitpunkt (linearer und quadratischer Term) sowie für Abschlussmonats-spezifische Effekte. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle A9: Schwellenerreichung für Männer und Frauen im 2. Examen.

Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Schwellenerreichung alle Werte (4,0; 6,5; 9,0; 11,5)				Schwellenerreichung 4,0 und 6,5 9,0 und 11,5		Schwellenerreichung 9,0 11,5	
Frauen	-0,013** [0,014]	-0,023*** [0,001]	-0,026*** [0,000]	-0,018* [0,051]	-0,010 [0,176]	-0,061*** [0,000]	-0,035** [0,034]	-0,158*** [0,000]
Frauen x FPK		0,009 [0,228]	0,008 [0,284]	0,013 [0,218]	-0,004 [0,610]	0,037** [0,030]	0,024 [0,187]	0,077* [0,098]
Männer x FPK		-0,018** [0,020]	-0,020** [0,012]	-0,020* [0,067]	-0,016* [0,071]	-0,031* [0,067]	-0,014 [0,456]	-0,082* [0,064]
Note 1. Examen				0,042*** [0,000]				
Abstand zum Schwellenwert (Schriftliche Note)	-0,260*** [0,000]	-0,260*** [0,000]	-0,259*** [0,000]	-0,228*** [0,000]	-0,232*** [0,000]	-0,302*** [0,000]	-0,284*** [0,000]	-0,390*** [0,000]
Notenstufe 4,0	0,026*** [0,000]	0,026*** [0,000]	0,030*** [0,000]	0,088*** [0,000]	0,049*** [0,000]			
Notenstufe 9,0	-0,013* [0,075]	-0,014* [0,073]	-0,016** [0,041]	-0,093*** [0,000]		0,183*** [0,000]		
Notenstufe 11,5	-0,203*** [0,000]	-0,203*** [0,000]	-0,206*** [0,000]	-0,372*** [0,000]				
Zusätzliche Kontrollvariablen	Nein	Nein	Ja	Ja	Ja	Ja	Ja	Ja
N	15.031	15.031	15.031	7.677	10.452	4.579	3.666	913
R ²	0,231	0,231	0,241	0,268	0,223	0,237	0,202	0,347
F-test (p-Werte)	-	0,012	0,010	0,028	0,330	0,005	0,144	0,015

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Abhängige Variable ist ein Indikator dafür, Gesamtnoten „über“ den vier Notenstufen (Modelle 1-4), über den zwei „unteren“ (Modelle 5-6) bzw. „oberen“ Notenstufen (Modelle 7-8) zu erreichen. Alle Modelle inkludieren eine Konstante (nicht in Tabelle aufgenommen). Die Variable „FPK“ ist ein Indikator für die Beteiligung von mindestens einer Frau in der Prüfungskommission. F-test dokumentiert p-Werte für Tests der Null-Hypothese „Frauen X FPK = Männer x FPK“ (also der Hypothese, dass die Präsenz einer Frau in der Kommission auf Männer und Frauen den gleichen Effekt hat). Die Notenstufen-Indikatoren bilden ab, ob die schriftliche Note eines Studierenden rund um den jeweiligen Notenwert liegt. (In Modellen 1-4 ist die Referenzgruppe die Notenstufe 6,5.) Modelle (3) – (8) kontrollieren für das Alter zum Prüfungszeitpunkt (linearer und quadratischer Term) sowie für Abschlussmonats-spezifische Effekte. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Abbildung A1: Verteilung der mündlichen Noten relativ zum jeweiligen Schwellenwert.

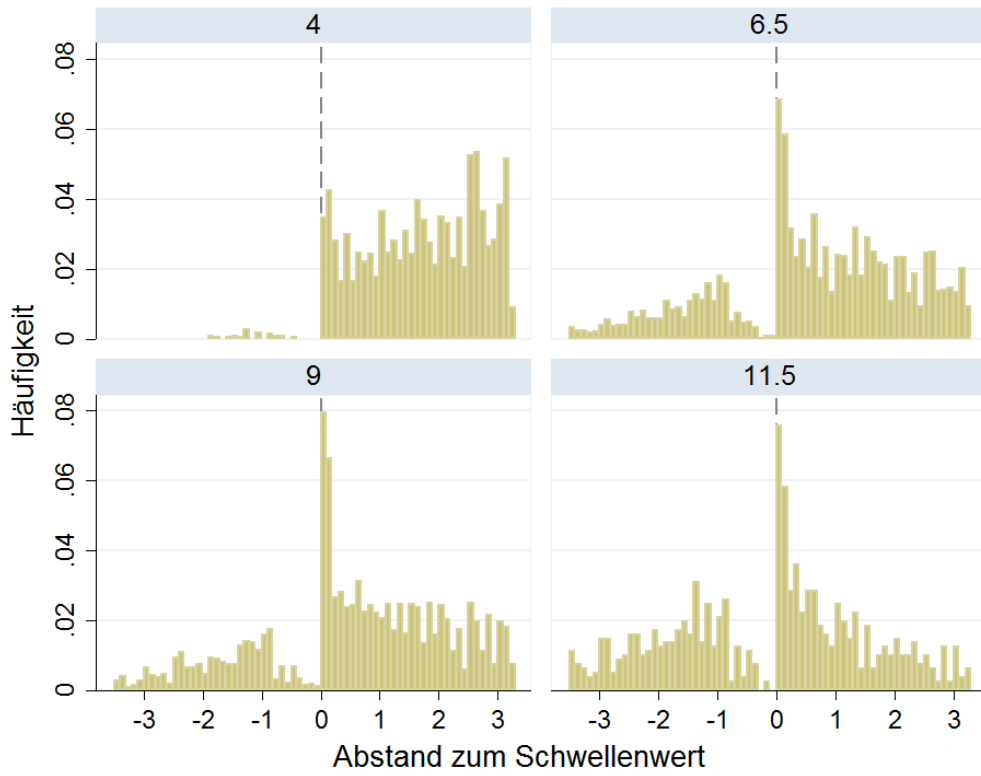
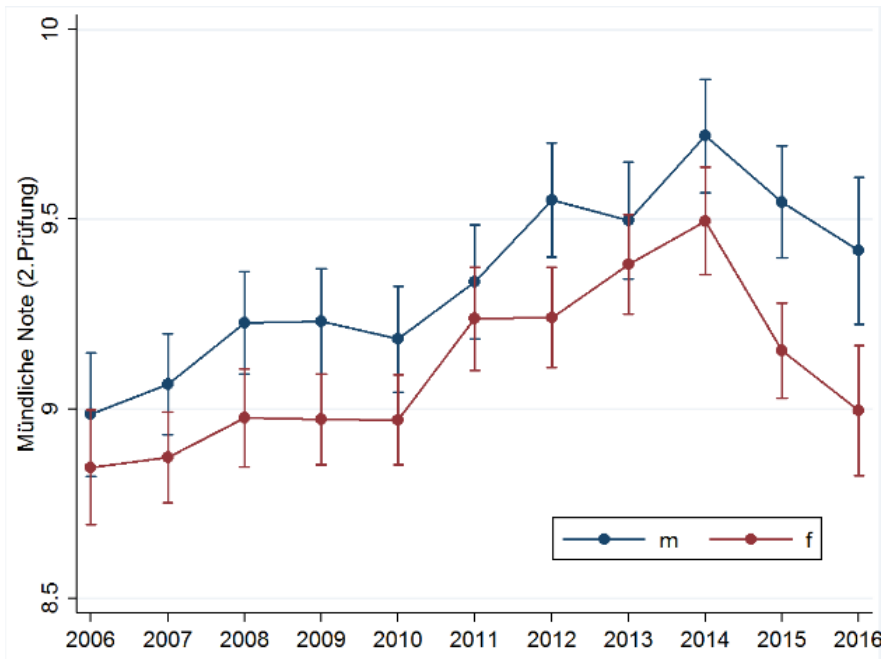


Abbildung A2: Zeitliche Entwicklung der Geschlechterunterschiede in den mündlichen Noten im zweiten Examen.



APPENDIX B: MIGRATIONSEFFEKTE

Tabelle B1: Effekte der Migrationsindikatoren auf Noten im 1. Examen (Analyse A)

Abhäng. Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Gesamtnote (erste Prüfung)				staatliche Pflichtfachprüfung			universitäre Prüfung		
I1: Onomastik	-0,792*** [0,000]	-0,180** [0,032]	-0,030 [0,371]	0,097 [0,494]	-0,395*** [0,000]	-0,179* [0,059]	0,295 [0,413]	-0,395*** [0,003]	-0,190 [0,119]	0,474 [0,359]
I2: Geburtsort	-0,366*** [0,000]	-0,440*** [0,000]	-0,080* [0,054]	-0,076* [0,088]	-0,390*** [0,001]	-0,462*** [0,000]	-0,427*** [0,001]	-0,437*** [0,003]	-0,502*** [0,001]	-0,473*** [0,003]
I3: Staatsangeh.	-0,259** [0,021]	-0,610*** [0,000]	-0,037 [0,570]	-0,059 [0,377]	-0,420** [0,015]	-0,602*** [0,001]	-0,626*** [0,002]	-0,316 [0,150]	-0,435* [0,062]	-0,522** [0,032]
Frauen		-0,359*** [0,000]	0,033* [0,076]	0,032* [0,082]		-0,493*** [0,000]	-0,494*** [0,000]		-0,143** [0,034]	-0,147** [0,031]
Abiturnote		-1,561*** [0,000]	-0,261*** [0,000]	-0,263*** [0,000]		-1,547*** [0,000]	-1,546*** [0,000]		-1,668*** [0,000]	-1,676*** [0,000]
Note staatliche Pflichtfachpr.			0,855*** [0,000]	0,855*** [0,000]						
MI: Region 1				-0,376* [0,054]			-0,283 [0,599]			-1,401** [0,047]
MI: Region 2				0,030 [0,853]			-0,567 [0,168]			-0,175 [0,767]
MI: Region 3				-0,113 [0,451]			-0,498 [0,191]			-0,635 [0,247]
MI: Region 4				-0,167 [0,281]			-0,549 [0,159]			-0,843 [0,134]
MI: Region 5				-0,136 [0,375]			-0,601 [0,131]			-0,762 [0,168]
MI: Region 6				-0,246 [0,137]			-0,190 [0,652]			-0,918 [0,123]
Konstante	7,930*** [0,000]	11,632*** [0,000]	2,192*** [0,000]	2,197*** [0,000]	7,401*** [0,000]	11,156*** [0,000]	11,154*** [0,000]	9,220*** [0,000]	13,061*** [0,000]	13,080*** [0,000]
N	9.430	4.596	4.252	4.241	4.296	4.254	4.243	4.294	4.252	4.241
R ²	0,032	0,273	0,898	0,899	0,015	0,242	0,242	0,010	0,186	0,188

MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa.
Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B2: Effekte der Migrationsindikatoren im 1. Examen (Analyse B).

Abhängige Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Gesamtnote (erste Prüfung)			staatliche Pflichtfachprüfung			universitäre Prüfung			
MI: 1. Generation	-1,093*** [0,000]	-0,801*** [0,000]	-0,120** [0,011]	-0,093 [0,472]	-0,832*** [0,000]	-0,839*** [0,000]	-0,682** [0,039]	-0,835*** [0,000]	-0,830*** [0,000]	-0,658 [0,165]
MI: ≥ 2. Generation	-0,953*** [0,000]	-0,252*** [0,005]	-0,041 [0,230]	-0,014 [0,906]	-0,520*** [0,000]	-0,234** [0,022]	-0,038 [0,905]	-0,534*** [0,000]	-0,258** [0,042]	-0,067 [0,876]
MI: Dt./Geb.Ausland	-0,642*** [0,000]	-0,344** [0,014]	-0,076 [0,205]	-0,076 [0,207]	-0,522*** [0,003]	-0,336** [0,038]	-0,336** [0,038]	-0,607*** [0,005]	-0,428** [0,048]	-0,427** [0,048]
Frauen		-0,361*** [0,000]	0,033* [0,075]	0,033* [0,080]		-0,495*** [0,000]	-0,498*** [0,000]		-0,144** [0,033]	-0,148** [0,030]
Abiturnote		-1,560*** [0,000]	-0,261*** [0,000]	-0,263*** [0,000]		-1,547*** [0,000]	-1,544*** [0,000]		-1,667*** [0,000]	-1,674*** [0,000]
Note staatl. Pflichtfachpr.			0,855*** [0,000]	0,855*** [0,000]						
MI: Region 1				-0,265 [0,136]			0,125 [0,805]			-0,820 [0,199]
MI: Region 2				0,136 [0,338]			-0,153 [0,685]			0,389 [0,457]
MI: Region 3				-0,010 [0,937]			-0,204 [0,552]			-0,139 [0,767]
MI: Region 4				-0,065 [0,630]			-0,144 [0,690]			-0,294 [0,554]
MI: Region 5				-0,045 [0,727]			-0,440 [0,220]			-0,377 [0,427]
MI: Region 6				-0,139 [0,337]			0,134 [0,730]			-0,392 [0,452]
Konstante	7,941*** [0,000]	11,630*** [0,000]	2,191*** [0,000]	2,194*** [0,000]	7,406*** [0,000]	11,154*** [0,000]	11,149*** [0,000]	9,227*** [0,000]	13,059*** [0,000]	13,077*** [0,000]
N	9.430	4.596	4.252	4.241	4.296	4.254	4.243	4.294	4.252	4.241
R ²	0,033	0,271	0,898	0,899	0,014	0,240	0,240	0,010	0,185	0,187
F-test (p-Werte)	0,168	0,000	0,162	0,224	0,043	0,000	0,001	0,141	0,005	0,013

MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. F-test dokumentiert die p-Werte für Tests der Null-Hypothese dass sich die Effekte für MI: 1. Generation und MI: ≥2.Generation nicht unterscheiden. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B3: Effekte der Migrationsindikatoren auf Stufenerreichung 9 Punkte im 1. Examen (Analyse A).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Gesamtnote über Schwellenwert (9,0)				Note in staatlicher Pflichtfachprüfung		
I1: Onomastik	-0,105*** [0,000]	-0,015 [0,446]	0,017 [0,255]	0,015 [0,756]	-0,042* [0,057]	-0,003 [0,870]	0,090 [0,260]
I2: Geburtsort	-0,079*** [0,000]	-0,111*** [0,000]	-0,045** [0,012]	-0,050** [0,013]	-0,079*** [0,001]	-0,090*** [0,000]	-0,093*** [0,000]
I3: Staatsangeh.	-0,067*** [0,002]	-0,114*** [0,003]	-0,006 [0,844]	-0,002 [0,933]	-0,053 [0,127]	-0,074* [0,059]	-0,076* [0,074]
Frauen		-0,066*** [0,000]	0,013 [0,156]	0,013 [0,153]		-0,071*** [0,000]	-0,071*** [0,000]
Abiturnote		-0,309*** [0,000]	-0,056*** [0,000]	-0,056*** [0,000]		-0,268*** [0,000]	-0,268*** [0,000]
Note staatl. Pflichtfachpr.			0,168*** [0,000]	0,168*** [0,000]			
MI: Region 1				-0,073 [0,308]			-0,117 [0,287]
MI: Region 2				0,085 [0,167]			-0,096 [0,287]
MI: Region 3				-0,009 [0,863]			-0,102 [0,217]
MI: Region 4				-0,014 [0,790]			-0,085 [0,326]
MI: Region 5				0,009 [0,863]			-0,122 [0,154]
MI: Region 6				-0,007 [0,914]			-0,053 [0,599]
Konstante	0,309*** [0,000]	1,023*** [0,000]	-0,827*** [0,000]	-0,827*** [0,000]	0,250*** [0,000]	0,894*** [0,000]	0,891*** [0,000]
N	9.430	4.596	4.252	4.241	4.296	4.254	4.243
R ²	0,015	0,182	0,590	0,590	0,007	0,148	0,148

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Abhängige Variable ist ein Indikator der angibt, ob die Gesamtnote (bzw. ob die Note in der staatlichen Pflichtfachprüfung) über dem Wert von 9,0 liegt. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B4: Effekte der Migrationsindikatoren auf Stufenerreichung 6,5 Punkte im 1. Examen (Analyse A).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Gesamtnote über Schwellenwert (6,5)				Note in staatlicher Pflichtfachprüfung		
I1: Onomastik	-0,181*** [0,000]	-0,063*** [0,010]	-0,037* [0,065]	-0,067 [0,388]	-0,093*** [0,001]	-0,054** [0,039]	-0,090 [0,320]
I2: Geburtsort	-0,064*** [0,002]	-0,066** [0,022]	-0,010 [0,673]	-0,011 [0,656]	-0,077** [0,017]	-0,094*** [0,003]	-0,084** [0,013]
I3: Staatsangeh.	-0,012 [0,734]	-0,114** [0,029]	-0,025 [0,546]	-0,044 [0,310]	-0,088* [0,095]	-0,136** [0,014]	-0,137** [0,018]
Frauen		-0,053*** [0,000]	0,009 [0,408]	0,009 [0,398]		-0,089*** [0,000]	-0,089*** [0,000]
Abiturnote		-0,255*** [0,000]	-0,047*** [0,000]	-0,047*** [0,000]		-0,283*** [0,000]	-0,283*** [0,000]
Note staatl. Pflichtfachpr.			0,135*** [0,000]	0,135*** [0,000]			
MI: Region 1				-0,090 [0,454]			0,048 [0,759]
MI: Region 2				0,005 [0,955]			-0,015 [0,890]
MI: Region 3				0,018 [0,825]			0,043 [0,658]
MI: Region 4				0,059 [0,483]			0,021 [0,833]
MI: Region 5				0,021 [0,811]			0,016 [0,878]
MI: Region 6				0,087 [0,354]			0,143 [0,205]
Konstante	0,748*** [0,000]	1,364*** [0,000]	-0,132*** [0,002]	-0,133*** [0,002]	0,667*** [0,000]	1,355*** [0,000]	1,353*** [0,000]
N	9.430	4.596	4.252	4.241	4.296	4.254	4.243
R ²	0,026	0,133	0,413	0,414	0,012	0,139	0,139

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Abhängige Variable ist ein Indikator der angibt, ob die Gesamtnote (bzw. ob die Note in der staatlichen Pflichtfachprüfung) über dem Wert von 6,5 liegt. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd- Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau:

*** p<0,01; ** p<0,05; * p<0,1.

Tabelle B5: Effekte der Migrationsindikatoren auf Stufenerreichung 11,5 Punkte im 1. Examen (Analyse A).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Gesamtnote über Schwellenwert (11,5)				Note in staatlicher Pflichtfachprüfung		
I1: Onomastik	-0,018*** [0,001]	-0,005 [0,543]	0,002 [0,760]	0,058 [0,265]	-0,012 [0,116]	-0,003 [0,736]	0,025 [0,549]
I2: Geburtsort	-0,021*** [0,000]	-0,019** [0,011]	-0,003 [0,694]	-0,004 [0,700]	-0,026*** [0,000]	-0,028*** [0,000]	-0,031*** [0,001]
I3: Staatsangeh.	-0,018*** [0,000]	-0,024*** [0,002]	0,002 [0,784]	0,002 [0,822]	-0,003 [0,759]	-0,008 [0,512]	-0,004 [0,746]
Frauen		-0,032*** [0,000]	-0,014** [0,011]	-0,014** [0,011]		-0,033*** [0,000]	-0,033*** [0,000]
Abiturnote		-0,074*** [0,000]	-0,012*** [0,007]	-0,012*** [0,006]		-0,062*** [0,000]	-0,062*** [0,000]
Note staatl. Pflichtfachpr.			0,041*** [0,000]	0,041*** [0,000]			
MI: Region 1				-0,073 [0,188]			-0,036 [0,414]
MI: Region 2				-0,055 [0,302]			-0,027 [0,527]
MI: Region 3				-0,046 [0,378]			-0,023 [0,599]
MI: Region 4				-0,054 [0,325]			-0,022 [0,630]
MI: Region 5				-0,065 [0,202]			-0,048 [0,283]
MI: Region 6				-0,093* [0,081]			-0,025 [0,620]
Konstante	0,046*** [0,000]	0,224*** [0,000]	-0,230*** [0,000]	-0,229*** [0,000]	0,037*** [0,000]	0,196*** [0,000]	0,196*** [0,000]
N	9.430	4.596	4.252	4.241	4.296	4.254	4.243
R ²	0,004	0,061	0,200	0,201	0,003	0,049	0,049

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Abhängige Variable ist ein Indikator der angibt, ob die Gesamtnote (bzw. ob die Note in der staatlichen Pflichtfachprüfung) über dem Wert von 11,5 liegt. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B6: Effekte der indirekten Migrationsindikatoren auf Stufenerreichung 9 Punkte im 1. Examen (Analyse B).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Gesamtnote über Schwellenwert (9,0)				Note in staatlicher Pflichtfachprüfung		
MI: 1. Generation	-0,192*** [0,000]	-0,168*** [0,000]	-0,035* [0,073]	-0,058 [0,197]	-0,127*** [0,000]	-0,123*** [0,000]	-0,099 [0,177]
MI: ≥ 2. Generation	-0,126*** [0,000]	-0,022 [0,311]	0,020 [0,205]	0,015 [0,733]	-0,058** [0,014]	-0,007 [0,734]	0,042 [0,564]
MI: Dt./Geb.Ausland	-0,098*** [0,000]	-0,077** [0,018]	-0,035 [0,195]	-0,035 [0,196]	-0,097*** [0,005]	-0,067** [0,045]	-0,067** [0,044]
Frauen		-0,067*** [0,000]	0,013 [0,161]	0,013 [0,161]		-0,072*** [0,000]	-0,071*** [0,000]
Abiturnote		-0,310*** [0,000]	-0,056*** [0,000]	-0,056*** [0,000]		-0,269*** [0,000]	-0,267*** [0,000]
Note staatl. Pflichtfachpr.			0,168*** [0,000]	0,168*** [0,000]			
MI: Region 1				-0,065 [0,339]			-0,053 [0,614]
MI: Region 2				0,100* [0,078]			-0,024 [0,775]
MI: Region 3				-0,004 [0,924]			-0,055 [0,470]
MI: Region 4				0,005 [0,924]			-0,009 [0,906]
MI: Region 5				0,013 [0,793]			-0,091 [0,248]
MI: Region 6				-0,004 [0,952]			-0,003 [0,972]
Konstante	0,310*** [0,000]	1,023*** [0,000]	-0,827*** [0,000]	-0,827*** [0,000]	0,250*** [0,000]	0,893*** [0,000]	0,891*** [0,000]
N	9.430	4.596	4.252	4.241	4.296	4.254	4.243
R ²	0,015	0,181	0,590	0,590	0,006	0,148	0,148
F-test (p-Werte)	0,001	0,000	0,022	0,014	0,029	0,000	0,000

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Abhängige Variable ist ein Indikator der angibt, ob die Gesamtnote (bzw. ob die Note in der staatlichen Pflichtfachprüfung) über dem Wert von 9,0 liegt. F-test dokumentiert die p-Werte für Tests der Null-Hypothese, dass sich die Effekte für MI: 1. Generation und MI: ≥ 2. Generation nicht unterscheiden. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B7: Effekte der indirekten Migrationsindikatoren auf Stufenerreichung 6,5 Punkte im 1. Examen (Analyse B).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Gesamtnote über Schwellenwert (6,5)				Note in staatlicher Pflichtfachprüfung		
MI: 1. Generation	-0,215*** [0,000]	-0,153*** [0,000]	-0,043 [0,104]	-0,081 [0,283]	-0,180*** [0,000]	-0,188*** [0,000]	-0,231** [0,012]
MI: ≥ 2. Generation	-0,210*** [0,000]	-0,084*** [0,001]	-0,048** [0,023]	-0,087 [0,205]	-0,118*** [0,000]	-0,068** [0,016]	-0,112 [0,200]
MI: Dt./Geb.Ausland	-0,132*** [0,000]	-0,067 [0,103]	-0,026 [0,420]	-0,026 [0,420]	-0,104** [0,025]	-0,074* [0,100]	-0,074* [0,100]
Frauen		-0,054*** [0,000]	0,009 [0,393]	0,009 [0,387]		-0,089*** [0,000]	-0,090*** [0,000]
Abiturnote		-0,254*** [0,000]	-0,047*** [0,000]	-0,046*** [0,000]		-0,283*** [0,000]	-0,282*** [0,000]
Note staatl. Pflichtfachpr.			0,135*** [0,000]	0,135*** [0,000]			
MI: Region 1				-0,076 [0,507]			0,082 [0,595]
MI: Region 2				0,010 [0,910]			0,016 [0,881]
MI: Region 3				0,030 [0,690]			0,054 [0,568]
MI: Region 4				0,057 [0,464]			0,048 [0,623]
MI: Region 5				0,023 [0,765]			-0,002 [0,980]
MI: Region 6				0,103 [0,236]			0,161 [0,141]
Konstante	0,751*** [0,000]	1,363*** [0,000]	-0,134*** [0,001]	-0,135*** [0,001]	0,668*** [0,000]	1,354*** [0,000]	1,352*** [0,000]
N	9.430	4.596	4.252	4.241	4.296	4.254	4.243
R ²	0,027	0,132	0,413	0,414	0,011	0,138	0,138
F-test (p-Werte)	0,860	0,090	0,880	0,878	0,161	0,005	0,021

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Abhängige Variable ist ein Indikator der angibt, ob die Gesamtnote (bzw. ob die Note in der staatlichen Pflichtfachprüfung) über dem Wert von 6,5 liegt. F-test dokumentiert die p-Werte für Tests der Null-Hypothese, dass sich die Effekte für MI: 1. Generation und MI: ≥ 2. Generation nicht unterscheiden. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B8: Effekte der indirekten Migrationsindikatoren auf Stufenerreichung 11,5 Punkte im 1. Examen (Analyse B).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Gesamtnote über Schwellenwert (11,5)				Note in staatlicher Pflichtfachprüfung		
MI: 1. Generation	-0,038*** [0,000]	-0,031*** [0,000]	-0,001 [0,847]	0,037 [0,400]	-0,037*** [0,000]	-0,037*** [0,000]	-0,029 [0,393]
MI: ≥ 2. Generation	-0,026*** [0,000]	-0,007 [0,357]	0,003 [0,680]	0,045 [0,254]	-0,014 [0,112]	-0,002 [0,856]	0,016 [0,658]
MI: Dt./Geb.Ausland	-0,032*** [0,000]	-0,015 [0,195]	-0,001 [0,932]	-0,001 [0,935]	-0,029*** [0,002]	-0,021** [0,024]	-0,021** [0,024]
Frauen		-0,032*** [0,000]	-0,014** [0,011]	-0,014** [0,011]		-0,033*** [0,000]	-0,033*** [0,000]
Abiturnote		-0,074*** [0,000]	-0,012*** [0,007]	-0,012*** [0,006]		-0,063*** [0,000]	-0,063*** [0,000]
Note staatl. Pflichtfachpr.			0,041*** [0,000]	0,041*** [0,000]			
MI: Region 1				-0,058 [0,183]			-0,021 [0,570]
MI: Region 2				-0,039 [0,358]			-0,008 [0,815]
MI: Region 3				-0,032 [0,429]			-0,011 [0,752]
MI: Region 4				-0,037 [0,424]			-0,001 [0,987]
MI: Region 5				-0,051 [0,205]			-0,038 [0,283]
MI: Region 6				-0,080* [0,056]			-0,014 [0,757]
Konstante	0,047*** [0,000]	0,224*** [0,000]	-0,230*** [0,000]	-0,230*** [0,000]	0,037*** [0,000]	0,196*** [0,000]	0,196*** [0,000]
N	9.430	4.596	4.252	4.241	4.296	4.254	4.243
R ²	0,004	0,061	0,200	0,201	0,003	0,049	0,050
F-test (p-Werte)	0,076	0,017	0,638	0,662	0,007	0,000	0,014

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Abhängige Variable ist ein Indikator der angibt, ob die Gesamtnote (bzw. ob die Note in der staatlichen Pflichtfachprüfung) über dem Wert von 11,5 liegt. F-test dokumentiert die p-Werte für Tests der Null-Hypothese, dass sich die Effekte für MI: 1. Generation und MI: ≥ 2. Generation nicht unterscheiden. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B9: Effekte der direkten Migrationsindikatoren auf die Gesamtnote im 2. Examen (Analyse A).

	(1)	(2)	(3)	(4)	(5)
I1: Onomastik	-0,819*** [0,000]	-0,678*** [0,000]	-0,315*** [0,000]	-0,454** [0,012]	-0,431 [0,136]
I2: Geburtsort	-0,487*** [0,000]	-0,299*** [0,000]	-0,134** [0,016]	-0,238*** [0,000]	-0,282*** [0,000]
I3: Staatsangeh.	-0,012 [0,923]	0,047 [0,703]	0,165* [0,077]	0,188* [0,056]	0,242* [0,062]
Frauen		-0,347*** [0,000]	-0,037 [0,205]	-0,035 [0,228]	-0,052 [0,183]
Alter		-0,706*** [0,000]	-0,319*** [0,000]	-0,316*** [0,000]	-0,182* [0,079]
Alter ²		0,008*** [0,000]	0,003*** [0,000]	0,003*** [0,000]	0,001 [0,362]
Note 1. Examen			0,723*** [0,000]	0,721*** [0,000]	0,689*** [0,000]
Abiturnote					-0,263*** [0,000]
MI: Region 1				0,378 [0,219]	0,546 [0,268]
MI: Region 2				0,240 [0,268]	0,165 [0,623]
MI: Region 3				-0,075 [0,695]	0,057 [0,851]
MI: Region 4				0,420** [0,029]	0,261 [0,396]
MI: Region 5				0,052 [0,793]	0,098 [0,747]
MI: Region 6				0,259 [0,253]	0,119 [0,733]
Konstante	7,741*** [0,000]	22,156*** [0,000]	8,440*** [0,000]	8,404*** [0,000]	6,834*** [0,000]
N	8.757	8.757	8.504	8.474	4.239
R ²	0,032	0,117	0,584	0,585	0,601

Modelle (2) – (5) kontrollieren auch für Abschlussmonat-spezifische Effekte. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B10: Effekte der abgeleiteten Migrationsindikatoren auf Mittelwerte im 2. Examen (Analyse B).

	(1)	(2)	(3)	(4)	(5)
MI: 1. Generation	-1,097*** [0,000]	-0,775*** [0,000]	-0,284*** [0,000]	-0,450*** [0,005]	-0,445* [0,085]
MI: ≥ 2. Generation	-1,012*** [0,000]	-0,845*** [0,000]	-0,399*** [0,000]	-0,460*** [0,004]	-0,343 [0,157]
MI: Dt./Geb.Ausland	-0,884*** [0,000]	-0,659*** [0,000]	-0,367*** [0,000]	-0,368*** [0,000]	-0,370*** [0,000]
Frauen		-0,342*** [0,000]	-0,033 [0,252]	-0,032 [0,266]	-0,050 [0,205]
Alter		-0,698*** [0,000]	-0,314*** [0,000]	-0,312*** [0,000]	-0,177* [0,087]
Alter ²		0,007*** [0,000]	0,003*** [0,000]	0,003*** [0,000]	0,001 [0,385]
Note 1. Examen			0,722*** [0,000]	0,721*** [0,000]	0,688*** [0,000]
Abiturnote					-0,264*** [0,000]
MI: Region 1				0,324 [0,261]	0,403 [0,380]
MI: Region 2				0,116 [0,552]	-0,004 [0,989]
MI: Region 3				-0,097 [0,570]	-0,044 [0,866]
MI: Region 4				0,253 [0,141]	0,057 [0,834]
MI: Region 5				0,072 [0,687]	0,056 [0,829]
MI: Region 6				0,229 [0,272]	0,015 [0,962]
Konstante	7,755*** [0,000]	22,026*** [0,000]	8,370*** [0,000]	8,347*** [0,000]	6,773*** [0,000]
N	8.757	8.757	8.504	8.474	4-239
R ²	0,035	0,119	0,585	0,585	0,601
F-test (p-Werte)	0,433	0,491	0,129	0,913	0,430

Modelle (2) – (5) kontrollieren für Abschlussmonat-spezifische Effekte. F-test dokumentiert die p-Werte für Tests der Null-Hypothese, dass sich die Effekte für MI: 1. Generation und MI: ≥ 2. Generation nicht unterscheiden. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B11: Effekte der direkten Migrationsindikatoren auf Stufenerreichung 9 Punkte im 2. Examen (Analyse A).

	(1)	(2)	(3)	(4)	(5)
I1: Onomastik	-0,137*** [0,000]	-0,113*** [0,000]	-0,047*** [0,000]	-0,119** [0,018]	-0,124* [0,080]
I2: Geburtsort	-0,082*** [0,000]	-0,050*** [0,002]	-0,020 [0,143]	-0,030* [0,058]	-0,033 [0,131]
I3: Staatsangeh.	-0,023 [0,344]	-0,014 [0,565]	0,015 [0,511]	0,031 [0,187]	0,029 [0,396]
Frauen		-0,068*** [0,000]	-0,009 [0,272]	-0,009 [0,266]	-0,014 [0,223]
Alter		-0,137*** [0,000]	-0,066*** [0,000]	-0,066*** [0,000]	-0,089*** [0,001]
Alter ²		0,002*** [0,000]	0,001*** [0,000]	0,001*** [0,000]	0,001*** [0,003]
Note 1. Examen			0,139*** [0,000]	0,139*** [0,000]	0,135*** [0,000]
Abiturnote					-0,063*** [0,000]
MI: Region 1				0,126* [0,054]	0,070 [0,433]
MI: Region 2				0,127** [0,032]	0,087 [0,316]
MI: Region 3				0,073 [0,163]	0,079 [0,295]
MI: Region 4				0,085 [0,104]	0,020 [0,791]
MI: Region 5				0,046 [0,391]	0,050 [0,506]
MI: Region 6				0,040 [0,509]	0,010 [0,914]
Konstante	0,309*** [0,000]	3,005*** [0,000]	0,435* [0,071]	0,435* [0,070]	0,898** [0,042]
N	8.757	8.757	8.504	8.474	4.239
R ²	0,018	0,067	0,387	0,388	0,405

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Abhängige Variable ist ein Indikator der angibt, ob die Gesamtnote über 9,0 liegt. Modelle (2) – (5) kontrollieren für Abschlussmonat-spezifische Effekte. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B12: Effekte der direkten Migrationsindikatoren auf Stufenerreichung 6,5 Punkte im 2. Examen (Analyse A).

	(1)	(2)	(3)	(4)	(5)
I1: Onomastik	-0,171*** [0,000]	-0,141*** [0,000]	-0,079*** [0,000]	-0,075 [0,188]	-0,110 [0,200]
I2: Geburtsort	-0,084*** [0,000]	-0,046** [0,038]	-0,021 [0,278]	-0,038* [0,065]	-0,059** [0,048]
I3: Staatsangeh.	-0,022 [0,561]	-0,010 [0,792]	0,007 [0,822]	-0,007 [0,849]	-0,003 [0,952]
Frauen		-0,051*** [0,000]	0,001 [0,941]	0,001 [0,906]	0,008 [0,498]
Alter		-0,147*** [0,000]	-0,083*** [0,000]	-0,082*** [0,000]	-0,062* [0,066]
Alter ²		0,002*** [0,000]	0,001*** [0,000]	0,001*** [0,000]	0,001 [0,278]
Note 1. Examen			0,116*** [0,000]	0,115*** [0,000]	0,114*** [0,000]
Abiturnote					-0,036*** [0,002]
MI: Region 1				0,094 [0,356]	0,317** [0,030]
MI: Region 2				-0,016 [0,820]	0,036 [0,725]
MI: Region 3				-0,097 [0,113]	-0,020 [0,832]
MI: Region 4				0,061 [0,323]	0,108 [0,252]
MI: Region 5				0,023 [0,717]	0,074 [0,428]
MI: Region 6				0,074 [0,289]	0,116 [0,278]
Konstante	0,741*** [0,000]	3,751*** [0,000]	1,506*** [0,000]	1,495*** [0,000]	1,220** [0,022]
N	8.757	8.757	8.504	8.474	4.239
R ²	0,024	0,096	0,318	0,320	0,316

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Abhängige Variable ist ein Indikator der angibt, ob die Gesamtnote über 6,5 liegt. Modelle (2) – (5) kontrollieren für Abschlussmonat-spezifische Effekte. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B13: Effekte der direkten Migrationsindikatoren auf Stufenerreichung 11,5 Punkte im 2. Examen (Analyse A).

	(1)	(2)	(3)	(4)	(5)
I1: Onomastik	-0,020*** [0,000]	-0,016*** [0,002]	0,000 [0,952]	-0,043*** [0,000]	-0,045*** [0,006]
I2: Geburtsort	-0,014*** [0,008]	-0,008 [0,127]	-0,000 [0,947]	-0,007 [0,316]	0,006 [0,591]
I3: Staatsangeh.	-0,010* [0,090]	-0,008 [0,179]	-0,000 [0,972]	0,007 [0,366]	0,021 [0,148]
Frauen		-0,017*** [0,000]	-0,003 [0,496]	-0,003 [0,488]	-0,007 [0,245]
Alter		-0,019*** [0,001]	-0,002 [0,747]	-0,002 [0,763]	0,002 [0,855]
Alter ²		0,000*** [0,004]	0,000 [0,669]	0,000 [0,697]	-0,000 [0,882]
Note 1. Examen			0,035*** [0,000]	0,035*** [0,000]	0,029*** [0,000]
Abiturnote					-0,014*** [0,001]
MI: Region 1				0,095*** [0,004]	0,116* [0,068]
MI: Region 2				0,061*** [0,001]	0,056* [0,053]
MI: Region 3				0,046*** [0,000]	0,043** [0,017]
MI: Region 4				0,061*** [0,000]	0,040* [0,067]
MI: Region 5				0,029** [0,011]	0,030* [0,070]
MI: Region 6				0,024 [0,161]	0,009 [0,642]
Konstante	0,041*** [0,000]	0,420*** [0,000]	-0,217** [0,034]	-0,219** [0,032]	-0,205 [0,302]
N	8.757	8.757	8.504	8.474	4.239
R ²	0,003	0,014	0,129	0,131	0,128

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Abhängige Variable ist ein Indikator der angibt, ob die Gesamtnote über 11,5 liegt. Modelle (2) – (5) kontrollieren für Abschlussmonat-spezifische Effekte. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B14: Effekte der abgeleiteten Migrationsindikatoren auf Stufenerreichung 9 Punkte im 2. Examen (Analyse B).

	(1)	(2)	(3)	(4)	(5)
MI: 1. Generation	-0,198*** [0,000]	-0,144*** [0,000]	-0,053*** [0,000]	-0,099** [0,023]	-0,128** [0,045]
MI: ≥ 2. Generation	-0,165*** [0,000]	-0,135*** [0,000]	-0,053*** [0,000]	-0,092** [0,037]	-0,088 [0,144]
MI: Dt./Geb.Ausland	-0,133*** [0,000]	-0,092*** [0,000]	-0,040* [0,057]	-0,039* [0,058]	-0,026 [0,358]
Frauen		-0,067*** [0,000]	-0,009 [0,287]	-0,009 [0,278]	-0,014 [0,224]
Alter		-0,136*** [0,000]	-0,066*** [0,000]	-0,066*** [0,000]	-0,089*** [0,001]
Alter ²		0,002*** [0,000]	0,001*** [0,000]	0,001*** [0,000]	0,001*** [0,003]
Note 1. Examen			0,139*** [0,000]	0,139*** [0,000]	0,135*** [0,000]
Abiturnote					-0,063*** [0,000]
MI: Region 1				0,095 [0,112]	0,037 [0,653]
MI: Region 2				0,090* [0,092]	0,057 [0,468]
MI: Region 3				0,046 [0,325]	0,046 [0,481]
MI: Region 4				0,046 [0,321]	-0,006 [0,927]
MI: Region 5				0,025 [0,599]	0,025 [0,710]
MI: Region 6				0,011 [0,842]	-0,024 [0,772]
Konstante	0,311*** [0,000]	2,991*** [0,000]	0,430* [0,074]	0,430* [0,073]	0,898** [0,042]
N	8.757	8.757	8.504	8.474	4.239
R ²	0,018	0,067	0,387	0,388	0,405
F-test (p-Werte)	0,110	0,682	0,991	0,763	0,249

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Abhängige Variable ist ein Indikator der angibt, ob die Gesamtnote über 9,0 liegt. Modelle (2) – (5) kontrollieren für Abschlussmonat-spezifische Effekte. F-test dokumentiert die p-Werte für Tests der Null-Hypothese, dass sich die Effekte für MI: 1. Generation und MI: ≥ 2. Generation nicht unterscheiden. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B15: Effekte der abgeleiteten Migrationsindikatoren auf Stufenerreichung 6,5 Punkte im 2. Examen (Analyse B).

	(1)	(2)	(3)	(4)	(5)
MI: 1. Generation	-0,217*** [0,000]	-0,150*** [0,000]	-0,070*** [0,001]	-0,088 [0,116]	-0,102 [0,216]
MI: ≥ 2. Generation	-0,214*** [0,000]	-0,179*** [0,000]	-0,104*** [0,000]	-0,108** [0,042]	-0,099 [0,192]
MI: Dt./Geb.Ausland	-0,165*** [0,000]	-0,121*** [0,000]	-0,074*** [0,005]	-0,075*** [0,005]	-0,094** [0,013]
Frauen		-0,050*** [0,000]	0,001 [0,872]	0,002 [0,857]	0,009 [0,469]
Alter		-0,146*** [0,000]	-0,082*** [0,000]	-0,082*** [0,000]	-0,061* [0,070]
Alter ²		0,002*** [0,000]	0,001*** [0,000]	0,001*** [0,000]	0,001 [0,286]
Note 1. Examen			0,116*** [0,000]	0,115*** [0,000]	0,114*** [0,000]
Abiturnote					-0,036*** [0,002]
MI: Region 1				0,110 [0,269]	0,289** [0,037]
MI: Region 2				-0,020 [0,766]	-0,007 [0,941]
MI: Region 3				-0,076 [0,187]	-0,040 [0,626]
MI: Region 4				0,042 [0,470]	0,049 [0,573]
MI: Region 5				0,043 [0,482]	0,052 [0,538]
MI: Region 6				0,097 [0,146]	0,096 [0,333]
Konstante	0,744*** [0,000]	3,724*** [0,000]	1,491*** [0,000]	1,485*** [0,000]	1,206** [0,023]
N	8.757	8.757	8.504	8.474	4.239
R ²	0,027	0,097	0,319	0,320	0,317
F-test (p-Werte)	0,902	0,330	0,205	0,544	0,954

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Abhängige Variable ist ein Indikator der angibt, ob die Gesamtnote über 6,5 liegt. Modelle (2) – (5) kontrollieren für Abschlussmonat-spezifische Effekte. F-test dokumentiert die p-Werte für Tests der Null-Hypothese, dass sich die Effekte für MI: 1. Generation und MI: ≥2.Generation nicht unterscheiden. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B16: Effekte der abgeleiteten Migrationsindikatoren auf Stufenerreichung 11,5 Punkte im 2. Examen (Analyse B).

MI: 1. Generation	-0,031*** [0,000]	-0,022*** [0,000]	0,001 [0,801]	-0,039*** [0,000]	-0,031* [0,079]
MI: ≥ 2. Generation	-0,026*** [0,000]	-0,022*** [0,000]	-0,001 [0,857]	-0,029*** [0,007]	-0,027* [0,055]
MI: Dt./Geb.Ausland	-0,025*** [0,003]	-0,018** [0,035]	-0,004 [0,685]	-0,003 [0,698]	0,014 [0,331]
Frauen		-0,017*** [0,000]	-0,003 [0,502]	-0,003 [0,480]	-0,007 [0,243]
Alter		-0,019*** [0,001]	-0,002 [0,754]	-0,002 [0,757]	0,002 [0,857]
Alter ²		0,000*** [0,004]	0,000 [0,676]	0,000 [0,691]	-0,000 [0,883]
Note 1. Examen			0,035*** [0,000]	0,035*** [0,000]	0,029*** [0,000]
Abiturnote					-0,015*** [0,001]
MI: Region 1				0,081** [0,012]	0,101 [0,108]
MI: Region 2				0,049*** [0,007]	0,046 [0,116]
MI: Region 3				0,033*** [0,003]	0,028* [0,077]
MI: Region 4				0,050*** [0,001]	0,034 [0,118]
MI: Region 5				0,018* [0,095]	0,020 [0,158]
MI: Region 6				0,010 [0,525]	-0,007 [0,677]
Konstante	0,041*** [0,000]	0,417*** [0,000]	-0,218** [0,033]	-0,218** [0,033]	-0,204 [0,308]
N	8.757	8.757	8.504	8.474	4.239
R ²	0,003	0,014	0,129	0,130	0,128
F-test (p-Werte)	0,557	0,972	0,741	0,394	0,781

Schätzungen mit linearem Wahrscheinlichkeitsmodell (LPM). Abhängige Variable ist ein Indikator der angibt, ob die Gesamtnote über 6,5 liegt. Modelle (2) – (5) kontrollieren für Abschlussmonat-spezifische Effekte. F-test dokumentiert die p-Werte für Tests der Null-Hypothese, dass sich die Effekte für MI: 1. Generation und MI: ≥ 2. Generation nicht unterscheiden. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B17: Effekte der Migrationsindikatoren auf die durchschnittliche schriftliche Note im 2. Examen (Analyse A).

	(1)	(2)	(3)	(4)	(5)
I1: Onomastik	-1,106*** [0,000]	-0,832*** [0,000]	-0,334*** [0,000]	-0,459** [0,020]	-0,476* [0,099]
I2: Geburtsort	-0,519*** [0,000]	-0,265*** [0,000]	-0,148*** [0,009]	-0,277*** [0,000]	-0,302*** [0,000]
I3: Staatsangeh.	-0,031 [0,815]	-0,010 [0,935]	0,054 [0,570]	0,028 [0,774]	0,040 [0,790]
Frauen		-0,353*** [0,000]	-0,015 [0,632]	-0,010 [0,744]	-0,004 [0,920]
Alter		-0,773*** [0,000]	-0,383*** [0,000]	-0,379*** [0,000]	-0,200*** [0,000]
Alter ²		0,007*** [0,000]	0,004*** [0,000]	0,004*** [0,000]	0,001** [0,018]
Note 1. Examen			0,745*** [0,000]	0,743*** [0,000]	0,710*** [0,000]
Abiturnote					-0,222*** [0,000]
MI: Region 1				0,460 [0,124]	0,593 [0,191]
MI: Region 2				0,207 [0,356]	0,147 [0,655]
MI: Region 3				-0,161 [0,431]	0,034 [0,912]
MI: Region 4				0,464** [0,025]	0,292 [0,341]
MI: Region 5				0,251 [0,239]	0,324 [0,290]
MI: Region 6				0,202 [0,390]	0,101 [0,769]
Konstante	6,251*** [0,000]	22,842*** [0,000]	8,423*** [0,000]	8,369*** [0,000]	5,799*** [0,000]
N	9.744	9.744	9.387	9.355	4.566
R ²	0,047	0,188	0,586	0,588	0,580

MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B18: Effekte der Migrationsindikatoren auf die durchschnittliche mündliche Note im 2. Examen (Analyse A).

	(1)	(2)	(3)	(4)	(5)
I1: Onomastik	-0,952*** [0,000]	-0,807*** [0,000]	-0,326*** [0,000]	-0,405*** [0,000]	-0,273 [0,288]
I2: Geburtsort	-0,495*** [0,000]	-0,282*** [0,008]	-0,032 [0,696]	-0,114 [0,174]	-0,218** [0,020]
I3: Staatsangeh.	0,026 [0,879]	0,096 [0,566]	0,084 [0,533]	0,207 [0,127]	0,230 [0,115]
Frauen		-0,460*** [0,000]	-0,240*** [0,000]	-0,130*** [0,001]	-0,128*** [0,002]
Alter		-0,654*** [0,000]	-0,055 [0,444]	-0,231*** [0,002]	-0,231*** [0,002]
Alter ²		0,007*** [0,000]	0,000 [0,958]	0,002* [0,077]	0,002* [0,078]
Note 1. Examen				0,770*** [0,000]	0,767*** [0,000]
MI: Region 1					0,071 [0,872]
MI: Region 2					0,093 [0,770]
MI: Region 3					-0,355 [0,198]
MI: Region 4					0,134 [0,629]
MI: Region 5					-0,372 [0,180]
MI: Region 6					0,078 [0,813]
Schriftliche Note (2. Examen)			0,810*** [0,000]		
Konstante	9,523*** [0,000]	23,257*** [0,000]	5,918*** [0,000]	8,454*** [0,000]	8,464*** [0,000]
N	8.756	8.756	8.756	8.503	8.473
R ²	0,027	0,092	0,483	0,448	0,448

MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B19: Effekte der Migrationsindikatoren auf durchschnittliche schriftliche Note im 2. Examen, getrennt nach Rechtsgebiet (Analyse A).

Rechtsgebiete:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Zivilrecht			Strafrecht			Öffentliches Recht		
I1: Onomastik	-1,131*** [0,000]	-0,329*** [0,000]	-0,366* [0,094]	-1,037*** [0,000]	-0,306*** [0,000]	-0,236 [0,447]	-1,145*** [0,000]	-0,369*** [0,000]	-0,867*** [0,000]
I2: Geburtsort	-0,501*** [0,000]	-0,112* [0,094]	-0,267*** [0,000]	-0,535*** [0,000]	-0,171** [0,027]	-0,225*** [0,008]	-0,535*** [0,000]	-0,198** [0,011]	-0,347*** [0,000]
I3: Staatsangeh.	0,030 [0,838]	0,121 [0,292]	0,090 [0,440]	-0,056 [0,716]	-0,014 [0,909]	-0,087 [0,488]	-0,130 [0,408]	-0,012 [0,926]	0,018 [0,892]
Frauen		-0,108*** [0,003]	-0,102*** [0,005]		0,070* [0,087]	0,073* [0,074]		0,086** [0,047]	0,088** [0,042]
Alter		-0,406*** [0,000]	-0,401*** [0,000]		-0,407*** [0,000]	-0,404*** [0,000]		-0,314*** [0,000]	-0,311*** [0,000]
Alter ²		0,004*** [0,000]	0,004*** [0,000]		0,004*** [0,000]	0,004*** [0,000]		0,003*** [0,000]	0,003*** [0,000]
Note 1. Examen		0,761*** [0,000]	0,759*** [0,000]		0,693*** [0,000]	0,690*** [0,000]		0,768*** [0,000]	0,767*** [0,000]
MI: Region 1			0,357 [0,262]			0,326 [0,478]			0,871** [0,037]
MI: Region 2			0,192 [0,441]			-0,127 [0,709]			0,572* [0,054]
MI: Region 3			-0,288 [0,209]			-0,330 [0,304]			0,262 [0,300]
MI: Region 4			0,443* [0,058]			0,076 [0,812]			0,895*** [0,001]
MI: Region 5			0,220 [0,360]			0,158 [0,633]			0,404 [0,134]
MI: Region 6			-0,024 [0,928]			0,150 [0,692]			0,707** [0,022]
Konstante	6,365*** [0,000]	8,868*** [0,000]	8,797*** [0,000]	6,004*** [0,000]	9,020*** [0,000]	8,987*** [0,000]	6,288*** [0,000]	6,945*** [0,000]	6,905*** [0,000]
N	9.696	9.364	9.332	9.696	9.364	9.332	9.696	9.364	9.332
R ²	0,039	0,512	0,514	0,034	0,409	0,410	0,035	0,429	0,431

MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B20: Effekte der abgeleiteten Migrationsindikatoren auf die durchschnittliche schriftliche Note im 2. Examen (Analyse B).

	(1)	(2)	(3)	(4)	(5)
MI: 1. Generation	-1,422*** [0,000]	-0,934*** [0,000]	-0,391*** [0,000]	-0,568*** [0,001]	-0,601** [0,020]
MI: ≥ 2. Generation	-1,295*** [0,000]	-0,979*** [0,000]	-0,389*** [0,000]	-0,349** [0,043]	-0,249 [0,321]
MI: Dt. / Geb. Ausland	-0,958*** [0,000]	-0,617*** [0,000]	-0,311*** [0,000]	-0,314*** [0,000]	-0,268** [0,016]
Frauen		-0,347*** [0,000]	-0,012 [0,704]	-0,010 [0,754]	-0,005 [0,899]
Alter		-0,772*** [0,000]	-0,381*** [0,000]	-0,378*** [0,000]	-0,201*** [0,000]
Alter ²		0,007*** [0,000]	0,004*** [0,000]	0,004*** [0,000]	0,001** [0,018]
Note 1. Examen			0,745*** [0,000]	0,743*** [0,000]	0,710*** [0,000]
Abiturnote					-0,224*** [0,000]
MI: Region 1				0,334 [0,236]	0,385 [0,369]
MI: Region 2				0,062 [0,756]	-0,041 [0,890]
MI: Region 3				-0,280 [0,121]	-0,177 [0,507]
MI: Region 4				0,310* [0,088]	0,116 [0,671]
MI: Region 5				0,137 [0,474]	0,121 [0,660]
MI: Region 6				0,081 [0,703]	-0,113 [0,715]
Konstante	6,268*** [0,000]	22,823*** [0,000]	8,392*** [0,000]	8,363*** [0,000]	5,820*** [0,000]
N	9.744	9.744	9.387	9.355	4.566
R ²	0,050	0,190	0,586	0,588	0,580
F-test (p-Werte)	0,234	0,638	0,972	0,016	0,007

F-test dokumentiert die p-Werte für Tests der Null-Hypothese, dass sich die Effekte für MI: 1. Generation und MI: ≥ 2. Generation nicht unterscheiden. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B21: Effekte der abgeleiteten Migrationsindikatoren auf die durchschnittliche mündliche Note im 2. Examen (Analyse B).

	(1)	(2)	(3)	(4)	(5)
MI: 1. Generation	-1,168*** [0,000]	-0,821*** [0,000]	-0,222** [0,014]	-0,290*** [0,002]	-0,277 [0,236]
MI: ≥ 2. Generation	-1,196*** [0,000]	-1,019*** [0,000]	-0,427*** [0,000]	-0,538*** [0,000]	-0,480** [0,032]
MI: Dt. / Geb. Ausland	-1,008*** [0,000]	-0,744*** [0,000]	-0,256** [0,036]	-0,451*** [0,000]	-0,452*** [0,000]
Frauen		-0,454*** [0,000]	-0,237*** [0,000]	-0,125*** [0,002]	-0,123*** [0,003]
Alter		-0,643*** [0,000]	-0,050 [0,483]	-0,224*** [0,003]	-0,224*** [0,003]
Alter ²		0,007*** [0,000]	-0,000 [0,993]	0,002* [0,092]	0,002* [0,092]
Note 1. Examen				0,768*** [0,000]	0,767*** [0,000]
MI: Region 1					0,170 [0,680]
MI: Region 2					0,070 [0,811]
MI: Region 3					-0,204 [0,399]
MI: Region 4					0,028 [0,911]
MI: Region 5					-0,170 [0,500]
MI: Region 6					0,221 [0,465]
Schriftliche Note (2. Examen)			0,809*** [0,000]		
Konstante	9,542*** [0,000]	23,085*** [0,000]	5,857*** [0,000]	8,347*** [0,000]	8,364*** [0,000]
N	8.756	8.756	8.756	8.503	8.473
R ²	0,030	0,095	0,484	0,449	0,449
F-test (p-Werte)	0,850	0,158	0,066	0,029	0,167

F-test dokumentiert die p-Werte für Tests der Null-Hypothese, dass sich die Effekte für MI: 1. Generation und MI: ≥ 2. Generation nicht unterscheiden. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.

Tabelle B22: Effekte der abgeleiteten Migrationsindikatoren auf Mittelwert schriftliche Note im 2. Examen getrennt nach Rechtsgebiet. (Analyse B).

Rechtsgebiete:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Zivilrecht			Strafrecht			Öffentliches Recht	
MI: 1. Generation	-1,401*** [0,000]	-0,329*** [0,000]	-0,461** [0,018]	-1,413*** [0,000]	-0,440*** [0,000]	-0,379 [0,144]	-1,481*** [0,000]	-0,462*** [0,000]	-0,963*** [0,000]
MI: ≥ 2. Generation	-1,320*** [0,000]	-0,375*** [0,000]	-0,232 [0,242]	-1,193*** [0,000]	-0,336*** [0,000]	-0,183 [0,486]	-1,369*** [0,000]	-0,468*** [0,000]	-0,753*** [0,000]
MI: Dt. / Geb. Ausl.	-0,968*** [0,000]	-0,281*** [0,004]	-0,285*** [0,003]	-0,891*** [0,000]	-0,254** [0,029]	-0,258** [0,026]	-1,028*** [0,000]	-0,434*** [0,000]	-0,436*** [0,000]
Frauen		-0,105*** [0,004]	-0,101*** [0,005]		0,071* [0,082]	0,073* [0,075]		0,091** [0,037]	0,089** [0,040]
Alter		-0,403*** [0,000]	-0,401*** [0,000]		-0,406*** [0,000]	-0,404*** [0,000]		-0,310*** [0,000]	-0,310*** [0,000]
Alter ²		0,004*** [0,000]	0,004*** [0,000]		0,004*** [0,000]	0,004*** [0,000]		0,003*** [0,000]	0,003*** [0,000]
Note 1. Examen		0,760*** [0,000]	0,758*** [0,000]		0,693*** [0,000]	0,690*** [0,000]		0,767*** [0,000]	0,766*** [0,000]
MI: Region 1			0,217 [0,472]			0,259 [0,541]			0,718* [0,068]
MI: Region 2			0,045 [0,845]			-0,210 [0,472]			0,368 [0,166]
MI: Region 3			-0,418** [0,044]			-0,400 [0,142]			0,122 [0,570]
MI: Region 4			0,295 [0,163]			-0,021 [0,939]			0,665*** [0,004]
MI: Region 5			0,109 [0,623]			0,068 [0,812]			0,264 [0,262]
MI: Region 6			-0,161 [0,507]			0,087 [0,798]			0,564** [0,043]
Konstante	6,383*** [0,000]	8,834*** [0,000]	8,790*** [0,000]	6,018*** [0,000]	9,009*** [0,000]	8,990*** [0,000]	6,306*** [0,000]	6,899*** [0,000]	6,888*** [0,000]
N	9.696	9.364	9.332	9.696	9.364	9.332	9.696	9.364	9.332
R ²	0,042	0,512	0,514	0,035	0,409	0,410	0,037	0,430	0,431

F-test dokumentiert die p-Werte für Tests der Null-Hypothese, dass sich die Effekte für MI: 1. Generation und MI: ≥2.Generation nicht unterscheiden. MI Regionen: (1) Afrika, (2) Süd/Ost- und Zentral-Asien, (3) West-Asien, (4) Ost-Europa, (5) Süd-Europa, (6) Andere/Unklare Herkunft, Referenzkategorie: Zentral- und Nord-Europa. Robuste p-Werte in eckigen Klammern; Signifikanz-Niveau: *** p<0,01; ** p<0,05; * p<0,1.