

# Statistische Korrekturen verbessern die Aussagekraft von Bachelornoten

## Statistical corrections improve the validity of bachelor grades

Working Paper, 2019

Jan Bergerhoff <sup>1 2</sup>, Luisa Bergholz <sup>3</sup>, Philipp Seegers <sup>1 2, 3</sup>, Stephan Stegt <sup>3</sup>

### Zusammenfassung

Aufgrund mangelnder Vergleichbarkeit eignen sich absolute Noten nur bedingt für die Zulassung von Personen mit Bachelorabschluss in Master-Programmen. Aus diesem Grund nutzt eine zunehmende Anzahl von Hochschulen Eignungstests für die Auswahl von Studierenden. Im Rahmen dieser Publikation wird evaluiert, in wie weit eine statistische Korrektur auf Basis von Kontextinformationen die Aussagekraft von Bachelornoten verbessern kann. Hierzu wurde eine Stichprobe mit 4.844 TM-WISO-Testergebnissen um CASE Scores erweitert. Der CASE Score berücksichtigt neben der programmspezifischen Notenvergabe ebenfalls IQ- und Persönlichkeitsergebnisse von Studierenden. Die Ergebnisse zeigen, dass eine statistische Korrektur die Aussagekraft von Bachelornoten erheblich verbessern kann: Die Korrelation zwischen Bachelornoten und TM-WISO steigt von  $r=0,27$  auf  $r=0,42$ , wenn eine Korrektur durch CASE vorgenommen wird.

### Abstract

Different grading standards as well as varying competitiveness lead to low comparability between higher education grades. As a result, absolute grade point averages (GPAs) are of limited use when admitting Bachelor graduates to Master programmes and an increasing number of higher education institutions relies on admission tests to select fitting applicants. This study analyses to which degree a statistical correction based on relevant context information can improve the explanatory power of GPAs achieved during Bachelor studies. The analysis is based on a unique sample of 4,844 completed TM-WISO admission tests. This sample has been enriched with CASE Scores based on the higher education institution at which the Bachelor degree was obtained, the subject studied, the year of graduation and the final GPA. The CASE Score statistically corrects those GPAs by using (1) programme-specific information on grading standards and (2) IQ and personality scores of students in those study-programmes. The results show that a statistical correction can significantly increase the explanatory power of GPAs. The correlation between TM-WISO test scores and Bachelor grades increases from  $r=-0.27$  to  $r=-0.42$  when GPAs are corrected by the CASE Score. When adjusting only for the respective grading standards but not for differences in competitiveness the correlation increases slightly from  $r=-0.27$  to  $r=-0.31$ .

**Keywords:** Masterzulassung, Abschlussvergleich, Notenkorrektur, Studierfähigkeit

**Keywords:** Master's admission, Degree comparison, Correction of grades, Ability to study

---

<sup>1</sup> Maastricht University, School of Business and Economics, Economics of Education

<sup>2</sup> candidate select GmbH (CASE)

<sup>3</sup> ITB Consulting GmbH

## 1. Einleitung

Absolute Bewertungsmaßstäbe unterliegen einem größeren Mess- und Interpretationsfehler als relative Vergleiche. Dieses Problem zeigt sich deutlich mit Blick auf die Notengebung in Bachelorstudiengängen in Deutschland. Während die Rangfolge innerhalb eines Studienprogramms<sup>4</sup> durchaus als treffender Indikator für eine erbrachte Leistung im Vergleich zu den direkten Mitstudierenden zu sehen ist, sind programmübergreifende Vergleiche basierend auf der absoluten Abschlussnote aus mehreren Gründen problematisch.

Zunächst setzt ein absolutes Notensystem keine verbindlichen Standards hinsichtlich der Nutzung der Skala. So schrieb der Wissenschaftsrat im Jahr 2012: „Arbeitgebern wird empfohlen, die fach- und hochschulspezifische Benotungspraxis bei der Einschätzung einer Prüfungsnote zu berücksichtigen. Nur dann erhalten sie ein aussagekräftiges Bild von der Leistungsfähigkeit der Bewerberin bzw. des Bewerbers“ (Wissenschaftsrat, 2012). Die Unterschiede mit Blick auf die genutzten Notenstandards in Bachelorprogrammen in Deutschland sind dabei groß. Berücksichtigt man lediglich große Programme<sup>5</sup>, so variiert der Mittelwert der Abschlussnoten je Programm zwischen 1,27 und 3,25.<sup>6</sup>

Zweitens ist der Wettbewerb um gute Noten ebenfalls abhängig vom Studienprogramm. In weniger selektiven Programmen kann es sein, dass im Gesamtkontext betrachtet eher schwache Leistungen schon für eine durchschnittliche Platzierung innerhalb des genutzten Notenspektrums ausreichen. In Programmen mit harten Zulassungskriterien<sup>7</sup> ist allein aufgrund der initialen Selektion von einem größeren Wettbewerb auszugehen.

Beide Effekte führen dazu, dass absolute Noten nur begrenzt vergleichbar sind. Hieraus resultieren eine Reihe von Problemen mit Blick auf Auswahlverfahren, welche sich auf ebenjene Noten stützen. Beispiele hierfür finden sich im Hochschulbereich bei der Vergabe von Masterplätzen wie auch am Arbeitsmarkt bei der Rekrutierung von Personen mit Abschluss. Um dennoch die notwendige Vergleichbarkeit zwischen Bewerbern herstellen zu können, haben sich vor allem Testverfahren in solchen Situationen etabliert.

Der Nutzen und die Wirksamkeit von Testverfahren konnten bereits in vielen Studien belegt werden (Hell, Trapmann & Schuler, 2008; Schult, Hofmann & Stegt, angenommen). Richtig angewandt sind Testverfahren objektiv und weisen bei Studierenden eine hohe Akzeptanz auf (Stegt, Didi, Zimmerhofer & Seegers, 2018). Allerdings lösen Testverfahren nicht das eigentliche Problem der mangelnden Vergleichbarkeit von Studienleistungen. Stattdessen führen sie eine weitere, besser vergleichbare Leistungsmessung ein.

Diese Publikation untersucht, in wie weit eine statistische Korrektur von Noten in der Lage ist, die Vergleichbarkeit von Studienleistungen zu verbessern. Hierzu wurden die Bachelorabschlüsse von 4.844 TM-WISO-Teilnehmenden mit Hilfe des CASE Algorithmus analysiert, welcher auf Basis von Kontextinformationen Abschlüsse vergleichbarer machen soll.<sup>8</sup> Die Personen hatten den fachspezifischen Studierfähigkeitstest TM-WISO (Test für Masterstudiengänge in Wirtschafts- und Sozialwissenschaften) im Zuge der Bewerbung für einen Masterstudienplatz im Bereich Wirtschaftswissenschaften bearbeitet. Somit kann evaluiert werden, wie stark der Zusammenhang zwischen Bachelornoten bzw. korrigierten Bachelornoten und einem etablierten studienfachspezifischen Eignungstest ist.

Die von CASE für die Korrektur von Abschlussnoten genutzte Methodik, der TM-WISO und die genutzten Daten werden im nächsten Abschnitt beschrieben. Im dritten Abschnitt werden die Ergebnisse aufgezeigt.

## 2. Methodik und Stichprobe

Die Stichprobe basiert auf anonymisierten Daten von TM-WISO-Teilnehmenden. Dieser Datensatz wurde durch CASE Scores angereichert, indem die Bachelorabschlüsse durch den CASE Algorithmus bewertet wurden. Insgesamt liegen Informationen zu 4.844 Personen mit Bachelorabschluss vor, welche sich in den Jahren 2010

---

<sup>4</sup> Kombination aus Hochschule, Studienfach, Abschlussart und Abschlussjahr

<sup>5</sup> Programme mit mindestens 100 Abschlüssen pro Jahr

<sup>6</sup> Mit Blick auf die deutsche Notenskala von 1 (sehr gut) bis 4 (ausreichend)

<sup>7</sup> z.B. Numerus Clausus oder Zulassungstest

<sup>8</sup> Eine ausführliche Erklärung zum CASE Algorithmus folgt in Kapitel 2.1

bis 2016 für ein Masterstudium im Bereich der Wirtschaftswissenschaften an einer den TM-WISO für die Platzvergabe nutzenden Hochschule beworben haben.<sup>9</sup>

Diese Selektivität in den Zielhochschulen der Studierenden führt dazu, dass die erworbenen Bachelorabschlüsse in der Stichprobe nicht repräsentativ für die Gesamtzahl der in Deutschland im Bereich der Wirtschaftswissenschaften erworbenen Bachelorabschlüsse sind.<sup>10</sup> Diese Verzerrung stellt allerdings kein größeres Problem für die Analyse dar, da die Größe des Datensatzes eine ausreichende Variabilität in den Abschlusshochschulen erlaubt und in der Realität auch keine Selektionsentscheidungen auf Basis einer zur gesamten Studierendenschaft repräsentativen Bewerbergruppe zu erwarten sind.<sup>11</sup>

In der Stichprobe liegen Angaben zu dem Bachelorabschluss (Hochschule, Studienfach, Abschlussjahr und Abschlussnote), der Abiturnote und den TM-WISO Testergebnissen, bestehend aus Gesamtergebnis und Ergebnissen für die vier Untertests, vor. Die Informationen zum Bachelorabschluss basieren auf freiwilligen Angaben der TM-WISO-Teilnehmenden, die diese bei der Anmeldung zum Test machen. Die Abschlussinformationen wurden durch den CASE Algorithmus interpretiert und der Datensatz um drei weitere Metriken angereichert: (1) Lokales Notenperzentil im Studienprogramm, (2) CASE Subject Score und (3) CASE Score. Deskriptive Statistiken zu dem Datensatz finden sich in Tabelle 1.

»Tabelle 1 hier einfügen«

## 2.1 Korrektur von Abschlussnoten durch CASE

Durch CASE werden Abschlussnoten im jeweiligen Studienkontext interpretiert. Zuerst wird die Endnote eines Abschlusses mit der Notenverteilung, die im Abschlussjahr im spezifischen Studienfach und Abschluss an der jeweiligen Hochschule vergeben wurde, verglichen. Die relative Position einer Note zu direkten Vergleichsnoten wird in einem zweiten Schritt nachjustiert, um der unterschiedlichen Leistungsdichte in verschiedenen Studiengängen an unterschiedlichen Hochschulen Rechnung zu tragen. Die relative Gewichtung dieser beiden Faktoren wurde durch CASE auf Basis von Lebenslaufdaten optimiert.

Die Einschätzung der lokalen Leistungsfähigkeit basiert auf der Notenverteilung, die im Abschlussjahr in der jeweiligen Fachrichtung an der spezifischen Universität vergeben wurde. Hierbei handelt es sich um die amtlichen Prüfungsergebnisse aus den Prüfungsämtern der Hochschulen. Für die Berechnung der Notenverteilungen wird ein gewichteter Durchschnitt genutzt, der Werte von bis zu drei zurückliegenden Jahren miteinbezieht. Hierdurch werden die Werte geglättet und die Analyse zukünftiger Abschlüsse ermöglicht. Die stetige Notenverteilung wird mithilfe einer schrittweisen kubischen Polynominterpolation angenähert. Die Information zum lokalen Notenperzentil entspricht der Position einer Abschlussnote in der jeweiligen Notenverteilung.

Um einen Vergleich zwischen Abschlüssen verschiedener Hochschulen, Studienfächer und Abschlussarten zu ermöglichen, wird die Leistungsdichte je Studienprogramm geschätzt. Dabei werden unter anderem Daten zur Intelligenz und zur Persönlichkeit der Studierenden in verschiedenen Studienprogrammen herangezogen. Die Daten hierzu stammen aus der Studienreihe „Fachkraft 2030“, welche in insgesamt 13 Erhebungen seit September 2012 knapp 300.000 Studierende befragt hat (Bergerhoff, Hartmann, Seegers, Knappe, 2018). Persönlichkeit wird dabei auf Basis eines 50-Item Big5 Fragebogens aus dem International Personality Item Pool erhoben (IPIP)<sup>12</sup>, welcher auf Goldberg (1999) zurückgeht. Zur Einschätzung der kognitiven Fähigkeiten wird

<sup>9</sup> Den TM-WISO nutzten in diesem Zeitraum die folgenden Hochschulen: (1) RWTH Aachen, (2) Universität Hamburg, (3) HAW Hamburg, (4) Universität zu Köln, (5) Kühne Logistics University, (6) Leuphana Universität Lüneburg, (7) OVGU Magdeburg, (8) Hochschule Neu-Ulm und (9) WHU

<sup>10</sup> Tabelle 1 zeigt, dass die Abschlüsse in der Bewertung durch den CASE Subject Score im Durchschnitt leicht besser sind als das erwartete Populationsmittel aller wirtschaftswissenschaftlichen Abschlüsse.

<sup>11</sup> Die Stichprobe beinhaltet Bachelorabschlüsse von insgesamt 186 deutschen Hochschulen. Für 125 dieser Hochschulen sind mehr als drei Bachelorabschlüsse in der Stichprobe vorhanden.

<sup>12</sup> International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences (<http://ipip.ori.org/>). Internet Website.

ein aus drei Untertests bestehender und von CASE in Kooperation mit der Universität Bonn und der Maastricht University entwickelter IQ-Test verwendet.<sup>13</sup>

Basierend auf den psychometrischen Testergebnissen der Studierenden wird für jedes Studienprogramm der erwartete Arbeitsmarkterfolg der Studierenden geschätzt. Der Zusammenhang ist in der Arbeitsmarktökonomie gut beschrieben. Konkret werden die Korrelate aus der Studie von Almlund, Duckworth, Heckman & Kautz (2011) genutzt. Auf Basis der Ergebnisse wird ein Ranking der verschiedenen Studienprogramme erstellt.

Für die finale Korrektur der Abschlussnote werden die zwei Informationen Position in der lokalen Notenverteilung und erwarteter Arbeitsmarkterfolg der Studierenden miteinander kombiniert. Die Gewichtung der verschiedenen Skalen wird dabei durch CASE basierend auf rund 100.000 realen Lebenslaufdaten optimiert. Hierdurch wird die Frage beantwortet, wie ein lokal guter Abschluss einer weniger kompetitiven Institution im Verhältnis zu einem lokal weniger guten Abschluss einer kompetitiven Institution einzuschätzen ist.

Aus dieser Analyse ergeben sich drei Metriken, welche einen Abschluss mit Bezug auf unterschiedliche Vergleichsgruppen einschätzen:

- (1) *Lokales Notenperzentil*: Die genaue Position eines Abschlusses in der Notenverteilung des jeweiligen Studienprogramms (Hochschule, Fach, Abschluss, Jahr) basierend auf den amtlichen Prüfungsergebnissen.
- (2) *CASE Subject Score*: Eine Einschätzung, wie gut ein Abschluss im Vergleich zu allen Abschlüssen der gleichen Fachrichtung in Deutschland ist, basierend auf einem Vergleich verschiedener Studienprogramme durch IQ- und Persönlichkeitstests aus der Studienreihe „Fachkraft 2030“.
- (3) *CASE Score*: Eine Einschätzung, wie gut ein Abschluss im Vergleich zu allen Abschlüssen deutschlandweit ist, ebenfalls basierend auf einem Vergleich zwischen Studienprogrammen durch ebenjene IQ- und Persönlichkeitstestwerte.

## 2.2 Fachspezifischer Studierfähigkeitstest TM-WISO

Der TM-WISO wurde im Jahr 2009 in Kooperation mit der WISO-Fakultät der Universität Hamburg entwickelt. Die Entwicklung basiert auf einer Anforderungsanalyse nach der Methode der kritischen Ereignisse (Flanagan, 1954). Der Test setzt sich aus insgesamt vier Aufgabengruppen zusammen:

### **Planen in Studium und Beruf (PSB):**

Die Aufgabengruppe prüft Organisations- und Planungskompetenzen, die in Studium und Beruf wichtig sind. Mehrere komplexe Szenarien (z. B. Projektarbeit, empirische Studie) werden geschildert, in denen verschiedene Tätigkeiten geplant und organisiert werden müssen.

### **Texte analysieren (TA):**

Mit dieser Aufgabengruppe wird die Kompetenz, komplexe, fachbezogene Texte zu strukturieren, zu verstehen und richtige Schlussfolgerungen abzuleiten, geprüft. Die etwa halbseitigen Texte beziehen sich auf wirtschafts- und sozialwissenschaftliche Sachverhalte.

### **Wirtschaftliche Zusammenhänge formalisieren (WZF):**

Hier wird quantitatives Problemlösen bei wirtschaftsbezogenen Fragestellungen untersucht. Verbal beschriebene Sachverhalte müssen durchdrungen und in eine mathematische Darstellung übertragen oder aus mehreren Modellgleichungen inhaltliche Schlussfolgerungen gezogen werden.

### **Wirtschaftsgrafiken interpretieren (WGI):**

Bei diesem Aufgabentyp geht es um die Kompetenz zur Interpretation von Diagrammen und Tabellen, wie sie auch in wirtschaftswissenschaftlichen Lehrbüchern und Publikationen zu finden sind. Dies erfordert sowohl quantitative als auch verbale Fähigkeiten und die Integration von grafisch dargebotenen Informationen mit kurzen Begleittexten.

---

<sup>13</sup> Der Test setzt sich aus drei Untertests zusammen: (1) Matrix Items (Fluide Intelligenz, Kultur-Neutral; vergleichbar mit dem Raven Matrices Test), (2) Spatial Puzzles (Räumliches Denken, Kultur-Neutral; vergleichbar mit einem Untertest aus dem WAIS-IV), (3) Quiz Items (Kristalline Intelligenz, Sprachverständnis, Memory; Vergleichbar mit einem Untertest aus dem WAIS-IV)

Der Test wird online in deutscher oder englischer Sprache unter Aufsicht deutschlandweit in zertifizierten Testzentren sowie in Einzelfällen im Ausland durchgeführt. Jährlich nehmen ca. 2.000 Personen am TM-WISO teil. Die (Weiter-)Entwicklung, Durchführung, Auswertung und Evaluation erfolgen nach nationalen und internationalen Standards der Testentwicklung (Kersting & Hornke, 2003). Die Reliabilität des Gesamttests liegt bei .87 (Testhalbierungsreliabilität mit Spearman-Brown-Korrektur). Zudem konnte im Längsschnitt eine gute Prognosekraft für den späteren Studienerfolg (Abschlussnoten im Masterstudium BWL oder VWL) nachgewiesen werden (Stegt & Bergholz, 2018). Alle Aufgabengruppen sagen die Abschlussnoten gut vorher, die stärkste Prognosekraft haben die Aufgabengruppen PSB und WZF.

### 3. Ergebnisse

Die Untersuchung der vorliegenden Stichprobe konzentriert sich auf die Analyse des Zusammenhangs zwischen (korrigierten) Bachelorabschlussnoten und TM-WISO. Hierzu wird im ersten Schritt eine einfache Korrelationsanalyse vorgenommen. In einem zweiten Schritt wird mit Hilfe von multiplen Regressionsanalysen untersucht, in wie weit die Voraussagekraft von Bachelornoten in Kombination mit Abiturnoten steigt, wenn die Bachelornoten korrigiert werden.

#### 3.1 Korrelationsanalyse

Das Gesamtergebnis im TM-WISO und alle vier Aufgabengruppen weisen eine signifikante Korrelation mit Abiturnote und Bachelorabschlussnote auf.<sup>14</sup> Die Korrelation zwischen TM-WISO-Ergebnis und Abiturnoten ( $r=-0,32$ ) ist stärker ausgeprägt als die Korrelation mit Bachelorabschlussnoten ( $r=-0,27$ ). Dies ist überraschend: Da es sich beim TM-WISO um einen fachspezifischen Studierfähigkeitstest handelt und die Abiturnoten eine weiter in der Vergangenheit liegende Leistung beschreiben, könnte ein stärkerer Zusammenhang des TM-WISO mit Bachelornoten erwartet werden. Mögliche Gründe für den stärkeren Zusammenhang mit Abiturnoten könnten ein im Vergleich zu Bachelornoten höherer Grad an Standardisierung sowie eine größere Homogenität der Lehrinhalte sein.

Eine Korrektur der Bachelorabschlussnoten durch CASE führt zu einem klaren Anstieg der Korrelation mit dem TM-WISO ( $r=-0,42$ ).<sup>15</sup> Dieses Ergebnis bestätigt sich auch für alle Aufgabengruppen: Korrigierte Bachelornoten weisen auch hier grundsätzlich höhere Korrelationen als unkorrigierte Bachelornoten bzw. die Abiturnote auf. Zusätzlich steigt durch die Korrektur auch die Korrelation zwischen dem Abschneiden im Bachelor und Abiturnoten ( $r=-0,33$  unkorrigiert,  $r=-0,42$  korrigiert).

Dies deutet darauf hin, dass die Korrektur sowohl den Anteil der im Bachelor erbrachten Leistungen, welcher im TM-WISO gemessen wird, als auch den Teil der Leistungen, welche mit dem Abschneiden im Abitur zusammenhängen, besser darstellen kann als die unkorrigierte Bachelornote. Die Gründe hierfür wurden eingangs bereits beschrieben. Unterschiede in Notenstandards und Programmselektion führen dazu, dass Noten nicht nur die Leistung im Studium widerspiegeln, sondern, statistisch gesprochen, auch eine programmspezifische Noise-Komponente.

»Tabelle 2 hier einfügen«

Bei den Untertests sind die Ergebnisse mit Blick auf die beiden für den Masterabschluss validesten Aufgabengruppen PSB und WZF besonders interessant: Das Ergebnis in der Aufgabengruppe WZF korreliert hoch mit Abiturnoten, mit den korrigierten Bachelornoten sowie mit Masterabschlussnoten. Außerdem hat hier die CASE Korrektur den größten Effekt. Die mit dieser Aufgabengruppe gemessenen quantitativen Fähigkeiten scheinen für den Erfolg in allen Ausbildungsabschnitten (Abitur, Bachelorstudium und Masterstudium)

<sup>14</sup> Da niedrige Noten (1: sehr gut) besser sind als hohe Noten (4: ausreichend), ist die Korrelation wie erwartet negativ.

<sup>15</sup> CASE Scores spiegeln die Perzentile in der Gesamtverteilung aller Studierenden wieder. Niedrige Werte (z.B. Top 10%) sind besser als höhere Werte (z.B. Top 50%).

besonders relevant zu sein. Das Ergebnis in der Aufgabengruppe PSB hingegen weist die niedrigsten Korrelationen zu Abitur- und Bachelornoten auf und die Korrektur hat nur einen kleinen Effekt. Der Zusammenhang mit der Masterabschlussnote hingegen ist sehr hoch. Dies könnte bedeuten, dass die Fähigkeiten zum Planen und Organisieren im Masterstudium eine deutlich wichtigere Rolle spielen als in früheren Bildungsabschnitten (Abitur und Bachelorstudium).

Ein weiteres interessantes Ergebnis offenbart der Blick auf das ebenfalls durch CASE bestimmte Leistungsperzentil innerhalb des jeweiligen Studienprogramms einer Hochschule. Hierbei werden bereits Unterschiede in der Benotung zwischen verschiedenen Studienprogrammen berücksichtigt, allerdings noch keine Vergleiche hinsichtlich der Leistungsfähigkeit der Studierenden in den verschiedenen Programmen angestellt. Im Vergleich zu unkorrigierten Bachelornoten steigt die Korrelation leicht an ( $r=-0,27$  unkorrigiert,  $r=-0,31$  lokales Perzentil). Eine lokale Notenkorrektur kann somit die Aussagekraft leicht steigern, stellt allerdings für sich genommen keine große Verbesserung dar.

Tabelle 2 listet alle paarweisen Korrelationen auf. Abbildungen 1 und 2 zeigen Streudiagramme mitsamt einer Regressionslinie zwischen korrigierten Bachelornoten (CASE Score) und dem TM-WISO samt Untertests.

»Abbildung 1 hier einfügen«

»Abbildung 2 hier einfügen«

### 3.2 Multiple Regressions-Analyse

Die Analyse der Korrelationen liefert eindeutige Ergebnisse hinsichtlich der möglichen Verbesserung der Aussagekraft von Bachelornoten durch eine auf Kontextinformationen basierende statistische Korrektur. Die simultane Analyse von Abitur- und Bachelornoten hinsichtlich des Zusammenhangs mit den TM-WISO Testergebnissen ermöglicht zweierlei: Erstens eine Analyse, ob die Aussagekraft von Abiturnoten und Bachelornoten insgesamt durch eine Korrektur der Bachelornoten gesteigert werden kann. Zweitens einen Vergleich, in wie weit die Kombination von Abitur- und (korrigierten) Bachelornoten eine bessere Voraussage hinsichtlich des TM-WISO liefern kann als deren individuelle Betrachtung.

Bezüglich der ersten Fragestellung zeigt sich, dass auch in Kombination mit der Abiturnote die Aussagekraft von Bachelornoten stark von einer statistischen Korrektur profitiert ( $R^2=0,13$  unkorrigiert und  $R^2=0,19$  korrigiert). Dieses Ergebnis zeigt sich neben dem TM-WISO-Gesamtergebnis auch bei allen Aufgabengruppen.

Mit Blick auf die zweite Fragestellung zeigt sich, dass sich die Aussagekraft von unkorrigierten Bachelornoten durch die Kombination mit Abiturnoten eindeutig steigern lässt ( $r=-0,27 \rightarrow R^2=0,07$  nur Bachelornote und  $R^2=0,13$  Bachelornote in Kombination mit Abiturnote). Der Determinationskoeffizient liegt hier allerdings dennoch unterhalb des Determinationskoeffizienten, der bei einer ausschließlichen Betrachtung korrigierter Bachelornoten zu finden ist ( $r=-0,42 \rightarrow R^2=0,18$ ). Entsprechend steigt die Voraussagekraft hinsichtlich des TM-WISO auch nur unwesentlich, wenn neben korrigierten Bachelornoten auch die Abiturnote betrachtet wird ( $R^2=0,19$ ).

Bei den einzelnen Aufgabengruppen zeigt sich ein ähnliches Bild wie bereits bei der Analyse der Korrelationen: WZF lässt sich mithilfe der Abiturnoten und der Bachelornoten gut vorhersagen und die Korrektur der Bachelornoten bringt eindeutigen Zugewinn in der Prognosegüte. PSB hängt hingegen schwächer als die anderen Aufgabengruppen mit Abitur- und Bachelornoten zusammen, selbst nach Korrektur der Bachelornoten. TA und WGI hängen etwas weniger als WZF mit Abitur- und Bachelornoten zusammen und die Korrektur hat mittelgroße Auswirkungen.

Tabelle 3 stellt die Ergebnisse der Regressionsanalyse für den TM-WISO Gesamtscore und die TM-WISO Untertests dar.

»Tabelle 3 hier einfügen«

## 4. Fazit

Auf Grund der eingeschränkten Vergleichbarkeit von Bachelorabschlüssen spiegeln Abschlussnoten nur unzureichend die erbrachte Studienleistung wider. Die vorliegende Publikation untersucht, ob eine statistische Korrektur von Bachelornoten unter Betrachtung des jeweiligen Studienkontextes die Aussagekraft verbessern kann. Hierzu werden unkorrigierte, bzw. korrigierte Bachelorabschlussnoten mit den Testergebnissen aus dem TM-WISO verglichen.

Die Ergebnisse zeigen, dass sich die Voraussagekraft von Bachelornoten klar erhöht, wenn diese für den jeweiligen Studienkontext korrigiert werden. Die Korrelation von unkorrigierten Bachelornoten und TM-WISO beträgt  $r=0,27$ . Mit Blick auf korrigierte Abschlussnoten steigt diese Korrelation auf  $r=0,42$  an. Eine Multiple Regressionsanalyse zeigt, dass die Korrektur von Bachelorabschlussnoten allein bereits eine bessere Voraussage hinsichtlich des TM-WISO bietet, als wenn neben der unkorrigierten Bachelornote zusätzlich noch absolute Abiturnoten genutzt werden. Als Folge der Korrektur der Bachelornoten steigt auch der Zusammenhang mit Abiturnoten merklich an ( $r=0,33$  unkorrigiert,  $r=0,42$  korrigiert).

Dies zeigt, dass der Zusammenhang zwischen Studienleistungen im Bachelor und dem Abschneiden in einem fachspezifischen Studierfähigkeitstest, bzw. im Rahmen des Abiturs, stärker ist als ein erster Blick auf die Bachelornoten vermuten lässt. Inkohärente Notenstandards, wie auch Unterschiede in der Kompetitivität verschiedener Studienprogramme produzieren statistisch gesehen ein Rauschen, wodurch sich die niedrigere Korrelation unmittelbar erklären lässt. Außerdem wird deutlich, dass Bachelornoten Informationen über einige Kompetenzen enthalten, die auch im Masterstudium wichtig sind. Ein Unterschied zwischen Bachelor und Master scheint allerdings darin zu bestehen, dass die Kompetenz zum Planen in Studium und Beruf für Masterstudiengänge eine deutlich höhere Rolle spielt.

Gleichzeitig zeigt diese Studie auch die Möglichkeiten auf, welche eine statistische Korrektur von Abschlussnoten bietet. Durch den CASE Score kann die Aussagekraft von Noten basierend auf Kontextinformationen gesteigert werden. Angesichts der Ergebnisse raten die Autoren von einer alleinigen Verwendung absoluter Auswahlkriterien auf Basis von Bachelorabschlussnoten ohne Korrekturen oder die Hinzunahme weiterer Kriterien ab.

## Literatur

- Almlund, M., Duckworth, A.L., Heckman, J.J. & Kautz, T.D. (2011). Personality psychology and economics. In Handbook of the Economics of Education (Vol. 4, pp. 1-181).
- Bergerhoff, J., Hartmann, S., Seegers, P.K., Knappe, A. (2018). Fachkraft 2030: Erhebung zur wirtschaftlichen und allgemeinen Lebenssituation der Studierenden in Deutschland. Köln / Maastricht: Studitemps GmbH / Maastricht University.
- Flanagan, J.C. (1954). The Critical Incident Technique. In: Psychological Bulletin, 51 (4), pp. 327-358.
- Goldberg, L.R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), Personality Psychology in Europe, Vol. 7 (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.
- Hell, B., Trapmann, S., Schuler, H. (2008). Synopse der Hohenheimer Metaanalysen zur Prognostizierbarkeit des Studienerfolgs und Implikationen für die Auswahl und Beratungspraxis. In: Schuler, H. & Hell, B. (Hg.): Studierendenauswahl und Studienentscheidung. Göttingen, S. 43-54.
- Kersting, M. & Hornke, L.F. (2003). Qualitätssicherung und Optimierung in der Diagnostik: Die DIN33430 und notwendige Begleit- und Folgeinitiativen. In: Psychologische Rundschau, 54, S.175-178. ; International Test Commission, 2011. ITC Guidelines for Quality Control in Scoring, Test Analysis, and Reporting of Test Scores. [www.intestcom.org](http://www.intestcom.org).
- Schult, J., Hofmann, A. & Stegt, S. J. (angenommen). Leisten fachspezifische Studierfähigkeitstests im deutschsprachigen Raum eine valide Studienerfolgsprognose? Ein metaanalytisches Update. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie.
- Stegt, S.J., Didi, H.-J., Zimmerhofer, A. & Seegers, P.K. (2018). Akzeptanz von Auswahlverfahren zur Studienplatzvergabe. Zeitschrift für Hochschulentwicklung, 13(4), 15-35.
- Stegt, S.J. & Bergholz, L. (2018). Vorhersage des Studienerfolgs in konsekutiven Masterstudiengängen mithilfe eines kognitiven Eignungstests. Hochschulmanagement, 13(4), 101-107.
- Wissenschaftsrat (2012). Prüfungsnoten an Hochschulen im Prüfungsjahr 2010. Arbeitsbericht mit einem Wissenschaftspolitischen Kommentar des Wissenschaftsrates (Drs. 2627-12).



**Tabelle 1: Deskriptive Statistiken**

Variable		Obs	Mean	Std. Dev.	Min	Max
<b>ITB</b>	<b>TM-WISO</b>	4.844	101,05	10,07	70,00	130,00
	<b>PSB</b>	4.844	101,35	9,66	70,00	130,00
	<b>TA</b>	4.844	100,41	10,79	70,00	130,00
	<b>WZF</b>	4.844	100,18	9,72	70,00	130,00
	<b>WGI</b>	4.844	101,34	10,02	70,00	130,00
<b>Vor- benotung</b>	<b>Note Abitur</b>	4.844	2,14	0,56	1,00	4,00
	<b>Note Bachelor</b>	4.844	2,16	0,44	1,00	3,50
<b>CASE</b>	<b>Perzentil lokal</b>	4.844	37,58	21,75	1,00	99,00
	<b>Subject Score</b>	4.844	40,56	25,78	1,00	100,00
	<b>CASE Score</b>	4.844	49,75	23,78	1,00	100,00

**Tabelle 2: Korrelationstabelle** (mit einem \* werden signifikante Korrelationen bei  $\alpha = 0,05$  gekennzeichnet)

	ITB					Vorbenotung		CASE		
	TM-WISO	PSB	TA	WZF	WGI	Note Abitur	Note Bachelor	Perzentil lokal	Subject Score	CASE Score
<b>TM-WISO</b>	1,0000									
<b>PSB</b>	0,7399*	1,0000								
<b>TA</b>	0,7911*	0,4389*	1,0000							
<b>WZF</b>	0,7980*	0,4484*	0,5293*	1,0000						
<b>WGI</b>	0,8228*	0,4516*	0,5778*	0,5489*	1,0000					
<b>Note Abitur</b>	-0,3226*	-0,2186*	-0,2695*	-0,2979*	-0,2376*	1,0000				
<b>Note Bachelor</b>	-0,2662*	-0,1636*	-0,2356*	-0,2248*	-0,2245*	0,3305*	1,0000			
<b>Perzentil lokal</b>	-0,3091*	-0,2055*	-0,2595*	-0,2682*	-0,2444*	0,3309*	0,8717*	1,0000		
<b>Subject Score</b>	-0,3771*	-0,2354*	-0,3168*	-0,3500*	-0,2927*	0,4024*	0,5049*	0,6205*	1,0000	
<b>CASE Score</b>	-0,4190*	-0,2530*	-0,3489*	-0,3976*	-0,3294*	0,4195*	0,5017*	0,6101*	0,9635*	1,0000

**Tabelle 3: Multiple Korrelationsanalysen TM-WISO gegen Abitur- und (korrigierte) Bachelornoten (Standardfehler in Klammern)**

	Gesamt TM-WISO	
	MLR.1	MLR.2
<b>Abiturnote</b>	-4,70 (0,26)	-3,48 (0,27)
<b>Bachelornote</b>	-4,21 (0,33)	---
<b>CASE Score</b>	---	-0,14 (0,01)
<b>R<sup>2</sup></b>	0,13	0,19
<b>N</b>	4.443	4.443

	PSB TM-WISO		TA TM-WISO	
	MLR.1A	MLR.2A	MLR.1B	MLR.2B
<b>Abiturnote</b>	-3,18 (0,27)	-2,44 (0,28)	-4,11 (0,29)	-3,03 (0,30)
<b>Bachelornote</b>	-2,33 (0,33)	---	-4,17 (0,36)	---
<b>CASE Score</b>	---	-0,08 (0,01)	---	-0,12 (0,01)
<b>R<sup>2</sup></b>	0,06	0,08	0,10	0,13
<b>N</b>	4.443	4.443	4.443	4.443

	WZF TM-WISO		WGI TM-WISO	
	MLR.1C	MLR.2C	MLR.1D	MLR.2D
<b>Abiturnote</b>	-4,34 (0,26)	-2,89 (0,26)	-3,27 (0,27)	-2,29 (0,28)
<b>Bachelornote</b>	-3,24 (0,33)	---	-3,74 (0,34)	---
<b>CASE Score</b>	---	-0,13 (0,01)	---	-0,11 (0,01)
<b>R<sup>2</sup></b>	0,11	0,17	0,08	0,11
<b>N</b>	4.443	4.443	4.443	4.443

**Abbildung 1: Streudiagramm und Regressionslinie – CASE Score und TM-WISO Gesamtergebnis**

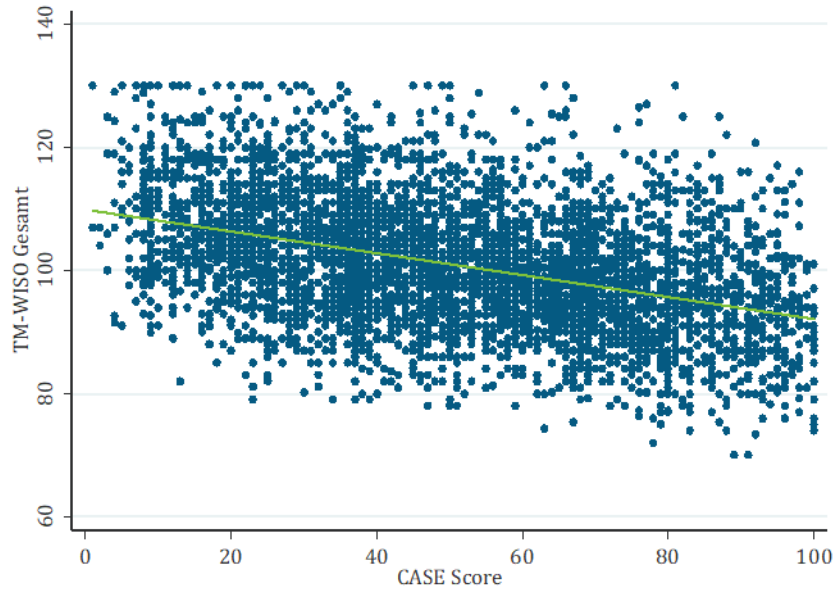


Abbildung 2: Streudiagramm und Regressionslinie – CASE Score und TM-WISO Untertests

