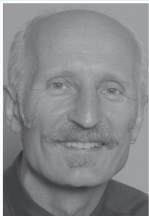


NOTENINFLATION

Leo Schubert



Prof. Dr. Leo Schubert
BWL-Studienschwerpunkte
an der Universität Augsburg:

*Marketing und Unternehmensforschung,
1985: Promotion über Methoden der
Datenanalyse*

bis 1991: CEO-Stabsstelle einer Großbank

*seit 1991: Professor für Marketing an der HTWG
Konstanz*

*Forschungsaufenthalte und Kurzzeitdozenturen
in Mittelamerika.*

*Forschungsschwerpunkte: Marktforschung
(insb. Kapitalmarkt- und Zufriedenheits-
forschung)*

1 EINLEITUNG

Vor über zwei Jahren hat der Wissenschaftsrat¹ auf der ersten Seite der Süddeutschen Zeitung Bedenken bzgl. der Benotung von akademischen Abschlüssen geäußert. Diese haben sich in den letzten 10 Jahren stetig verbessert, sodass in diesem Zeitraum der Anteil der Absolventen, die ihr Studium mit „gut“ oder „sehr gut“ beendeten von 70% auf 80% anstieg. „Der Trend zu besseren Noten darf nicht so weitergehen“ fordert der Vorsitzende des Wissenschaftsrates. Der studentische Ansatz zur Lösung des Problems zu guter Noten wäre trivial. Doch dieser Appell richtet sich nicht an die Studierenden, sondern ist als Rüge an die Prüfenden an den Hochschulen gedacht, die durch „Aufweichung der Bewertungsstandards“ eine „schleichende Noteninflation“ bewirken.

Nicht nur die Ausbildung gehört zu den Aufgaben einer Hochschule, sondern auch die verlässliche Bewertung des Wissens und der Fähigkeiten der Studierenden, um Unternehmen bzw. anderen Hochschulen eine Orientierung zu bieten bei der Auswahl von Personal bzw. Studenten für das Masterstudium. Deshalb ist eine konkretere Betrachtung des Problems angebracht. In diesem Beitrag stehen die Prüfungsergebnisse von zwei Studiengängen im Mittelpunkt. Auch wenn dabei nicht erwartet werden kann, dass der Effekt von den verschiedenen potentiellen Einflussfaktoren durch die Analyse dieser Daten ermittelt wird, ist das explorative Projekt ein erster Schritt, sich mit dem Problem auseinanderzusetzen und ein wenig Transparenz zu bieten.

2 URSACHEN

Die Darstellung der Ursachen der sog. „Noteninflation“ beschränkt sich in den Medien vorwiegend auf Vermutungen und z.T. Schuldzuweisungen. Auch drückt die sog. „Aufweichung der Bewertungsstandards“ aus, dass für die gleiche Leistung²

heute bessere Noten vergeben werden, ohne die Möglichkeit in Erwägung zu ziehen, dass Studierende heute leistungsbe-reiter sind und entsprechend bessere Leistungen vorweisen können. Deshalb sind die im Folgenden aufgeführten Ursachen als potentielle Ursachen zu verstehen:

a) Die Umstellung vom Diplom zu Bachelor und Masterabschlüsse könnte zu einer Verbesserung der Studienabschlüsse führen, da mit dieser Umstellung vermehrt Prüfungsformen wie Referate, Projektarbeit, Studienarbeit etc. in die Studiengänge implementiert wurden, die meist mit einer besseren Bewertung einhergehen, als z.B. die klassische Prüfungsform der Klausur.

Werden bei diesem Vergleich die Noten aus dem Diplomstudium mit denen aus dem Bachelor- und Masterstudium verglichen, ergibt sich durch das zweifache Erfassen der besseren Studenten sowohl mit ihrem Bachelor als auch deren Masterabschluss rein rechnerisch ein Notenverbessernder Effekt, da die Studenten mit schlechteren Noten meist kein Masterstudium durchführen.

b) Eine weitere Auswirkung der Umstellung auf die Bachelor- und Masterstudien wird in den erhöhten Anforderungen und einer verkürzten Studienzeit gesehen, da diese die Abbrecherquote erhöht und dadurch nur noch die Besseren das Studium abschließen³.

c) Studierende haben heute an vielen Hochschulen die Möglichkeit, ihre Dozenten im Rahmen eines KV-Prozesses zu bewerten. Falls diese Bewertung zur Anregung für Dozenten dient, ihre Lehrveranstaltungen zu verbessern, ist ein positiver Effekt durch verbesserte Lehrveranstaltungen auf Noten zu erwarten. Schielen dagegen Dekane, Studienkommission Berufungskommission oder gar die Hochschulleitung auf diese Bewertungen um die Lehrveranstaltungen an die Wünsche der Studierenden anzupassen oder gar in den

W-Besoldungsgruppen die Leistungszulage daran zu orientieren, so muss bedacht werden, dass die Studierenden nicht nur an guten Lehrveranstaltungen interessiert sind, sondern auch an einer entsprechend guten Note⁴. Kontinuierlich verbesserte Noten führen zu zufriedeneren Kunden d.h. Studierenden. Aussagen von Studenten wie „Wer uns sehr streng bewertet, da sehen wir auch genauer hin“ deuten an, dass dieses Instrument wohl nicht valide ist und dass schielende Hochschulorgane die Souveränität, eine angemessene Note für die Leistung eines Studierenden festzulegen, untergraben.

d) Ein weiteres Argument, das als Ursache genannt wird, ist die personelle Überlastung an den Hochschulen. Diese führe zu „sehr oberflächlichen und pauschalen Korrekturen der Abschlussarbeiten“ stellte der Vorsitzende des Philologenverbandes Heinz-Peter Meidinger im Jahr 2007 fest⁵. Es handelt sich dabei um kein singuläres Problem der Hochschulen, sondern betrifft unterschiedliche einst souverän entscheidende Instanzen unserer Gesellschaft, deren Souveränität im Rahmen des „new public managements“ ausgehöhlt wurde⁷.

e) Als weitere Ursache wird die „Bequemlichkeit“ der prüfenden Dozenten vermutet. „Bei einer Eins vor dem Komma hat sich bislang noch kein Student beschwert“ ist als Kommentar einer Studentin in der FAZ zu lesen⁸. Als Versuch diese „Bequemlichkeit“ zu belegen, wird eine besonders gut bewertete Hausarbeit (Note: 1,0) erwähnt, in der keine Seite geknickt und kein Kommentar im Fließtext stand⁹. Bessere Noten bedeuten weniger Diskussion oder Erläuterungen in der Sprechstunde und mehr Zeit für die eigene Forschung, die mehr zur universitären Karriere von Professoren beiträgt als die Lehre.

f) Eine Ursache der Notenentwicklung, zu der die Studierenden beitragen ist das angestiegene Interesse an einem

akademischen Beruf das mehr Leistungsbereitschaft bewirkt. Einerseits stellt dabei die Reputation eines akademischen Berufes ein Motiv andererseits scheint der Arbeitsmarkt dieses Interesse stark beeinflusst zu haben. Ein akademischer Beruf entspricht laut einer Studie des HIS-Institutes für Hochschulforschung quasi einer Jobgarantie, da nur 1 % der Absolventen nach 10 Jahren noch arbeitslos ist. Zudem erzielen 80 % der Akademiker Einkommen, die über dem Durchschnitt der Beschäftigten lag¹⁰. Mittlerweile ist das Verhältnis von Auszubildenden zu Studienanfängern zu Gunsten der Studienanfänger gekippt und der Deutsche Industrie und Handelstag fordert die Akademisierung um jeden Preis zu stoppen¹¹.

g) Die Unterstützung durch die digitalen Medien bei der Erstellung von Referaten und Studienarbeiten ermöglichen, bessere Präsentationen in kürzerer Zeit zu erstellen. Aber auch bei der Vorbereitung auf eine Klausur bieten diese Medien schnelle Informationsmöglichkeiten um Wissenslücken zu schließen.

h) Die Diskussion um die Ursachen der Noteninflation klammert den in den letzten Jahren gestiegenen Anteil an weiblichen Studentinnen und Dozentinnen aus. Da die Art, wie ein Dozent eine Leistung bewertet, dem Datenschutz unterliegt, musste der Einfluss, den das Geschlecht auf die Bewertung ausübt, ausgeklammert werden. Lediglich hinsichtlich des Geschlechts der Studierenden konnte bei den Prüfungen differenziert werden.

i) Auch die zeitweise regional eingeführten und wieder abgeschafften Studiengebühren können bessere Noten bewirken. In den USA und GB, deren Hochschulen wesentlich höhere Studiengebühren von ihren Studierenden verlangen, ist dieses Problem der Bestnotenvermehrung seit Jahren bekannt¹². Wer viel bezahlt, erwartet auch einen ordentlichen Abschluss. In Deutschland scheint weniger die Er-

wartungshaltung aufgrund der Höhe der Studiengebühr ursächlich. Die durch die Studiengebühren bzw. Qualität sichern den Mittel finanzierten Maßnahmen zur Verbesserung der Studiensituation sollten jedoch Spuren hinterlassen haben.

Neben den genannten allgemeinen potentiellen Ursachen sind die lokalen mindestens ebenso wirksam und erschweren eine klare Begründung der beobachteten Noteninflation einer Fakultät. Zu diesen möglichen lokalen Ursachen zählt eine hohe Bewerberzahl als Ausdruck guter Vermarktung der Leistung einer Hochschule und ihrer Standortbedingungen. Diese Bewerberzahl müsste sich insbesondere bei einem effektiven Auswahlprozess auf das Notenniveau auswirken. Zudem kann ein anspruchsvolles Studium entsprechend profilierte Studierende ansprechen, die erwartungsgemäß bessere Leistungen vorweisen.

3 KONSEQUENZEN

Auf die sog. Noteninflation sind bereits Reaktionen beobachtbar:

a) Hochschulen, die Masterstudenten auswählen, müssen Auswahlverfahren installieren, die garantieren, dass die Qualität der Kandidaten einem Masterstudium entspricht. Diese Auswahlmühen können auch verlagert werden, indem bis zu drei Professorengutachten zum davor gelegenen Studium von Studierenden eingefordert werden. Obgleich dies mehr als Hürde denn als Qualitätsgarantie gesehen werden kann, da ja dieselben Professoren, die um die Gutachten gebeten werden, bereits die vermuteten zu guten Bewertungen abgaben.

b) Unternehmen, die Studierende einstellen wollen, werden ebenfalls Auswahlverfahren modifizieren, um sich ein eigenes Bild des Bewerbers zu machen. Dies kann durch einen um Wissensfragen erweiterten Assessment-Prozess gesche-

hen, oder durch die Beschränkung auf Bewerber, die bereits durch Praktika oder Abschlussarbeiten im entsprechenden Unternehmen ihr Können unter Beweis stellen konnten. Manche Personalverantwortliche sprechen bereits von einer relativen Bedeutungslosigkeit der Noten, die nur noch als grobes Filterkriterium dienen.

Das Outsourcing des Auswahlprozesses an spezielle Beratungsunternehmen ginge mit einer Standardisierung der Lehrinhalte einher. Dass dies lukrativ sein kann, zeigt der sog. TOEFL-Test, der die Englischbewertung des Abiturs völlig in den Schatten gestellt hat. Für die unterschiedlichen akademischen Fachrichtungen wäre eine Standardisierung des Basiswissens durch eine externe Organisation aufwendiger und würde den akademischen Lehrbetrieb quasi indirekt einen Lehrplan verpassen.

c) Für gute Studenten bedeutet eine Schwemme guter und sehr guter Abschlüsse, dass andere Felder der Profilierung gefunden werden müssen um noch herausragend zu sein. Manche hoffen deshalb, z.B. durch Preise für akademische und nicht akademische Leistungen oder durch Sprachenvielfalt mehr Aufmerksamkeit auf dem Arbeitsmarkt zu bewirken. Der Masterabschluss hätte eine derartige Differenzierung sein können. Doch statt (wie geplant) 30 % der Bachelorstudierenden in ein Masterstudium aufzunehmen studieren heute etwa 80 % in Masterstudiengängen³. Entsprechend stark hat das Bestreben zu Promovieren zugenommen.

4 DATENBASIS DER AUSWERTUNG

Die Datenbasis für die folgenden Auswertungen sind die Prüfungsdaten der Studiengänge Betriebswirtschaft und Wirtschaftssprachen des Fachbereiches Wirtschafts- und Sozialwissenschaften der letzten 24 Jahre, die von der Hochschulverwaltung dafür zur Verfügung gestellt wurden und insgesamt 81994 Prüfungsfälle im

Zeitraum vom WS 1991 bis zum WS 2013 ergaben (vgl. Tabelle 1). Diese Prüfungen beziehen sich ausschließlich auf mit Noten bewertete Leistungen. In diesen Fällen sind 1961 Abschlussarbeiten enthalten. Die generierte Leistung „Gesamtnote“ die bei Abschluss des Studiums für 1959 Studierende ausgestellt wurde (vgl. Tabelle 2), ist nicht in diesen Fällen enthalten.

Die Prüfungen werden nach Prüfungsformen differenziert betrachtet. Diese sind in folgende Kategorien eingeteilt: Klausur, mündliche Prüfung, Referat, sonstige schriftliche oder praktische Arbeiten und Haus- oder Studienarbeiten. Die oben erwähnten Abschlussarbeiten zählen zur letzten Kategorie. Die Zuordnung zu diesen Kategorien wurde in der Datenbasis in ca. 3% der Fälle nicht vorgenommen, so dass für Auswertung hinsichtlich der Prüfungsform lediglich 79515 Prüfungsfälle zur Verfügung standen⁴.

5. GLEITENDE DURCHSCHNITTE ZUR GESAMTNOTE

Um die Veränderung der Abschlussnoten zu vergleichen wurden die Jahre 2003

und 2013 ausgewählt. Zu diesen Jahren zählt das Sommersemester des Jahres und das folgende Wintersemester. Die Gesamtnote des Studienabschlusses der beiden wirtschaftswissenschaftlichen Studiengänge zeigt eine Verbesserung von 0,16 vom Jahr 2003 bis 2013. Diese Verbesserung ist die Differenz aus den beiden Mittelwerten der Tabelle 3. Grundlage waren dabei die Gesamtnoten der 112 Diplomabschlüsse im Jahr 2003 und der 135 Bachelorabschlüsse im Jahr 2013. Eine lineare Regression über den gesamten Zeitraum (1995-2013), in dem Abschlüsse vorlagen, ergab einen Steigerungsfaktor von 0,015 p.a.. Dies würde für einen Zeitraum von 10 Jahren eine Verbesserung der Noten um 0,15 bedeuteten.

Das 90% Perzentil gibt die Notenschwelle an, die 90% der Absolventen im entsprechenden Jahr mindestens erreichten. Dementsprechend zeigt die Note 2,3 des 90% Perzentils, dass mindestens 90% der Studierenden einen guten bzw. sehr guten Bachelorabschluss im Jahr 2013 erhielten.

Um die Veränderung der Noten im

		Abschluss			Summe
		Diplom	Bachelor	Master	
		Anzahl	Anzahl	Anzahl	
Betriebswirtschaft	Prüfungen (davon Abschlussarbeiten)	32532 (1033)	20078 (247)	1529 (50)	54139 (1330)
Wirtschaftssprachen	Prüfungen (davon Abschlussarbeiten)	10508 (294)	16368 (300)	979 (37)	27855 (631)
Summe (davon Abschlussarbeiten)		43040 (1327)	36446 (547)	2508 (87)	81994 (1961)

TAB. 1: Anzahl aller originärer Prüfungen und Abschlussarbeiten (von WS91/92 bis inkl. WS13/14)

		Abschluss			Summe
		Diplom	Bachelor	Master	
		Anzahl	Anzahl	Anzahl	
Betriebswirtschaft		1031	247	50	1328
Wirtschaftssprachen		294	300	37	631
		1325	547	87	1959

TAB. 2: Anzahl aller abgeschlossenen Studien (von WS95/96 bis inkl. WS13/14)

Jahr		2003 (Dipl.)	2013 (Bac.)
Anzahl	Gültig	112	135
Mittelwert		2,0241	1,8593
Standardabweichung		0,43	0,37
Perzentile	10	1,40	1,30
	20	1,66	1,50
	30	1,80	1,70
	70	2,30	2,10
	80	2,40	2,20
	90	2,60	2,30

TAB. 3: Perzentile und durchschnittliche Gesamtnote im Jahr 2003 und 2013

Zeitverlauf darzustellen wurden gleitende Durchschnitte bestimmt. Dabei werden beim Diplom und Bachelor jeweils arithmetische Mittelwerte zu 100 Werten und beim Master je 30 Werte einbezogen. Jeder Mittelwert sollte mindestens 2 Semesterstärken umfassen, um die Schwankungen zwischen WS und SS auszugleichen. Der erste Mittelwert bezieht sich auf die ersten 100 bzw. 30 Abschlüsse. Der zweite Mittelwert berechnet sich aus den 2. – 101. bzw. 2. – 31. Gesamtnoten etc.. In diesem Sinne gleitet der Mittelwert über die Zeitreihe. Die einzelnen Mittelwerte wurden durch eine Linie verbunden. Für die beiden Studiengänge ergab sich der in Abbildung 1 dargestellte Verlauf. Die Grafiken zu den Abschlussarten Diplom, Bachelor und Master überlagern sich, da in den beiden Studiengängen zu unterschiedlichen Zeitpunkten die Umstellung auf Bachelor bzw. Master vorgenommen wurde. Da sich die Prüfungen nicht linear am Datum orientierten, sondern an einer Reihenfolge der Prüfungen, ist die Zeitachse ebenfalls nicht linear und lediglich mit dem ersten und letzten Abschlussemester angedeutet.

Die gleitenden Durchschnitte zeigen über den Zeitverlauf, dass bereits während des Diplomstudiums die Notendurchschnitte zu fallen begannen und nicht erst mit dem Beginn der Bachelorabschlüsse, wie eine der potentiellen Ursachen annahm. Das Masterstudium wird, wie

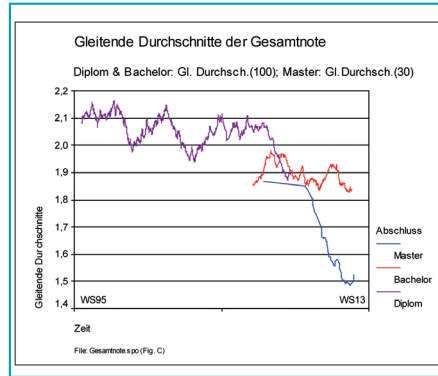


ABB. 1: Gleitende Durchschnitte zu den Noten der einzelnen Abschlussarten.

erwartet besser als die andern beiden Abschlussarten abgeschlossen. Da hier nur eine geringe Anzahl an Abschlüssen vorliegt, kann zu den Masterabschlüssen noch keine Tendenzaussage abgeleitet werden. Die ersten Masterabschlüsse waren auf Niveau der Bachelorabschlüsse. Die besseren Abschlüsse in den letzten Mastersemestern haben den gleitenden Mittelwert deutlich verlagert. Im Bachelor ist eine leichte Abwärtstendenz hin zu besseren Noten erkennbar.

6. DAS NOTENFELD

Um die Durchschnittsnote und deren Streuung bei einzelnen Prüfungsformen zu betrachten, wurde das Feld möglicher Kombinationen aus Durchschnittsnote und Streuung ermittelt. Diese Streuung ist durch die Beschränkung der Notenskala auf die Noten 1,0 bis 5,0 ebenfalls beschränkt. Wie Abbildung 2 zeigt, hängt die Form des rechten Rands des Feldes von der Anzahl der Teilnehmer einer Prüfung ab. Bei zwei Prüfungskandidaten nimmt das Feld die Form eines Dreiecks an. Bereits bei 15 Teilnehmern nähert sich der „blasenförmige“ Rand des Notenfeldes stark dem Grenzverlauf (mit der Anzahl 1000 angedeutet) an. Die Funktion zur Standardabweichung

$$s(\bar{x}) = \sqrt{-\bar{x}^2 + 6 \cdot \bar{x} - 5}, \text{ mit } 1,0 \leq \bar{x} \leq 5,0$$

zeigt den Grenzverlauf in Abhängigkeit

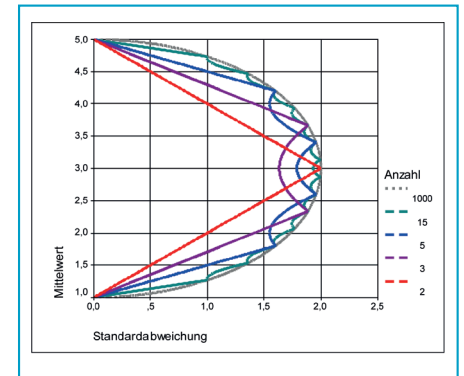


ABB. 2: Maximale Standardabweichungen bei Teilnehmerzahlen von n = 2, 3, 5, 15 und 1000.

von dem jeweiligen Notenmittelwert \bar{x} einer Prüfung¹⁵. Der Grenzverlauf entspricht beim Streumaß Varianz, also s^2 , einer quadratischen Funktion. Um auch die Perzentile im Notenfeld zu kennzeichnen, wurden die sog. TSP-Linien zu 70%, 80% und 90% integriert. TSP steht dabei für Target-Shortfall-Probability¹⁶. Um mindestens eine gute oder sehr gute Abschlussnote zu erzielen, muss ein Student eine „Target-Note“ von mindestens 2,5 erzielen. Für eine Prüfung zeigt die TSP-Linie, wo der Mittelwert und die Standardabweichung der Noten liegen müssen, damit 70%, 80% bzw. 90% das Ziel, mindestens 2,5 zu erzielen, erreicht haben (vgl. Abbildung 3). Hier muss darauf hingewiesen werden, dass durch die Beschränkung der Noten auf das Intervall von 1,0 bis 5,0 die dabei getroffene Annahme normal verteilter Noten oft nicht erfüllt

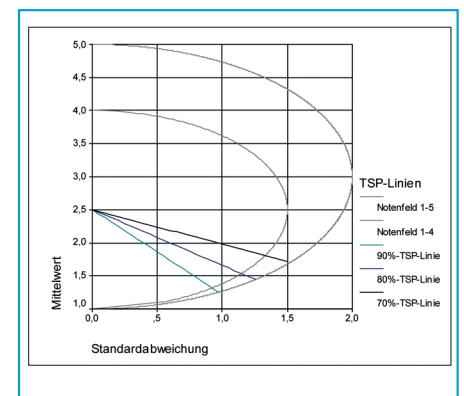


ABB. 3: Target-Shortfall-Probability Linien im Notenfeld.

ist, mit der Konsequenz, dass dann die TSP-Linien, die den resultierenden Verteilungen entsprechen würden, flacher verlaufen. Die %- bzw. Wahrscheinlichkeitsangaben der abgebildeten TSP-Linien sind in diesen Fällen höher anzusetzen¹⁷.

Bezieht man lediglich Notendurchschnitte von bestandenen Prüfungen (Note $\geq 4,0$) ein, wie bei dieser Untersuchung, so reduziert sich das Notenfeld durch die beim Mittelwert von 4,0 startende Kurve (vgl. Abbildung 3, 5a – 6b).

7. DIE PRÜFUNGSFORMEN IM NOTENFELD

Die differenzierte Betrachtung der durchschnittlichen Bewertungen nach einzelnen Prüfungsformen in Abbildung 4a zeigt, dass die Klausur bei der Abschlussart Diplom deutlich über der 70%-TSP-Linie

liegt und die Prüfungsformen des Referats, der mündlichen Prüfung und der Hausarbeit diese bereits unterschritten hatten. Im Vergleich mit der Abschlussart Bachelor stellt die Klausur einen Fixpunkt dar, während sich die anderen Prüfungsformen in den Bereich zwischen der 80%- und 90%-TSP-Linie bewegt bzw. überschritten haben, wie im Falle der sonstigen schriftlichen und praktischen Arbeiten (vgl. Abbildung 4b). Wie Abbildung 4c verdeutlicht, hat auch die Klausurbewertung beim Masterstudium die 80%-TSP-Linie passiert, jedoch auch hier deutlich abgesetzt von den Referaten, mündlichen Prüfungen und sonstigen schriftlichen und praktischen Arbeiten, die die 90%-TSP-Linie hinter sich gelassen haben.

Die einzelnen Diplom-Klausurfächer in Abbildung 5a zeigen eine Punktwolke, die entsprechend der Abbildung 4a um die mittlere Note 2,4 streut. Einige der mittleren Fächernoten fallen deutlich unter die 80%-TSP-Linie und einige werden mit Noten schlechter als 3,0 im Mittel bewertet. Betrachtet man eine horizontale Linie durch die Punktwolke, so zeichnen sich die Klausuren am linken Rand durch in etwa gleiche Bewertungen aus; d.h. die Noten streuen wenig um die mittlere Fachnote. Dies kann bei schlechten Noten auf ein permanent zu hohes Anspruchsniveau hinweisen und bei sehr guten Noten auf ein geringes. Bei Klausuren kann das Argument „Bequemlichkeit“ des Dozenten

kaum unterstellt werden, da i.d.R. eine vorab festgelegte Bepunktung angewendet wird, die kaum Verhandlungsspielräume bietet. Die mittleren Fächernoten am rechten Rand der Punktwolke sind Fächer mit überdurchschnittlich stark streuenden Noten. Dies kann z.B. bei mathematischen Fächern der Fall sein, bei denen gelegentlich fast Zweipunktverteilungen beobachtet werden können. Liegt im Extrem ein Fachmittel nahe am rechten Rand des Notenfeldes, so wurden primär Noten nahe 1,0 oder 4,0 bzw. 5,0 vergeben. Bei der Punktwolke der Abbildung 5a zu den Diplom-Klausurfächern der beiden Studiengänge scheint kein Fach extreme Streuungen aufzuweisen; diese könnten jedoch bei der mittleren Fachnote einzelner Semester auftreten. Ob die wenigen sehr guten durchschnittlichen Benotungen mit dem Ziel entstanden, auch durch die Studenten im Rahmen des KVPs entsprechend bewertet zu werden, ist Spekulation. Es könnte sich ebenso um eine sehr gute Lehrveranstaltung handeln, bei der die Studierenden motiviert mehr Zeit investierten als für andere Fächer. Die schlechten Durchschnittsnoten können mit einem ähnlichen Deutungspaar kommentiert werden: z.B.: schlechte Lehrveranstaltung oder großes Selbstbewusstsein, das das eigene Fach als das wichtigste sieht, das mehr Aufmerksamkeit erfordert. Wird das Notenfeld eines einzelnen Semesters betrachtet, ergeben sich weitere Argumente zu Extrempositionen, da sich z.B. viele sehr gute Studenten – oder eben nicht – in einem Semester befinden können.

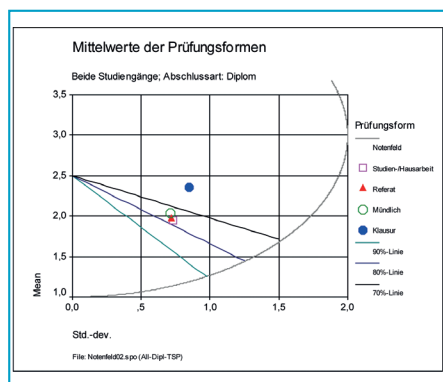


ABB. 4a: Mittelwerte der Noten nach Prüfungsform beim Diplom beider Studiengänge.

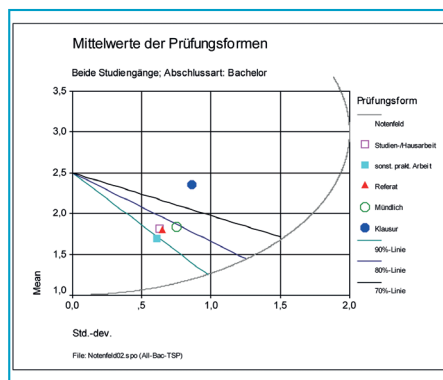


ABB. 4b: Mittelwerte der Noten nach Prüfungsform beim Bachelor.

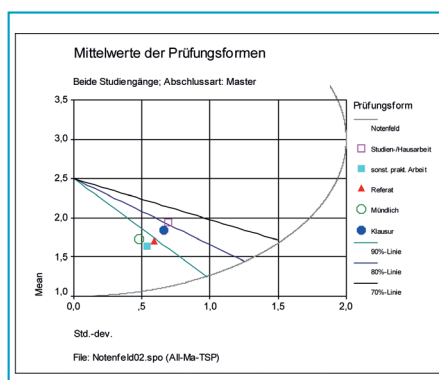


ABB. 4c: Mittelwerte der Noten nach Prüfungsform beim Master.

Im Vergleich zum Diplom besitzen die mittleren Fächernoten der Prüfungsform Klausur des Bachelors, wie Abbildung 6a illustriert, ähnliche Positionen. Einige Fächer haben die 90%-TSP-Linie überschritten und ein Fach zeigt eine relativ gute Durchschnittsbewertung mit starker Reduktion der Streuung. Hier muss angemerkt werden, dass im Bachelor neben den Pflichtfächern sog. Wahlfächer eingeführt wurden, die im Durchschnitt um ca. 0,32 besser

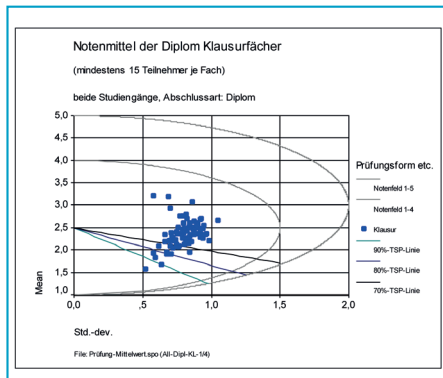


ABB. 5a: Durchschnittsnoten einzelner Diplom-Klausurfächer beider Studiengänge.

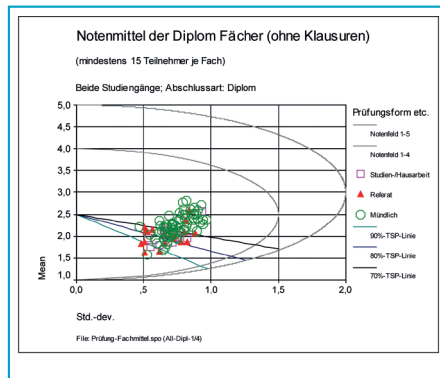


ABB. 5b: Durchschnittsnoten einzelner Diplom-Fächer nach Prüfungsart.

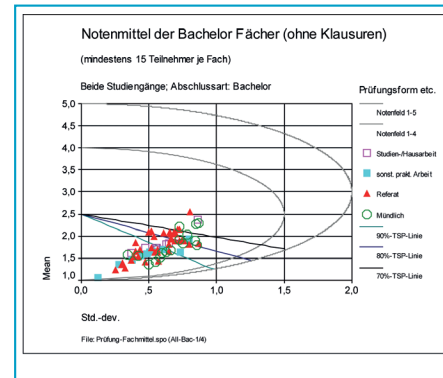


ABB. 6b: Durchschnittsnoten einzelner Bachelor-Fächer nach Prüfungsart.

bewertet werden. Da ein Wahlfach nur stattfindet, wenn sich mindestens 5 Studierende dafür einschreiben, kann das bessere Resultat auch durch das „Buhlen um Teilnehmer“ via Note zustande kommen. Ebenso nahe liegend ist das Argument, dass hier auf Dozenten- wie Studierenden-seite die Lieblingsfächer angeboten bzw. ausgewählt wurden, für die gerne mehr Zeit investiert wird.

Die mittleren Fächerbewertungen der Prüfungsformen Referat, mündliche Prüfung und Haus- oder Studienarbeit zum Diplom zeigt die Abbildung 5b. Wie bereits in Abbildung 4a ersichtlich, deutet die Lage der Punktwolke der Fächer etwas bessere Notendurchschnitte an, wobei die mündlichen Prüfungen eher schlechter als die anderen beiden Prüfungsformen bewertet werden. Auch in dieser Punktwolke sind keine extremen Ausreißer erkennbar. Die Abbildung 6b zeigt analog die Bewertung der Fächer nach Prüfungsformen zum Bachelor. Hier ist eine deutliche Notenverbesserung bei vielen Fächern erkennbar, die mit ihrem Mittelwert unter der 90%-TSP-Linie positioniert sind. In das Bachelorstudium wurde neu die Prüfungsform der sonstigen schriftlichen oder praktischen Arbeiten integriert. Die erkennbar verbesserte Bewertung der abgebildeten Prüfungsformen hat bereits ca. 8 Jahre vor dem Bachelorstudium begonnen. Insbesondere bei Referaten und schriftlichen Ar-

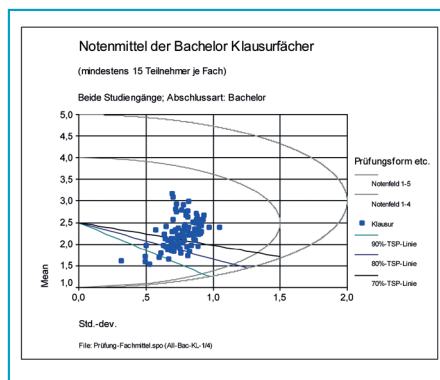


ABB. 6a: Durchschnittsnoten einzelner Bachelor Klausurfächer.

beiten konnten je nach Studiengang eine Verbesserung der Benotung von bis zu 0,6 Notenstufen bei diesen Prüfungsformen festgestellt werden. Hier muss angemerkt werden, dass diese positive Entwicklung auch die Klausuren betraf und in eine Zeit mit extrem hohen Bewerberzahlen fiel (z.T. 10-20 Bewerber je Studienplatz). Referate, Studienarbeiten oder praktische Arbeiten fordern im Gegensatz zu Klausuren nicht von jedem Studierenden die gleiche Aufgabe zu bearbeiten, sondern oft individuelle Aufgabenstellungen. Dadurch ergibt sich ein schwierigerer Vergleich der erbrachten Leistung der Studierenden, der häufig mit „in dubio pro“ entschieden wird. Hinzu kommen die IT-Unterstützung und das Internet, mit deren Hilfe heute schneller attraktiv gestaltete Referate oder Hausarbeiten „gesampelt“ werden können. Bei den sehr gut bewerteten Fächern könnte das

Wahlfachargument oder das der Lieblingsfächer ursächlich sein. Falls ein Einfluss auf die Note durch KVP Bewertungen existiert, so würde dieser hier besonders stark ausfallen, da eine schlechte Bewertung mehr Subjekt bezogen gesehen wird und darüber hinaus nicht nur die allgemeine Bewertungsstrenge eines Dozenten sanktioniert wird, sondern in vielen Fällen die Note selbst, da diese oft direkt nach Erbringung der Leistung mitgeteilt wird.

8. DER KLEINE UNTERSCHIED

Betrachtet man die Gesamtnote des Studienabschlusses differenziert nach Geschlecht, so zeigt sich ein Unterschied von männlichen und weiblichen Studenten. In den Werten der Tabelle 4 sind alle Abschlussarten berücksichtigt. Während die 913 männlichen Studenten im Durchschnitt die Gesamtnote von 2,03 erzielten, schnitten die Studentinnen mit 1,94 insgesamt um 0,09 besser ab. Geht man davon aus, dass die einzelnen Ergebnisse unabhängig zustande kamen, so kann der Unterschied bei Anwendung des approximativen Zweistichproben Gaußtests¹⁸ als hoch signifikant bezeichnet werden. Leider stammen die Noten von einem relativ kleinen Kreis von Dozenten, sodass die signifikante Qualität in Frage gestellt ist. Trotzdem ist der Unterschied – deskriptiv betrachtet – in den beiden Studiengängen bemerkenswert, da sich dieser fast kons-

tant in allen Prüfungsformen wieder findet. Der Vergleich auf Prüfungsebene findet mit einer hohen Anzahl von Prüfungen statt (vgl. Tabelle 5). Trotzdem wird hier auf einen Signifikanztest verzichtet, da die Prüfungsergebnisse einzelner Studierender nicht als unabhängig bezeichnet werden können. Abbildung 7 veranschaulicht die Ergebnisse der Prüfungsformen nach Geschlecht.

Die Entwicklung des Anteils der von Studentinnen abgelegter Prüfungen ist in den letzten 20 Jahren stetig angestiegen. Die Anteilswerte der Tabelle 6 zeigen für 1993 noch 36,6% weibliche und 63,4% männliche Studenten in den Prüfungen. 20 Jahre später ist das Verhältnis umgekehrt mit 62,2% und 37,8%. Rein rechnerisch müsste mit diesem Anstieg des um 0,09 Notenstufen besseren Geschlechts

die Durchschnittsnote um 0,05 Notenstufen steigen.

9. SCHLUSS

Die Untersuchung der Prüfungsdaten belegt eine deutliche Noteninflation in den letzten 20 Jahren. Im Vergleich zu damals würde ein Examen heute mit ca. 0,3 Notenstufen besser bewertet – warum auch immer. Der Beitrag der Klausuren zur Noteninflation stellt sich als gering und der der anderen Prüfungsformen als deutlich größer dar. Das Ziel der Untersuchung beschränkte sich auf die deskriptive Betrachtung der Situation in zwei wirtschaftswissenschaftlichen Studiengängen. In den einzelnen Abschlussarten streuen die Mittelwerte der Noten einzelner Fächer zum Teil stark. Einige fallen sehr gut aus, andere wieder nur befriedigend. Diese Streuungen

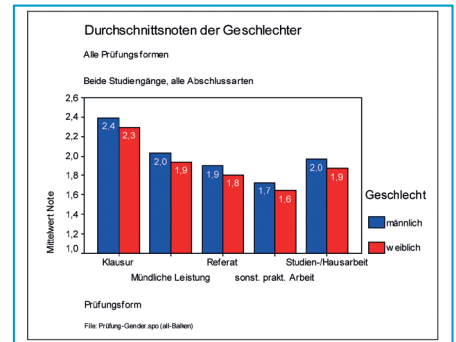


ABB. 7: Vergleich der Gender Notenmittelwerte nach Prüfungsform (Geschlechter Kennzeichnung durch: W-Prüfungsform bzw. M-Prüfungsform).

durch Vorgaben einzuengen würde auch die Möglichkeit, gerecht zu bewerten, reduzieren und darüber hinaus Kollateralschäden bewirken. Die Noteninflation selbst kann auch als Kollateralschaden von Effizienzbemühungen im Hochschulbereich interpretiert werden. Obgleich die Daten keine allgemeinen Schlüsse zulassen, hat eine derartige deskriptive Untersuchung, die nicht oberflächliche Schuldzuweisungen vornimmt, ihre Berechtigung, da sie Transparenz schafft und Reflexion der eigenen Bewertung ermöglicht.

Geschlecht	Anzahl	Noten-Mittelwert	Std. Abw.	Std. Abw. Mittelwert
männlich	913	2,0333	0,4498	0,01489
weiblich	1046	1,9459	0,4002	0,01489
gesamt	1959	1,9866	0,4262	0,00963

TAB. 4: Vergleich der Mittelwerte der Gesamt-Abschlussnoten nach Geschlecht.

Prüfungsform	Geschlecht					
	männlich			weiblich		
	Anzahl	Mittelwert	Std. Abw.	Anzahl	Mittelwert	Std. Abw.
Klausur	23717	2,39	0,86	28084	2,30	0,85
Mündliche Leistung	3894	2,03	0,73	4575	1,94	0,71
Referat	5172	1,90	0,71	6996	1,81	0,65
sonst. schriftl./prakt. Arbeit	816	1,73	0,61	1009	1,64	0,59
Studien-/Hausarbeit	2340	1,97	0,75	2912	1,87	0,67

TAB. 5: Vergleich der Gender Notenmittelwerte nach Prüfungsform

Prüfungsjahr	Geschlecht				
	männlich		weiblich		
	Anzahl	%	Anzahl	%	
1993	599	63,4 %	346	36,6 %	
2003	1554	42,6 %	2094	57,4 %	
2013	2243	37,8 %	3695	62,2 %	

TAB. 6: Anteil der Prüfungen, die in den Jahren 1991 bis 2013 von Studenten bzw. Studentinnen abgelegt wurden unter Einbezug aller Abschlussarten.

FUSSNOTEN

- [1] Preuss R.: Zu gute Noten an deutschen Hochschulen, SZ, 11.11.2012, S. 1.
- [2] Für die Abiturprüfung im Fach Mathematik wurde diese Vermutung von fünf Wissenschaftlern bestätigt (vgl. Mathe-Aufgaben immer leichter, SK, 4.4.14, S. 4). Auch deutete bereits der auf einer Fachtagung am 5. Juli 2012 in Esslingen verfasste „Mindestanforderungskatalog Mathematik der Hochschulen Baden-Württembergs für ein Studium von MINT- oder Wirtschaftsfächern (WiMINT)“ des cosh (Kooperation Schule – Hochschule) darauf hin, dass Hochschulen bei den heutigen Studienanfängern Mathematikkenntnisse nachrüsten müssen.
- [3] Vgl. Plöter K.: Der Leistungsdruck lässt die Examensnoten steigen, DIE ZEIT online, 14.12.2014, <http://www.zeit.de/studium/2012-12/leserartikel-noteninflation->

examen (abgerufen am 1.7.2014))

[4] Bereits 2007 vermutete der damalige Vorsitzende des Deutschen Philologenverbandes Heinz-Peter Meidinger einen gewissen Kuhhandel, bei dem „gute Abschlussnoten gegen gute Veranstaltungsbewertungen seitens der Studierenden“ getauscht werden. Die „Vertrautheit zwischen Professoren und Abschlusskandidaten“ verhindern ein Ausschöpfen der Notenskala (vgl. <http://www.spiegel.de/unispiegel/studium/einsen-fuer-alle-kuschelnoten-kuhhandel-kumpanei-a-460388.html>, 17. Januar 2007 (abgerufen 1.7.2014)).

[5] Vgl. <http://www.spiegel.de/unispiegel/studium/einsen-fuer-alle-kuschelnoten-kuhhandel-kumpanei-a-460388.html>, 17. Januar 2007 (abgerufen 1.7.2014)).

[6] Bartmann Ch., *Leben im Büro*, 2012.

[7] Deutsche Richter wurden im März 2013 vom Bundesverfassungsgericht gerügt, da diese zur Verkürzung von Prozessen häufiger untransparente „Deals“ in Strafrechtsfällen eingehen. Die Aufgabe eines Gerichts, die Wahrheit zu ermitteln, um auf deren Grundlage Recht zu sprechen, konkurriert hier mit der Prozesskostenreduktion (vgl. auch W. Janisch: Viele Richter kungeln am Strafrecht vorbei, *SZ*, 2.11.2012, S. 1 bzw. J. Cáceres: Die Justiz soll schneller werden, *SZ*, 28.3.2013). Auch dem an Gymnasien ausgetesteten Abitur trauen immer weniger (vgl.

U. Fricker: *Muster ohne Wert*, *Südkurier*, 21.3.2013, S. 2) in Anbetracht des von 2006 bis 2012 um 40% gestiegenen Anteils an 1,0 Abiturnoten (vgl. *Südkurier*, 16.6.14, S. 4).

[8] A. Schäfer: *Noteninflation – Einsen für alle*, *FAZ-online*, 25.3.2013 (<http://www.faz.net/aktuell/beruf-chance/campus/noteninflation-einser-fuer-alle-12115313.html>) (abgerufen am 1.7.2014)).

[9] Hier muss angemerkt werden, dass manche Fakultäten die Abschlussarbeit als juristisch relevantes Dokument sehen und deshalb darin keine Kommentare anbringen.

[10] Vgl. Hochschulabschluss fast wie Jobgarantie, *Südkurier* vom 25.7.2013, S. 2. Eine Studie des Nürnberger Institutes für Arbeitsmarkt- und Berufsforschung (IAB) berechnete folgende Lebenseinkommen (in Mio €): ohne Berufsausbildung: 1, mit Lehre: 1,3, mit Abitur: 1,6, FH-Absolventen: 2, Uniabsolventen: 2,3 (vgl. Studie: *Ausbildung zahlt sich aus*, *Südkurier*, 24.1.14, S. 7).

[11] Vgl. *Wirtschaft warnt vor Akademisierung*, *Südkurier* vom 9.9.2013, S. 6.

[12] Das Problem der „Inflation der Bestnoten“ ist in den USA kein neues Phänomen und gehört für David Forster Wallace zur Charakterisierung der Befindlichkeit der amerikanischen Gesellschaft (siehe D.F. Wallace: *Unendlicher Spass*, 2009, S. 145).

[13] vgl. Bergmann L.: *Warum ist schnelles Lernen out?*, *brand eins*, 02/2014, S. 10ff.

[14] Neben den nicht vollständig registrier-

ten Prüfungsformen besteht eine weitere Unschärfe in den Daten, da einige Fächer Wahlmöglichkeiten hinsichtlich der Prüfungsform vorsahen und in den letzten Semestern vermehrt „Modulprüfungen“ als Note erfasst wurden und nicht mehr detailliert die einzelnen Modulteilprüfungen und deren Prüfungsformen erkennen ließen. Diese Unschärfen bewirken bei der Bestimmung der mittleren Noten der einzelnen Prüfungsformen, dass deren Differenzen etwas kleiner ausfallen als diese in Wirklichkeit sind.

[15] Vgl. Schubert L. (2014): *Mean-Variance - Space of Evaluations*, 2014, <http://opus.htwg-konstanz.de/frontdoor/index/index/docId/272>

[16] Die Target-Shortfall-Probability ist ein Risikomaß in der Kapitalmarkttheorie und wurde zur Messung des Anteils der Studierenden mit guten oder sehr guten Abschlüssen übertragen. Ein Beispiel zur Anwendung der TSP in der Portfoliooptimierung zeigt z.B. Schubert L. (2002): *Portfolio Optimization with Target-Shortfall-Probability-Vector*, *Economics Analysis Working Papers*, Vol. 1 No 3, 2002, ISSN15791475.

[17] Vgl. Schubert L. (2014): *Mean-Variance - Space of Evaluations*, <http://opus.htwg-konstanz.de/frontdoor/index/index/docId/272>

[18] Vgl. Bamberg G., Baur F., Krapp, M.: *Statistik*, 2012, S. 179.