

Neue SAS-Prozedur GLMSELECT: Gehaltsanalyse und Studiengebührenmodell

Dr. Jakob Margolis
Analytiker

Frizstr. 9
70734 Fellbach
Jakob_margolis@yahoo.de

Anna Margolis
Studentin European Business
FH Münster / University of
Portsmouth
Frizstr. 9
70734 Fellbach
anna_margolis@yahoo.de

Zusammenfassung

Die neue SAS-Prozedur GLMSELECT (experimentell, V9.1) bietet bestimmte Vorteile gegenüber den klassischen SAS-Prozeduren GLM, REG, LOGISTIC: Man kann sowohl stetige als auch binäre Zielvariablen nutzen und automatisch selektierte Variablen in der Makrovariable `&_GLSIND` ausgeben. Der Vortrag zeigt, wie SAS-Anwender die neue Prozedur GLMSELECT in Kombination mit existierenden SAS-Prozeduren effektiv anwenden und Modelle mit Interaktionen und quadratischen Termen optimieren können. Außerdem werden zwei neue Methoden, LASSO und LAR, eingeführt. Diese werden im Vergleich zu den traditionellen Methoden FORWARD, BACKWARD, STEPWISE und den beiden Variablenselektionskriterien (Schwarz Bayesian criterion - SBC und Mallows statistic - CP) untersucht.

Als Beispiel für die Anwendung dient ein Studiengebührenmodell für den europäischen MBA-Markt (Prototyp-Markt). Aus diesem Modell werden mit GLMSELECT Variablen selektiert, die für die Bestimmung von Studiengebühren relevant sind. Auf der Basis dieser wird anschließend ein kombiniertes Modell gebildet, das als differenziertes Studiengebührenmodell für deutsche staatliche BWL-Fakultäten (neuer BWL-Markt) dienen könnte.

Das Ergebnis der Untersuchung ist ein Vorschlag für die Landes- und Bundesbildungsministerien. Im differenzierten Studiengebührenmodell werden die Studiengebühren in Abhängigkeit von signifikanten Kriterien bestimmt. Dazu gehört das durchschnittliche Absolventengehalt, welches als objektives Arbeitsmarktsignal wirkt, sowie das Image der Fakultät bzw. der Uni und die Studentenbetreuung. Als Höchstgrenze für Top-Fakultäten dient der vom Ministerium vorgegebene Höchstsatz von €500. Das differenzierte Modell bietet eine Kompromisslösung zwischen der subjektiven Meinung radikaler Studenten (keine Studiengebühren!) und der verallgemeinernden Forderung der Universitäten (500 Euro für alle!).

Schlüsselwörter: GLMSELECT, Variablenselektion, LASSO, LAR, signifikante Variablen, Studiengebühren, Studiengebührenmodell, BWL, Steuerungssystem

1 Einleitung

Das Thema „Studiengebühren“ ist für deutsche staatliche Universitäten und Fachhochschulen aktueller denn je. Viele Hochschulen kassieren seit dem Sommersemester 2007

Studiengebühren i.H.v. 500 € pro Semester von allen Studierenden. Die Studenten protestieren und wollen selbstverständlich keine Studiengebühren bezahlen. Das Ziel dieser Studie ist eine Kompromisslösung zu finden.

Die Grundidee der Untersuchung ist einen existierenden Studiengebühren-Markt ins Visier zu nehmen, diesen bereits existierenden Markt zu analysieren, signifikante Variablen für Studiengebühren mit der neuen SAS-Prozedur GLMSELECT zu selektieren und auf Basis der selektierten Variablen ein Studiengebührenmodell für deutsche staatliche BWL-Fakultäten zu entwickeln. BWL-Fakultäten werden gewählt aufgrund von zahlreich vorhandenen internationalen Ratings für betriebswirtschaftliche Studiengänge.

Als Prototyp für den deutschen BWL-Markt wird der europäische MBA-Markt bestimmt. Dazu gibt es mehrere Gründe. Beide Märkte sind sich von der Struktur her ähnlich. Allerdings wird der Preis auf dem MBA-Markt meist durch Nachfrage und nicht durch landesspezifische Vorgaben reguliert. Außerdem existieren auf dem europäischen MBA-Markt langjährige stabile statistische Daten, die zur Grundlage dieser Untersuchung werden. Die Daten stammen aus dem weltweiten MBA-Ranking der Financial Times sowie einem Ranking des europäischen MBA-Markts der Karriere (Handelsblattverlag). Diese werden mit der neuen SAS-Prozedur GLMSELECT analysiert um signifikante Variablen zu selektieren.

Die Kompromisslösung heißt Differenzierung. Es ist wenig fair und weitsichtig, gleiche Studiengebühren an allen Hochschulen zu erheben. Die Studiengebühren sollten viel mehr entsprechend der Studienqualität, der Arbeitsmarktreaktion (Gehalt) und dem Image der Hochschule gestaltet werden. Anspruch auf maximale Studiengebühren (z.Z. €500 pro Semester) sollten nur die besten BWL-Fakultäten Deutschlands haben.

Anhand des differenzierten Studiengebührenmodells, könnten die Landes- und Bundesbildungsministerien die Gebühren in regelmäßigen Abständen neu berechnen, um diese Veränderungen im Bildungsmarkt und der Arbeitsmarktsituation anzupassen. Solche „freie Marktpolitik“ würde BWL-Fakultäten sicherlich zur Verbesserung motivieren.

2 GLMSELECT – Test: Gehaltsanalyse für Baseballspieler

Dieser Teil stellt die Funktion GLMSELECT zunächst anhand von Daten vor, die von SAS vorgegeben sind und eine Gehaltsanalyse für Baseballspieler ermöglichen.

2.1 GLMSELECT-Test

Der erste GLMSELECT-Test der SAS-Daten mit Method=Stepwise und Criterion=SBC sieht folgendermaßen aus:

```

*Daten von SAS: Salary for baseball players *;
ods html; ods graphics on;
proc glmselect data=baseball plots=all;
  class league division;
  model logSalary = nAtBat nHits nHome nRuns nRBI nBB
    yrMajor crAtBat crHits crHome crRuns crRbi
    crBB league division nOuts nAssts nError
  / details=all stats=all;
run; * default heisst Method=Stepwise Criterion=SBC Schwarz Bayes **;
ods graphics off; ods html close;

```

Nach fünf Schritten erhält man folgende Ergebnisse:

Tabelle 1: GLMSELECT mit Method = Stepwise und Schwarz Bayes Criterion

The SAS System									
The GLMSELECT Procedure									
Stepwise Selection Summary									
Step	Effect Entered	Effect Removed	Number Effects In	NumberParms In	Model R-Square	Adjusted R-Square	AIC	AICC	BIC
0	Intercept		1	1	0.0000	0.0000	-60.7762	0.7767	-60.6397
1	crRuns		2	2	0.4187	0.4165	-201.4609	0.2419	-200.7872
2	nHits		3	3	0.5440	0.5405	-263.2959	0.0071	-261.8807
3	yrMajor		4	4	0.5705	0.5655	-277.0208	-0.0448	-275.3333
4		crRuns	3	3	0.5614	0.5581	-273.5517	-0.0319	-271.9095
5	nBB		4	4	0.5818	0.5770	* -284.0690	* -0.0716	* -282.1700

`%put &_GLSIND.;` zeigt uns im SAS-Log den Inhalt der Macrovariable `&_GLSIND` und zwar drei selektierte Variablen: **nHits nBB yrMajor**.

Warum aber wird die beste Variable crRuns (Step1 ergibt crRuns max. $r = 0,647$ mit Zielvariable logSalary) im Step4 entfernt? Die Ursache liegt in der starken Korrelation zwischen crRuns und yrMajor $r = 0,879$. Das heißt bei automatischer Selektion mit GLMSELECT (Method=Stepwise und Criterion = SBC) können stark korrelierende Variablen bzw. Informationen verloren gehen und man erhält zwar eine gute, aber nicht die bestmöglich Modellqualität (F-Value ist 120 und R-Square ist 0,5818).

Wie kann das entwickelte lineare Modell verbessert werden?

2.2 Erste Modellverbesserung

Die erste Möglichkeit zur Modellverbesserung ist Hauptkomponentenanalyse mit proc PRINCOMP und `&_GLSIND`:

```
ods html; graphics on;
proc princomp data=WORK.BASEBALL
  (keep= &_GLSIND logsalary)
run;
ods graphics off; ods html close;
```

Und dann kann man mit zwei Variablen prin1 und prin2 ein besseres Modell entwickeln:

```
ods html; ods graphics on;
proc glmselect data= work.baseball_princomp;
** Modell logsalary von zwei Hauptkomponenten prin1 prin2 ;
  model logsalary = prin1 prin2;
run;
ods graphics off; ods html close;
```

Die Variablen prin1 und prin2 liefern ca. 90% der gesamten Information, F Value entspricht 165.

2.3 Zweite Verbesserungsmöglichkeit

Die zweite Möglichkeit zur Modellverbesserung ist GLMSELECT mit Interaktionen höherer Ordnung @2 oder @3 zu verbinden. Diese bietet aber keine bessere Modellqualität:

```
proc glmselect data=baseball plots=all; ** mit Ordnung @2 **;
  class league division;
  model logSalary = yrMajor|nHits|nBB @2
  /details=all stats=all;
run;
```

Ordnung @2 ergibt F Value von 91,35, welcher nicht besser ist, als bei automatischer Selektion.

```
proc glmselect data=baseball plots=all; ** mit Ordnung @3 **;
  class league division;
  model logSalary = yrMajor|nHits|nBB @3
  / details=all stats=all;
  run;
```

Ordnung @3 ergibt F Value 92,63, auch hier ist die Modellqualität nicht besser als bei automatischer Selektion.

2.4 Dritte Verbesserungsmöglichkeit

Die dritte Möglichkeit ist GLMSELECT mit Selection = LASSO und Criterion = SBC:

```
proc glmselect data=baseball plots=CriterionPanel;
  class league division;
  model logSalary = nAtBat nHits nHome nRuns nRBI nBB yrMajor
    crAtBat crHits crHome crRuns crRbi
    crBB league division nOuts nAssts nError
  / selection = LASSO details=all stats=all; * Criterion=SBC *;
  ** (choose=CP steps=20) details=all stats=all; * zwete choose *;
run;
```

Tabelle 2: GLMSELECT mit LASSO und SCB

The GLMSELECT Procedure									
LASSO Selection Summary									
Step	Effect Entered	Effect Removed	Number Effects In	Model R-Square	Adjusted R-Square	AIC	AICC	BIC	CP
0	Intercept		1	0.0000	0.0000	-60.7762	0.7767	-60.6397	375.9275
1	crRuns		2	0.0774	0.0738	-79.9545	0.7039	-80.7747	328.6492
2	crHits		3	0.2204	0.2144	-122.2615	0.5433	-123.5259	239.5392
3	nHits		4	0.3892	0.3821	-184.4300	0.3072	-185.2317	134.0374
4	nBB		5	0.4275	0.4186	-199.4461	0.2505	-200.3894	111.6638
5	crRbi		6	0.4776	0.4674	-221.5489	0.1669	-222.1516	81.7296
6	yrMajor		7	0.4912	0.4793	-226.5085	0.1485	-227.2147	75.0428
7	nRBI		8	0.5644	0.5524	* -265.3338 *	* 0.0014 *	* -264.1725 *	* 30.4494 *

*** Optimal Value Of Criterion**

Selection stopped at a local minimum of the SBC criterion.

Tabelle 3: Korrelierende Variablen yrMajor und crRuns im Modell

The selected model is the model at the last step (Step 7).

Effects:	Intercept nHits nRBI nBB yrMajor crHits crRuns crRbi
-----------------	--

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Value
Model	7	121.53980	17.36283	51.71
Error	255	85.61393	0.33574	
Corrected Total	262	207.15373		

Parameter Estimates		
Parameter	DF	Estimate
Intercept	1	4.579485
nHits	1	0.005085
nRBI	1	0.000748
nBB	1	0.004768
yrMajor	1	0.038867
crHits	1	0.000335
crRuns	1	0.000050967
crRbi	1	0.000007971

GLMSELECT mit LASSO und SCB ergibt kein besseres Modell (Tab.2). Zwei stark korrelierende Variablen yrMajor und crRuns werden jedoch beibehalten (Tab.3).

The GLMSELECT Procedure

LASSO Selection Summary								
Step	Effect Entered	Effect Removed	Number Effects In	Model R-Square	Adjusted R-Square	AIC	AICC	BIC
0	Intercept		1	0.0000	0.0000	-60.7762	0.7767	-60.8397
1	crRuns		2	0.0774	0.0738	-79.9545	0.7039	-80.7747
2	crHits		3	0.2204	0.2144	-122.2615	0.5433	-123.5259
3	nHits		4	0.3892	0.3821	-184.4300	0.3072	-185.2317
4	nBB		5	0.4275	0.4186	-199.4461	0.2505	-200.3894
5	crRbi		6	0.4776	0.4674	-221.5489	0.1669	-222.1516
6	yrMajor		7	0.4812	0.4793	-226.5085	0.1485	-227.2147
7	nRBI		8	0.5644	0.5524	-265.3338	0.0014	-264.1725
8	division_East		9	0.5683	0.5547	-265.6742	0.0008	-264.4629
9	nOuts		10	0.5790	0.5640	-270.2846	-0.0161	-268.6706
10		crRuns	9	0.5859	0.5729	-276.6717	-0.0411	-274.7215
11		crRbi	8	0.5884	0.5771	-280.2332	-0.0552	-278.1825
12	nError		9	0.5892	0.5763	-278.7805	-0.0491	-276.6858

Tabelle 4: GLMSELECT mit Method Lasso und Selektionskriterium CP

Die Anwendung der Methode Lasso bei Criterium=SBC ermöglicht also stark korrelierende Variablen in der Selektionsliste beizubehalten, während bei Criterium=CP (Mallow) eine von zwei stark korrelierenden Variablen aus der Selektionsliste entfernt wird (Tab.4).

Method LASSO ist in diesem Beispiel nicht so „hart“ wie STEPWISE und übernimmt mehr Variablen ins Modell, inkl. zwei stark korrelierender Variablen (SBC). LASSO bietet nicht die beste Modellqualität, F Value = 43 (Choose=CP) oder 51 (Choose=SBC). Die Selektionskriterien sehen bei LASSO jedoch weniger „monoton“ aus als bei STEPWISE.

Als nächstes wird GLMSELECT mit Selection = LAR und criterion = SBC getestet:

```
proc glmselect data=baseball plots=CriterionPanel;
  class league division;
  model logSalary = nAtBat nHits nHome nRuns nRBI nBB yrMajor
    crAtBat crHits crHome crRuns crRbi
    crBB league division nOuts nAssts nError
  / selection = LAR (choose = SBC ) details=all stats=all; run;
```

Tabelle 5: GLMSELECT mit Method LAR und Selektionskriterium SBC

The GLMSELECT Procedure									
LAR Selection Summary									
Step	Effect Entered	Number Effects In	Model R-Square	Adjusted R-Square	AIC	AICC	BIC	CP	SBC
0	Intercept	1	0.0000	0.0000	-80.7762	0.7767	-80.6397	375.9275	-57.2041
1	crRuns	2	0.0774	0.0738	-79.9545	0.7039	-80.7747	328.6492	-72.8102
2	crHits	3	0.2204	0.2144	-122.2615	0.5433	-123.5259	239.5392	-111.5450
3	nHits	4	0.3892	0.3821	-184.4300	0.3072	-185.2317	134.0374	-170.1414
4	nBB	5	0.4275	0.4186	-199.4461	0.2505	-200.3894	111.6638	-181.5853
5	crRbi	6	0.4776	0.4674	-221.5489	0.1669	-222.1516	81.7296	-200.1160
6	yrMajor	7	0.4912	0.4793	-226.5065	0.1485	-227.2147	75.0428	-201.5014
7	nRBI	8	0.5644	0.5524 *	-265.3338 *	0.0014 *	-264.1725 *	30.4494 *	-236.7565 *

* Optimal Value Of Criterion

Selection stopped at a local minimum of the SBC criterion.

Tabelle 6: GLMSELECT mit Method LAR und Selektionskriterium CP

Parameter Estimates		
Parameter	DF	Estimate
Intercept	1	4.184985
nHits	1	0.006757
nHome	1	0.004416
nRBI	1	-0.000659
nBB	1	0.006065
yrMajor	1	0.064205
crHits	1	0.000364
crRuns	1	-0.000185
crRbi	1	-0.000060753
league_National	1	0.058717
division_East	1	0.127452
nOuts	1	0.000162
nError	1	-0.005938

Proc GLMSELECT mit Methoden = LAR und den Criteria SBC und CP behält stark korrelierende Variablen crRuns und yrMajor jeweils in der Variablenliste bei., d.h. LAR wirkt in diesem Beispiel noch „weicher“ als LASSO.

3 GLMSELECT im Einsatz: Studiengebührenanalyse für den europäischen MBA - Markt

Für die Studiengebührenanalyse wurden zwei unabhängige Datenquellen gewählt und zusammengeführt - die MBA-Weltrangliste der Financial Times (FT) und die europäische MBA-Rangliste der Karriere, Handelsblattverlag.

Die Variablenselektion mit GLMSELECT sieht folgendermaßen aus:

```
proc glmselect data= mba.mba_17bs_sas ;
  model studgeb =
    gehalt forsch nennung profstud ausstud ft_rank;
  title 'GLMSELECT Method=Stepwise '; run;
```

Tabelle 7: GLMSELECT für den MBA-Markt mit Method Stepwise

GLMSELECT Method=Stepwise					
The GLMSELECT Procedure					
Stepwise Selection Summary					
Step	Effect Entered	Effect Removed	Number Effects In	SBC	
0	Intercept		1	304.5995	
1	gehalt		2	297.4957	
2	profstud		3	296.1557	
3	nennung		4	291.4028	
4		gehalt	3	290.0436	*
* Optimal Value Of Criterion					

Als Zielvariable dienen die Studiengebühren (studgeb, in €, Karriere); gehalt steht für Absolventengehalt drei Jahre nach dem MBA-Abschluss (in US\$, FT); forsch gibt den Forschungs-Rang an (FT); nennung bezieht sich auf MBA-Nennungen, welche sich durch den Peer-Rang aus der Bewertung durch andere MBA-Anbieter ergeben (in %, Karriere); profstud ist das Verhältnis zwischen den Anzahlen der Professoren und der Studenten (Karriere); ausstud ist der Anteil ausländischer Studierenden (in %, Karriere); ft_rank ist der Rang in der Weltrangliste (FT).

Variable „gehalt“ fliegt nach Step4 raus. Es ist jedoch aus dem Beispiel oben bekannt, dass das an der starken Korrelation zwischen den selektierten Variablen liegen kann. Tatsächlich existiert zwischen den Variablen „gehalt“ und „nennung“ eine starke Korrelation ($r = 0,73$):

Correlationsmatrix:

		gehalt	nennung
gehalt	FT:Absolventen-Gehalt 3.J.nach MBA-Abschl. in \$	1.0000	0.7321
nennung	Karriere:Nennungen-Image Peer-Ranking	0.7321	1.0000

Eigenvalues of the Correlation Matrix

Prin	Eigenvalue	Difference	Proportion	Cumulative
1	1.73209965	1.46419929	0.8660	0.8660
2	0.26790035		0.1340	1.0000

Das Zusammenfassen der zwei korrelierenden Variablen gehalt und nennung in eine latente Variable $\text{prin1} = A1 \cdot \text{gehalt} + A2 \cdot \text{nennung}$ ergibt 87% der gesamten Information (Varianz). Auf Basis einer signifikanten Originalvariable „profstud“ und der latenten Variable prin1 wird mit proc GLMSELECT ein Kombi-Modell aus Originalvariablen und latenten Variablen entwickelt:

The GLMSELECT Procedure
Parameter Estimates

Selected Model

Parameter	DF	Estimate	Standard Error	t Value
Intercept	1	28279	1760.779010	16.06
Prin1	1	4539.249444	840.918024	5.40
profstud	1	6421.964595	2006.319713	3.20

Zwischenergebnis nach GLMSELECT: Für Studiengebühren spielen das Absolventengehalt als objektiver Arbeitsmarktindikator, die Nennungen (Peer-Rang) als Imageindikator und das Verhältnis zwischen den Anzahlen Lernpersonal / Studenten als Studienqualitätsindikator eine wichtige Rolle.

4 Differenziertes Studiengebührenmodell für deutsche staatliche BWL-Fakultäten

Eine Expertenbewertung der staatlichen BWL-Fakultäten nach den oben selektierten Kriterien kann die Konstruktion eines differenzierten Studiengebührenmodells ermöglichen.

Die Karriere (5/2006) nimmt eine ähnliche Bewertung für 18 staatliche BWL-Fakultäten vor. Dabei nutzt die Karriere folgende Gewichtungen um den Rang zu bestimmen:

Zielvariable: BWL Rang	Gewichtung (in %)
Wertung Studenten	26,6
Wertung Absolventen	26,6
Nennungen Personal	26,6
Betreuung	6,66
Studiendauer	6,66
Internationalität	6,66
Absolventengehalt	0

Die Karriere beachtet nur den „akademischen Wert“ eines Studiums, in der praktischen Anwendung jedoch spielt das Absolventengehalt auf dem MBA- und BWL-Markt eine wesentliche Rolle. Man kann für den BWL-Markt eine ähnliche Variablenengewichtung wie für den MBA-Markt verwenden. Allerdings sind die Märkte nicht identisch, d. h. die gewählte Gewichtung für den BWL-Markt sollte ähnlich aber nicht exakt genauso sein wie für den MBA-Markt:

$$\text{index_SG} = 0.45 \cdot \text{bwl_gehalt} + 0.30 \cdot \text{nenn_pers} + 0.15 \cdot \text{wert_absol} - 0.10 \cdot \text{betreuung} ; \text{ (in standardisierten Variablen)}$$

Aufgrund dieser Ergebnisse, beruhend auf den Variablen aus der GLMSELECT-Analyse schlagen wir folgende Gewichtung der Variable für das Studiengebührenmodell vor:

Zielvariable: SG Modell	Gewichtung (in %)
Wertung Studenten	0
Wertung Absolventen	15
Nennungen Personal	30
Betreuung	10
Studiendauer	0
Internationalität	0
Absolventengehalt	45

Dieses Modell kann von den Ministerien beliebig an die Marktsituation angepasst werden, indem die Gewichtungen der Variablen von Experten verschoben werden. Dies wird in einem existierenden BWL-Markt in Zukunft sicherlich leichter fallen, da dann bereits praktische Erfahrungen und konkrete Daten vorliegen.

In unserem Vorschlag sollten die Studiengebühren an den 18 Top-BWL-Fakultäten zwischen €500 (Höchstgrenze des Bildungsministeriums) und €250 liegen. Studiengebühren an anderen BWL-Fakultäten sollten sich auf €100 bis €250 belaufen. Zur Bestimmung wird die folgende Formel verwendet:

$$\text{studgeb} = 250 + ((500 - 250) / (1.2223677 - (-1.4525896))) * (\text{index_sg} - (-1.4525896));$$

Das Ergebnis, das auf den Daten aus der Karriere (05/2006) und den oben vorgeschlagenen Gewichtungen beruht, ist in Tabelle 8 dargestellt.

Tabelle 8: Studiengebühren für 18 staatliche Top-BWL-Fakultäten (Stand: 2006)

Obs	bwl_name	bwl_gehalt_orig	nenn_pers_orig	wert_absol_orig	betreuung_orig	index_sg	studgeb
1	Uni_Mannheim	43900	61.5	60.3	24.6	1.22237	500
2	TU_Freiberg	46200	0.0	65.5	24.2	0.83382	464
3	Uni_Muenster	43900	37.4	44.4	20.7	0.60566	442
4	Uni_Hohenhei	45700	3.3	48.2	26.8	0.42272	425
5	Uni_Frankfur	45200	6.6	50.9	30.0	0.34693	418
6	FU_Berlin	45900	3.3	41.5	30.7	0.28687	413
7	Uni_Koeln	43500	38.5	38.9	31.8	0.23554	408
8	Uni_Ulm	44100	0.0	48.0	14.1	0.19619	404
9	Uni_Muenchen	42400	28.6	45.1	18.4	0.14183	399
10	Uni_Bayreuth	44000	12.1	45.9	25.7	0.13677	399
11	Uni_Regensbu	43200	4.4	54.0	15.9	0.12317	397
12	Uni_Passau	44100	5.5	43.2	26.6	-0.01381	384
13	Uni_Konstanz	44000	3.3	40.6	27.0	-0.12776	374
14	Uni_Wuerzbur	43000	1.1	35.4	27.6	-0.51019	338
15	Uni_Osnabrue	42300	0.0	51.4	36.8	-0.58573	331
16	Uni_Erlangen	41000	6.6	42.8	30.3	-0.82501	309
17	Uni_Marburg	40200	1.1	43.4	26.5	-1.03678	289
18	Uni_Jena	39600	0.0	27.9	25.5	-1.45259	250

bwl_gehalt_orig steht für das durchschnittliche Absolventengehalt (in €); nenn_pers_orig bezieht sich auf die Nennungen von 1000 Firmen-Personalchefs (in %); betreuung_orig stellt das Verhältnis der Anzahlen Studenten/Lernpersonal dar; wert_absol_orig ist die Wertung der Absolventen (in %); index_sg steht für den Index im Studiengebührenmodell; studgeb sind die Studiengebühren, die sich für die 18 Top-BWL-Fakultäten ergeben.

Der FOCUS (37/2005) bestätigt die oben gewählte Rangreihenfolge für deutsche BWL-Fakultäten.

BWL Mannheim profitiert vom hervorragenden Image (Nennungen Personalchefs: 61,5%), hat aber relativ gesehen nicht das höchste Absolventengehalt und nicht die beste Studentenbetreuung. Absolventen der TU Freiberg (Bergakademie), die sich auf Energie spezialisiert, profitieren von der aktuellen Arbeitsmarktsituation und weisen mit €46.200 pro Jahr das höchste Absolventengehalt sowie die höchste Zufriedenheitsrate auf. Allerdings steht die Hochschule in Bezug auf ihr Image hinter anderen BWL-Fakultäten. BWLer in Ulm können sich über die beste Studentenbetreuung (14 Studenten pro Lehrperson) und ein höheres durchschnittliches Einstiegsgehalt als die Mannheimer freuen. Dennoch sind die Ulmer weniger mit ihrem Studium zufrieden als die Mannheimer. Unter den 18 Top-BWL-Fakultäten erzielt die Uni Jena den letzten Rang aufgrund des relativ niedrigsten durchschnittlichen Absolventengehalts, des schwächeren

Image, der niedrigen Absolventenzufriedenheit, sowie einer nur mittelmäßigen Studentenbetreuung.

5 Studiengebühren-Steuerungssystem für staatliche BWL-Fakultäten

Die Studiengebühren könnten von den Landesministerien mit Hilfe des entwickelten differenzierten Modells in regelmäßigen Abständen berechnet werden. Der Wettbewerb würde die BWL-Fakultäten zur kontinuierlichen Verbesserung motivieren.

6 Ergebnisse:

Die neue SAS-Prozedur GLMSELECT hat mehrere Vorteile und einen Nachteil. Zu den Vorteilen zählen eine universelle Zielvariable, automatische Variablenselektion in Macrovariable &_GLSIND, zwei neue Methoden LASSO und LAR sowie eine hohe Effekanzahl. Nachteilig ist, dass stark korrelierende Variablen mit $|r| > 0,7$ bei Anwendung der Methoden STEPWISE und BACKWARD eliminiert werden. Deswegen sollte man GLMSELECT in Kombination mit den SAS-Prozeduren CORR und PRINCOMP sowie GLM, REG und RSREG nutzen.

Mit Hilfe von GLMSELECT wurde das Studiengebührenmodell für den existierenden Europäischen MBA-Markt entwickelt. Folgende signifikante Variablen wurden selektiert: Absolventengehalt als Arbeitsmarktindikator, Nennungen (Peer-Rang) als Imageindikator und Studentenbetreuung als Studienqualitätsindikator.

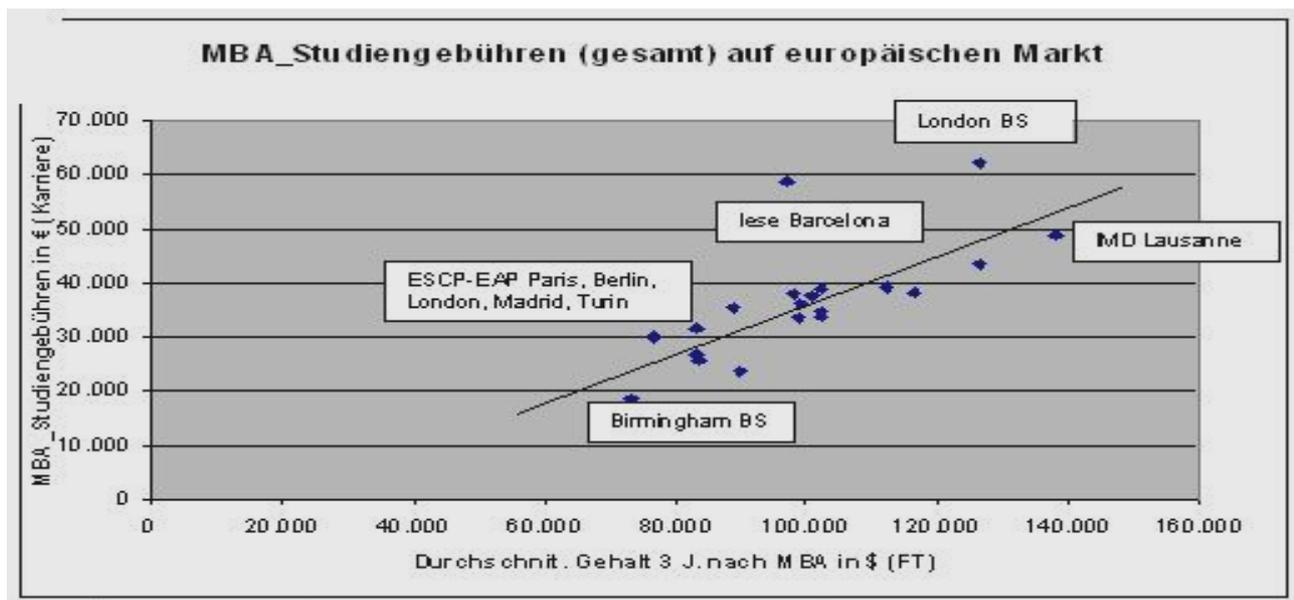


Abbildung 1: Europäischer MBA-Markt - Jede Business School hat ihren Preis

Als wichtigste Variable wurde oben das Absolventengehalt benannt. Die Abbildungen unten zeigen das Verhältnis zwischen dem Absolventengehalt und den Studiengebühren für den Prototyp-Markt an. BWL-Mannheim spielt auf dem deutschen BWL-Markt eine ähnliche Rolle wie die London Business School auf dem europäischen MBA-Markt: Die Absolventen dieser Hochschulen verdienen zwar nicht das maximale Gehalt, jedoch profitieren beide Hochschulen vom hervorragenden Image.

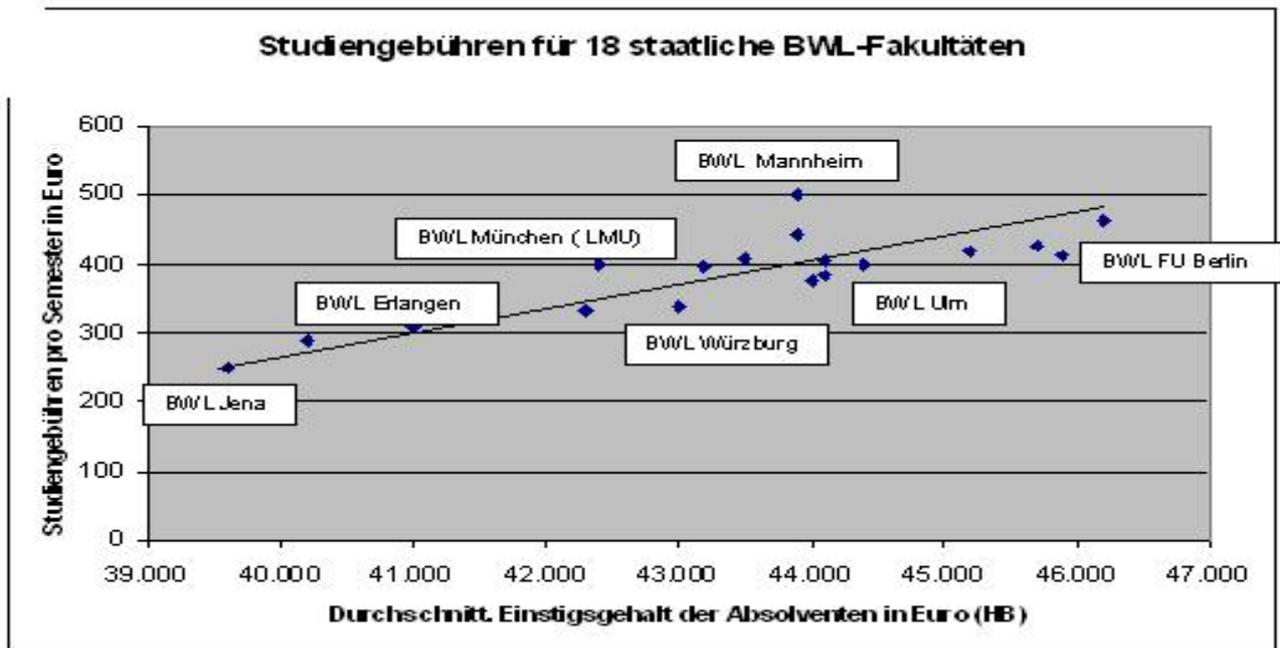


Abbildung 2: Potentieller deutscher BWL-Markt - Jede Fakultät muß ihren Preis haben

Auf der Basis der gefundenen Variablen und ihrer Gewichte wird für den neuen Deutschen BWL-Markt ein differenziertes Studiengebührenmodell vorgeschlagen. Das BWL-Studiengebührenmodell ist eine Kompromisslösung zwischen den Forderungen der Studenten „keine Studiengebühren!“ und den Forderungen der Hochschulen „€500 pro Semester für alle“.

Das Modell ist dynamisch, basiert auf Rückkopplung und reagiert somit auf Marktsignale (vgl. Spence, 1973) sowie die qualitativen Veränderungen innerhalb der BWL-Fakultäten. Landesbildungsministerien könnten das differenzierte Studiengebührenmodell nutzen um Studiengebühren regelmäßig neu zu bestimmen. Der resultierende Wettbewerb würde BWL-Fakultäten zur Verbesserung motivieren.

7 Weiterführende Diskussion

Vorgeschlagen wird ein differenziertes Studiengebührenmodell für staatliche BWL-Fakultäten. Als weiterführende Problemstellung könnte man überlegen, ob sich ein Modell, das auf Variablenselektion beruht, auch für andere Fakultäten anwenden ließe. Welche Methoden könnten alternativ für andere Fakultäten genutzt werden?

Literatur

- [1] Cohen, R. A. (2005): *Introducing the GLMSELECT Procedure for Model Selection*. SAS Inst. Inc., Cary, NC. Paper 207-31.
- [2] The GLMSELECT Procedure (2006, *Experim.*). SAS Inst. Inc., Cary, NC. Paper 1-101.
- [3] MBA-Weltrangliste. *Financial Times Deutschland*, Beilage, 31.01.2005.
- [4] TOP-30-Universitäten Wirtschaft. *Karriere*, 5/2006, S. 47.
- [5] Rangliste deutscher Universitäten (BWL). *FOCUS*, 37/2005, S. 68.
- [6] Spence M.A. (1973): Job Market Signaling. *Quarterly Journal of Economics*, Vol. 87, No. 3, pp. 355-374.
- [7] Strüby R., Reincke U. (2006): *Gegenüberstellung alternativer Methoden zur Variablenselektion*. Vortrag bei der 10.KSFE, Universität Hamburg.
- [8] Margolis J., Margolis A.(2006): *Rank Mining mit SAS für europäische Business Schools und deutsche Elite – Universitäten*. Vortrag bei der 10.KSFE, Universität Hamburg.
- [9] Tibshirani R. (1996): Regression Shrinkage and Selektion via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, Volume 58, Issue 1, S. 267-288.
Efron B., Hastie T., Johnstone I. and Tibshirani R. (2003): *Least Angle Regression*. *Statistical Department, Stanford University*, Januar 9, 2003.