

Diplomprüfung

Fach: Mikroökonomie

Prüfer: PD Dr. Joachim Wolff.

Name, Vorname	
Matrikelnr.	
Studiengang	
Semester	
Datum	14.07.2008
Raum	4.109
Unterschrift	

Vorbemerkungen:

Beachten Sie die **Anweisungen auf der Klausurangabe.**

Aufgabe	Punkte
Aufgabe 1	
Aufgabe 2	
Aufgabe 3	
Aufgabe 4	
Aufgabe 5	
Aufgabe 6	
Aufgabe 7	
Summe	
Note	

Klausur in Mikroökonomie am 14.07.2008**Hinweise:**

- (1) Die Klausur besteht aus 7 Aufgaben, die alle bearbeitet werden müssen.
- (2) Die Punktverteilung der Aufgaben entspricht dem empfohlenen zeitlichen Gewicht bei der Beantwortung, es stehen 120 Minuten zur Verfügung.
- (3) Erlaubte Hilfsmittel sind Taschenrechner und die anhängenden Verteilungstabellen.

Bitte tragen Sie Ihren Namen und Ihre Matrikelnummer ein.

Name

Matrikelnummer

Zahl der Kreditpunkte, die Sie erwerben möchten: 4 CP 6 CP

Aufgabe 1**[23 Punkte]**

Eine Forscherin untersucht, wie lange es dauert, bis ein Beschäftigungsverhältnis endet. Dafür nutzt sie Daten aus einer Umfrage, bei der die Befragten Angaben zum Startzeitpunkt eines Jobs und den Wochen, die der Job andauerte bzw. noch andauert, machen. Diese Angaben werden dabei für alle Jobs erfragt, die die Befragten seit ihrem 16. Geburtstag jemals ausgeübt haben. Ebenso werden viele weitere Merkmale erhoben, wie das Geschlecht, Geburtsdatum, Familienstand zum Beginn eines Jobs, Stundenlohn zum Beginn eines Jobs. Die Befragten sind im Alter von 18 bis 64 Jahren, die Befragung findet nur einmal statt. Ihre Stichprobe besteht also aus allen Jobsells dieser Personen.

- a) Warum wäre ein lineares Regressionsmodell eher ungeeignet, um mit diesen Daten die Effekte der verfügbaren Merkmale auf die Zufallsvariable Dauer eines Jobs zu schätzen? (6 Punkte)
- b) Definieren und erläutern Sie das Konzept der Hazardrate in Verweildauermodellen in stetiger Zeit. (5 Punkte)
- c) Die Forscherin schätzt ein Exponential-Modell, das zu den proportionalen Hazardratenmodellen gehört. Tabelle 1 zeigt ihre Schätzergebnisse (alle geschätzten Koeffizienten und t-Werte). Interpretieren Sie die Koeffizienten statistisch und inhaltlich. (6 Punkte)
- d) Wie könnten die Effekte der Kovariablen auf die Hazardrate in Tabelle 1 auch quantitativ ohne großen Aufwand anschaulich dargestellt werden? Erläutern Sie dies und bestimmen Sie auf diese Weise die Effekte des Stundenlohns und des Alters auf die Hazardrate. (6 Punkte)

Tabelle 1: Effekte der Kovariablen auf die Hazardrate (beliebiger Zielzustand)

Kovariablen	Koeffizient	t-Wert
Stundenlohn	0,03	2,54
Alter/10	-0,114	-2,68
verheiratet	-0,431	-1,65
Frau	0,801	6,21
Konstante	-4,334	-10,01
Anzahl der Spells		4329
LogLikelihood		-245,114
McFadden's Pseudo R ²		0,126

Alle Kovariablen werden zum Startzeitpunkt eines Jobs gemessen: Stundenlohn in Euro, Alter in Jahren, Dummy verheiratet=1 für Verheiratete und 0 für Unverheiratete, Dummy Frau = 1 für Frauen und 0 für Männer.

Aufgabe 2

[22 Punkte]

Tabelle 2 zeigt die Schätzergebnisse eines Multinomialen Logit-Modells zur Wahl zwischen den Berufen Arbeiter, Angestellter und Selbständiger. Die Schätzung beruht auf Querschnittsdaten berufstätiger Personen.

Tabelle 2: Schätzergebnisse eines Multinomialen Logit-Modells zur Berufswahl

Kovariablen	Abhängige Variable: Wahl des Berufs (Referenzkategorie: Arbeiter)			
	Angestellter		Selbständiger	
	Koeffizient	Standardfehler	Koeffizient	Standardfehler
Ausbildungsjahre	0,120	0,074	-0,050	0,009
Berufserfahrung	0,014	0,008	0,047	0,019
Frau	-0,110	0,069	-0,230	0,186
Vater selbständig	-0,040	0,029	0,150	0,039
Konstante	-1,100	0,269	-2,300	0,495
Anzahl der Beobachtungen	2301			
McFadden's Pseudo R ²	0,087			

Zusätzliche Angaben zu den Kovariablen:

- Berufserfahrung in Jahren,
 - Dummy Frau = 1 für Frauen und 0 für Männer,
 - Dummy Vater selbständig = 1, wenn der Vater der befragten Person jemals eine selbständige Tätigkeit ausgeübt hat, und 0 sonst.
- a) Interpretieren Sie die Ergebnisse für die einzelnen Einflussfaktoren der Berufswahl in Tabelle 2 und erläutern Sie, warum es schwierig ist, diese Ergebnisse zu interpretieren. (6 Punkte)
 - b) Erläutern Sie die „Independence of Irrelevant Alternatives“ Annahme anhand eines Beispiels und gehen Sie auf ihre Bedeutung für Parameterschätzungen, die durch die Maximierung der Likelihoodfunktion eines Multinomialen Logit Modells errechnet wurden, ein. (8 Punkte)
 - c) Erläutern Sie den Hausman Test für die „Independence of Irrelevant Alternatives“ Annahme beim Multinomialen Logit Modell (Formel und Intuition). (8 Punkte)

Aufgabe 3

[20 Punkte]

Ein Forscher interessiert sich für die Auswirkungen des Klimas auf den Mineralwasserkonsum von privaten Haushalten. Er hat jährliche Informationen über 10000 repräsentativ ausgewählte Haushalte in Deutschland für eine Periode von 10 Jahren mit folgenden Variablen:

Mineralwasserkonsum _{it}	Liter/Jahr
Erwachsene _{it}	Anzahl der Erwachsenen im Haushalt
Kinder _{it}	Anzahl der Kinder unter 18 Jahren im Haushalt
Einkommen _{it}	Jahreseinkommen des Haushalts (in 100€)
Tagestemperatur _{it}	Durchschnittstagestemperatur im Jahr t gemessen am Wohnort des Haushalts in Grad Celsius
N30Grad _{it}	Anzahl der Tage im Jahr mit einer Höchsttemperatur von mehr als 30 Grad Celsius am Wohnort des Haushalts
N20bis30Grad _{it}	Anzahl der Tage im Jahr mit einer Höchsttemperatur von 20 bis 30 Grad Celsius am Wohnort des Haushalts

Tabelle 3: Regressionsergebnisse für den Mineralwasserverbrauch privater Haushalte

Regressoren	Abhängige Variable: jährlicher Mineralwasserkonsum privater Haushalte (in Litern)			
	Random Effects		Fixed Effects	
	Koeffizient	Standardfehler	Koeffizient	Standardfehler
Erwachsene	296,53	6,86	281,38	23,50
Kinder	164,89	3,74	170,53	9,58
Einkommen	1,49	0,02	1,55	0,10
Tagestemperatur	43,73	1,98	22,01	2,80
N30Grad	1,10	0,37	1,10	0,39
N20bis30Grad	0,31	0,73	0,22	0,76
Konstante	-320,24	75,75	79,88	88,20

Hausman Test (χ^2 -Statistik) = 120,33 (mit p-Wert: 0,0000)

- a) Nennen und erläutern Sie die Vorteile von Paneldaten (insbesondere im Vergleich zu Querschnittsdaten und Daten aus wiederholten Querschnittsbefragungen). (6 Punkte)
- b) Random-Effects-Schätzer und Fixed-Effects-Schätzer sind zwei Typen von Panelregressionsschätzern für stetige abhängige Variable.
- i) Erläutern Sie, wieso bei einer within Schätzung keine Koeffizienten für zeitkonstante Größen ausgewiesen werden. (3 Punkte)
- ii) Wieso können sich die Ergebnisse der within und random effects Schätzung erheblich unterscheiden? (3 Punkte)
- c) Interpretieren Sie die Steigungskoeffizienten der Variablen in N30Grad und N20bis30Grad in Tabelle 3 für das Fixed-Effects-Modell. Der Forscher hat alle Regressoren in linearer Form in das Modell eingehen lassen. Nehmen Sie an, Sie müssen den Forscher bezüglich einer besser geeigneten Spezifikation für sein Vorhaben beraten und es stehen Ihnen keine weiteren Variablen zur Verfügung. Geben Sie eine Spezifikation an, die besser geeignet wäre und begründen Sie, warum dies der Fall ist. (8 Punkte)

Aufgabe 4**[11 Punkte]**

Gestutzte abhängige Variablen

- a) Was versteht man unter einer von unten gestutzten abhängigen Variablen? Geben Sie auch ein Beispiel für eine solche Variable an. (5 Punkte)
- b) Zeigen und erläutern Sie mit Hilfe einer geeigneten Abbildung, warum im Falle von gestutzten abhängigen Variablen eine Kleinstquadratanalyse ungeeignet ist, um den Parameter eines Regressors zu schätzen. (6 Punkte)

Aufgabe 5**[13 Punkte]**

Tabelle 4 zeigt Resultate einer Schätzung eines geordneten Logit-Modells. Es wurden die Determinanten der Lebenszufriedenheit untersucht, wobei die abhängige Variable drei mögliche Ausprägungen hat:

1: völlig unzufrieden

2: weder völlig unzufrieden noch völlig zufrieden

3: völlig zufrieden

Tabelle 4: Schätzergebnisse eines geordneten Logit Modells zu den Determinanten der Lebenszufriedenheit in fünf europäischen Ländern

Kovariablen	Abhängige Variable: Lebenszufriedenheit	
	Koeffizient	Standardfehler
Arbeitslos	-0,356	0,132
Weder arbeitslos noch beschäftigt	-0,04	0,0065
Beschäftigt	Referenzkategorie	
Alter	-0,015	0,005
Alter ²	0,00021	0,000031
pro-Kopf-Einkommen	0,025	0,01
Deutschland	0,33	0,02
Italien	-0,08	0,07
Spanien	-0,01	0,03
Großbritannien	0,042	0,014
Frankreich	Referenzkategorie	
1. Schwellenwert	0,31	0,0294
2. Schwellenwert	0,582	0,0137
McFadden's Pseudo R2	0,232	
Anzahl der Beobachtungen	27453	

Angaben zu den Kovariablen: Die Variablen Arbeitslos, Weder arbeitslos noch beschäftigt, Beschäftigt nehmen den Wert 1 an, wenn der jeweilige Arbeitsmarktstatus zutrifft und 0 sonst. Alter wird in Jahren gemessen. „pro-Kopf-Einkommen“ ist das jährliche Einkommen pro Haushaltsmitglied des Haushalts einer Person in 100 Euro. Die Ländervariablen sind Dummy-Variablen für das Heimatland der Befragten.

- a) Wie sind die Wahrscheinlichkeiten eines geordneten Logit-Modells für den Fall einer abhängigen Variablen mit 3 möglichen Kategorien definiert? (3 Punkte)
- b) Inwieweit ist es möglich mit Hilfe der Koeffizienten in Tabelle 4 die Effekte der Kovariablen auf die Wahrscheinlichkeit „weder völlig unzufrieden noch völlig zufrieden“ und die Wahrscheinlichkeit „völlig zufrieden“ zu sein, direkt zu interpretieren? Erläutern Sie. (6 Punkte)
- c) Wenn eine zuvor beschäftigte Person arbeitslos wird und dabei ihr pro-Kopf-Einkommen um 5000 Euro sinkt, wie würde sich dies gemäß Tabelle 4 auf die Zufriedenheit der Person auswirken? Wie könnte eine staatliche Transferpolitik dafür sorgen, dass ihre Zufriedenheit unverändert bleibt (4 Punkte)

Aufgabe 6

[16 Punkte]

Wahr oder falsch? Tragen Sie für zutreffende Aussagen den Buchstaben w (für wahr) und für nicht zutreffende Aussagen den Buchstaben f (für falsch) ein. Für jede richtige Antwort gibt es einen Punkt, für jede falsche Antwort wird ein Punkt abgezogen. Nicht beantwortete Fragen werden nicht berücksichtigt. Die Gesamtpunktzahl kann nicht negativ werden.

w/f	
	Erwartungstreue ist eine für große Stichproben definierte Eigenschaft eines Schätzers.
	Eine kontrafaktische Situation stellt für Personen mit treatment ab auf den Wert der abhängigen Variable ohne treatment.
	Der Matching-Schätzer basiert auf der Grundidee, jeder Person mit treatment mindestens eine Vergleichsperson ohne treatment zuzuordnen.
	Die Informationsmatrix beschreibt die Krümmung der Likelihoodfunktion.
	Werden Parameter von Determinanten einer gestutzten abhängigen Variablen mit einem Kleinstquadrat-Modell geschätzt, so tritt Heteroskedastie auf.
	Der Kaplan-Meier Schätzer ist ein parametrisches Verfahren.

	Bei zensierten Daten entspricht der Erwartungswert der Bevölkerung dem Produkt aus der Wahrscheinlichkeit, unzensiert zu sein und, solange $A=0$, dem Erwartungswert der gestutzten Verteilung.
	Die Ergebnisse des 2-stufigen Heckman-Verfahrens sind umso verlässlicher, je mehr Ausschlussrestriktionen gegeben sind.
	Um einen Likelihood-Ratio durchführen zu können, ist eine Schätzung unter Restriktionen ausreichend.
	Der Erwartungswert einer Zufallsvariable x sinkt, wenn x von unten gestutzt wird.
	Wird ein Matching-Verfahren mit Zurücklegen durchgeführt, so steigt die Varianz des Schätzers.
	Bei Linkszensierung kann der Anfang einer Episode nicht beobachtet werden.
	Der Butler-Moffit Ansatz implementiert einen fixed effects Logit-Schätzer.
	Das ‚grid-search‘ Verfahren führt auch ohne Kenntnis des Werterahmens der wahren Parameter immer zum globalen Optimum der Likelihoodfunktion.
	Das Maximum-Likelihood Verfahren sucht den Parameter-Vektor, der die Wahrscheinlichkeit maximiert, das zu beobachten, was man beobachtet.
	Erklärende Variablen werden als endogen bezeichnet, wenn sie untereinander korreliert sind.

Aufgabe 7

[15 Punkte]

Sind folgende Aussagen richtig? Geben Sie in der ersten Spalte zunächst w (für wahr) oder f (für falsch) an. Erläutern Sie in der zweiten Spalte stichwortartig Ihre Auffassung (Bsp.: „Stimmt, weil...“ oder „Stimmt nicht, weil...“). Nur bei korrekter Begründung wird die Antwort mit 1,5 Punkten je Frage honoriert.

w/f	Begründung
	Bei der Schätzung der Parameter von Verweildauer-Modellen mit einer Stichprobe aus zwei Gruppen, die sich bezüglich un beobachtbarer Determinanten der Verweildauer unterscheiden, kann negative duration dependence das Schätzergebnis sein, obwohl keine der beiden Gruppen diese aufweist.
	Im linearen Wahrscheinlichkeitsmodell kann ε_i für gegebenes x_i nur zwei Werte annehmen.
	Im linearen fixed effects Modell werden nur die Beobachtungen genutzt, die Variation über die Zeit aufweisen.
	Propensity-Score Matching ist exaktem Matching überlegen, wenn x aus einer großen Anzahl von Regressoren besteht.

	Das Multinomiale Logit-Modell kann für die Untersuchung von Verweildauern verwendet werden.
	Natürliche Experimente können nicht zur Untersuchung kausaler Effekte herangezogen werden.
	Im Logit-Modell hängt die Stärke des marginalen Effekts nicht von den Werten der erklärenden Variablen ab.
	Das gestufte Logit-Modell bietet eine Möglichkeit, Korrelationsstrukturen in den Ausprägungsalternativen kategorialer abhängiger Variablen aufzufangen.
	Liegt ein Messfehler in der abhängigen Variable vor, so sind die Schätzergebnisse des linearen Modells erwartungstreu, aber nicht effizient.
	Im Conditional Logit Modell ist die unterstellte Nutzenfunktion abhängig von den Eigenschaften des Individuums, wie beispielsweise dem Alter.

Tabelle 1: Perzentile der t-Verteilung

Zelleneintrag: x , so dass $\text{Prob}[t_n \leq x] = P$, mit n Freiheitsgraden

n \ P	0.75	0.9	0.95	0.975	0.99	0.995
1	1.000	3.078	6.314	12.706	31.821	63.657
2	0.817	1.886	2.920	4.303	6.965	9.925
3	0.765	1.638	2.353	3.182	4.541	5.841
4	0.741	1.533	2.132	2.776	3.747	4.604
5	0.727	1.476	2.015	2.571	3.365	4.032
6	0.718	1.440	1.943	2.447	3.143	3.707
7	0.711	1.415	1.895	2.365	2.998	3.500
8	0.706	1.397	1.860	2.306	2.896	3.355
9	0.703	1.383	1.833	2.262	2.821	3.250
10	0.700	1.372	1.813	2.228	2.764	3.169
15	0.691	1.341	1.753	2.131	2.602	2.947
20	0.687	1.325	1.725	2.086	2.528	2.845
25	0.684	1.316	1.708	2.060	2.485	2.787
30	0.683	1.310	1.697	2.042	2.457	2.750
35	0.682	1.306	1.690	2.030	2.438	2.724
40	0.681	1.303	1.684	2.021	2.423	2.705
45	0.680	1.301	1.679	2.014	2.412	2.690
50	0.679	1.299	1.676	2.009	2.403	2.678
60	0.679	1.296	1.671	2.000	2.390	2.660
70	0.678	1.294	1.667	1.994	2.381	2.648
80	0.678	1.292	1.664	1.990	2.374	2.639
90	0.677	1.291	1.662	1.987	2.368	2.632
100	0.677	1.290	1.660	1.984	2.364	2.626
∞	0.674	1.282	1.645	1.960	2.326	2.576

Quelle: In R generiert

Tabelle 2: Perzentile der χ^2 -Verteilung

Zelleneintrag: c , sodass $\text{Prob}[\chi_n^2 \leq c] = P$, mit n Freiheitsgraden

n \ P	0.005	0.01	0.025	0.05	0.1	0.25	0.5	0.75	0.9	0.95	0.975	0.99	0.995
1	0.00004	0.0002	0.001	0.004	0.016	0.102	0.455	1.323	2.706	3.842	5.024	6.635	7.879
2	0.01	0.02	0.05	0.10	0.21	0.58	1.39	2.77	4.61	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	0.58	1.21	2.37	4.11	6.25	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	1.06	1.92	3.36	5.39	7.78	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	1.61	2.67	4.35	6.63	9.24	11.07	12.83	15.09	16.75
6	0.68	0.87	1.24	1.64	2.20	3.45	5.35	7.84	10.64	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.22	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.39	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.55	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	7.58	10.34	13.70	17.28	19.68	21.92	24.73	26.76
12	3.07	3.57	4.40	5.23	6.30	8.44	11.34	14.85	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	9.30	12.34	15.98	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	10.17	13.34	17.12	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	11.04	14.34	18.25	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	11.91	15.34	19.37	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	12.79	16.34	20.49	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	13.68	17.34	21.60	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	14.56	18.34	22.72	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	15.45	19.34	23.83	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	16.34	20.34	24.93	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	17.24	21.34	26.04	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	18.14	22.34	27.14	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	19.04	23.34	28.24	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	19.94	24.34	29.34	34.38	37.65	40.65	44.31	46.93
30	13.79	14.95	16.79	18.49	20.60	24.48	29.34	34.80	40.26	43.77	46.98	50.89	53.67
35	17.19	18.51	20.57	22.47	24.80	29.05	34.34	40.22	46.06	49.80	53.20	57.34	60.27
40	20.71	22.16	24.43	26.51	29.05	33.66	39.34	45.62	51.81	55.76	59.34	63.69	66.77
45	24.31	25.90	28.37	30.61	33.35	38.29	44.34	50.98	57.51	61.66	65.41	69.96	73.17
50	27.99	29.71	32.36	34.76	37.69	42.94	49.33	56.33	63.17	67.50	71.42	76.15	79.49

Quelle: In R generiert