# Humboldt-Universität zu Berlin

Wirtschaftswissenschaftliche Fakultät
Institut für Statistik und Ökonometrie
Lehrstuhl Statistik


Bachelor Thesis „Generalized Linear Models“
Wintersemester 2004/2005

---

# Beeinflusst die Perzeption der Wirtschaftslage die Lebenszufriedenheit?

## Eine Anwendung des Logit-Modells mittels ALLBUS-Daten


| | |
|---|---|
| Eingereicht von: | Anna-Julika Frei |
| Eingereicht bei: | Prof. Dr. Bernd Rönz |
| MatrikelNr: | 3614247 |
| Studiengang: | Statistik |
| email: | annajulika@gmx.de |
| letzte Version am: | 8. November 2004 |

**Selbstständigkeitserklärung**

Ich versichere: Ich habe die vorliegende Arbeit selbständig verfasst. Andere als die angegebenen Hilfmittel und Quellen habe ich nicht benutzt.
Die Arbeit hat keiner anderen Prüfungsbehörde vorgelegen.
Mir ist bekannt: Bei Verwendung von Inhalten aus dem Internet habe ich diese zu kennzeichnen und einen Ausdruck davon mit Datum sowie der Internet-Adresse (URL) als Anhang der vorliegenden Arbeit beizufügen.

Berlin, den 8.11.2004 Unterschrift

**Zusammenfassung**

Die Untersuchung der Lebenszufriedenheit hat schon seit langem einen festen Platz in den Sozialwissenschaften und in der Psychologie.
Seit kurzem beschäftigt sich auch die Ökonomie mehr und mehr mit diesem Thema und versucht, das Konstrukt Lebenszufriedenheit mittels ökonomischer Indikatoren besser zu verstehen.
Diese Arbeit beschäftigt sich mit dem Zusammenhang verschiedener sozio-ökonomischer Indikatoren mit besonderem Augenmerk auf die Einschätzung der Wirtschaftslage auf die Lebenszufriedenheit. Innerhalb dieser Untersuchung kann nachgewiesen werden, dass die Einschätzung der Wirtschaftslage einen großen Einfluss auf die Lebenszufriedenheit hat; des weiteren werden Zusammenhänge zwischen anderen in die Untersuchung aufgenommenen Variablen wie zum Beispiel dem Alter und der Lebenszufriedenheit gefunden.
Eine weitere Fragestellung, mit der sich die vorliegende Untersuchung beschäftigt, ist, ob Daten, die aus unterschiedlichen Personenbefragungen stammen, bei der Berechnung gleicher Modelle auch gleiche Ergebnisse liefern. Dies wird hier an Hand des SOEP und des ALLBUS untersucht. Es wird gezeigt, dass diese beiden Datensätze durchaus vergleichbare Ergebnisse liefern.

# Inhaltsverzeichnis

**1 Einleitung** **1**

**2 Theoretischer Hintergrund** **6**

2.1 Lineares Wahrscheinlichkeitsmodell . . . . . 7

2.2 Formulierung des Logit-Modells . . . . . 9

**3 Daten und Methode** **11**

3.1 Der Datensatz . . . . . 11

3.2 Die abhängige Variable . . . . . 12

3.3 Die erklärenden Variablen . . . . . 13

**4 Das Modell** **16**

4.1 Bivariate Analysen . . . . . 16

4.2 Modellwahl . . . . . 18

4.3 Das resultierende Modell . . . . . 19

**5 Ergebnisse** **20**

5.1 Die Effekte der exogenen Variablen . . . . . 20

5.2 Zur Güte der Anpassung . . . . . 24

**6 Zusammenfassung und Fazit** **25**

6.1 Zusammenfassung der Ergebnisse im ALLBUS-Modell . . . . . 25

6.2 Vergleich der Ergebnisse der beiden Untersuchungen . . . . . 26

**7 Anhang** **29**

# Tabellenverzeichnis

5.1 Koeffizienten des Logit-Modells . . . . . 21
5.2 Odds Ratios des Logit-Modells . . . . . 22

7.1 Bivariate Analysen 1 . . . . . 30
7.2 Bivariate Analysen 2 . . . . . 31

# Abbildungsverzeichnis

3.1 Histogramm Lebenszufriedenheit . . . . . 12
3.2 allgemeine und persönliche Wirtschaftslage . . . . . 14

4.1 Wirtschaftslage und Lebenszufriedenheit . . . . . 16
4.2 Alter und Lebenszufriedenheit . . . . . 17
4.3 Einkommen und Lebenszufriedenheit . . . . . 18
4.4 Familienstand und Lebenszufriedenheit . . . . . 18

# Kapitel 1

# Einleitung

Sind Sie mit Ihrem Leben zufrieden? Und wenn ja, warum sind Sie zufrieden?

Zufriedenheit ist ein emotionaler Zustand eines jeden Menschen, in dem seine Bedürfnisse von ihm als befriedigt angesehen werden. Das Bedeutungswörterbuch des Duden definiert Zufriedenheit wie folgt:
a) innerlich ausgeglichen und nichts anderes verlangend als man hat.
b) mit den gegebenen Verhältnissen, Leistungen o.ä. einverstanden, nichts auszusetzend habend.

Zufriedenheit ist demnach kein eindeutiges Gefühl. Es kann beliebig differenziert oder mehr oder minder generell gemeint sein. Wie bei den meisten psychologischen Begriffen sind Erfassung und Erhebung noch nicht standardisiert und so gesehen ist die jeweilige Bedeutung weitgehender Unklarheit ausgesetzt. Die einleitend gestellte Frage: "Sind Sie mit Ihrem Leben zufrieden?" lässt sich dennoch von vielen Menschen schnell beantworten, denn die meisten haben ein intuitives Verständnis von dem Wort 'zufrieden' und seiner Bedeutung. Die zweite Frage: "Warum sind Sie zufrieden?" ist schon bedeutend schwieriger zu beantworten, da sich Lebenszufriedenheit auf Grund vieler verschiedener Indikatoren manifestieren kann, und es einem schwer fallen mag, die passenden Erklärungen für die eigene Lebenszufriedenheit zu finden.

Die vorliegende Arbeit beschäftigt sich nun mit der Frage, wodurch sich Lebenszufriedenheit erklären lässt und prüft insbesondere, inwieweit die Einschätzung der

Wirtschaftslage grundlegenden Einfluss auf die Lebenszufriedenheit hat. Dies geschieht unter Hinzunahme anderer sozio-demographischer Faktoren, von denen angenommen wird, dass sie Einfluss auf die Lebenszufriedenheit haben. Darüber hinaus wird an Hand dieser Fragestellung untersucht, ob die Daten aus unterschiedlichen Personenbefragungen - dem ALLBUS und dem SOEP - vergleichbare Ergebnisse liefern.

Die Berechnungen in dieser Arbeit erfolgten mit Daten aus dem ALLBUS[1], eine parallele Arbeit untersucht die gleiche Fragestellung mit Daten aus dem SOEP[2], um danach die Ergebnisse zu vergleichen, Gemeinsamkeiten herauszuarbeiten und Unterschiede festzustellen. Motiviert wurde diese Arbeit zum einen dadurch, dass versucht wurde, eine Fragstellung zu finden, die soziologische (die Lebenszufriedenheit) und ökonomische (die Wirtschaftslage) Aspekte miteinander verbindet.
Zum anderen wurde durch den ständigen Kontakt mit verschiedenen Personenbefragungen die Frage aufgeworfen, ob diese auch vergleichbare Ergebnisse liefern, wenn man dieselbe Fragestellung auf sie anwendet, ob sie demnach beide den Anspruch, repräsentative Umfragedaten über die Bevölkerung Deutschlands zu liefern, erfüllen.

Das Hauptaugenmerk dieser Arbeit liegt wie schon erwähnt auf der Frage, ob gerade die Perzeption der Wirtschaftslage wesentlichen Einfluss auf die Zufriedenheit hat. Die Annahme ist, dass Menschen, die ihre Wirtschaftslage als überdurchschnittlich gut einschätzen, tendenziell dazu neigen, auch ihre Lebenszufriedenheit höher einzuschätzen. Es ist hier natürlich gleich anzumerken, dass man diese Frage durchaus auch anders herum stellen kann: Nehmen Leute, die mit ihrem Leben zufrieden sind, die Wirtschaftslage positiver wahr als solche, die mit ihrem Leben unzufrieden sind? Diese Arbeit soll sich aber lediglich mit der erstgenannten Sichtweise beschäftigen. Die Frage nach der Einschätzung der Wirtschaftslage wird im ALLBUS auf zwei unterschiedliche Arten gestellt: Zum einem wird durch die Frage "Wie beurteilen Sie die Wirtschaftslage in der BRD heute?" die allgemeine Sicht auf die Wirtschaftslage gemessen, zum anderen durch die Frage "Wie beurteilen Sie Ihre persönliche Wirtschaftlage heute?" die Wahrnehmung der eigenen Wirtschaftslage. Beiden liegt eine 5-stufige Skala zugrunde, die von 'sehr gut' (1) bis 'sehr schlecht'

[1] Allgemeine Bevölkerungsumfrage der Sozialwissenschaften, angesiedelt am ZUMA Mannheim
[2] Sozio-oekonomisches Panel, angesiedelt am DIW Berlin

(5) reicht. Für diese Arbeit wurde die zweite Fragevariante als Variable in die Untersuchung miteinbezogen, da zu vermuten ist, dass sich die Lebenszufriedenheit eher an der persönlichen als an der allgemeinen Sicht der Wirtschaftslage festmacht. Im SOEP gibt es nur eine Frage zur allgemeinen Wirtschaftslage. Diese wird wie folgt gestellt: "Machen Sie sich Sorgen um die allgemeine wirtschaftliche Entwicklung?" Als Antwortmöglichkeiten gibt es hier drei Kateogrien: 'keine Sorgen', 'einige Sorgen' und 'große Sorgen'. Um eine spätere Vergleichbarkeit zu gewährleisten, wurde die Variable des ALLBUS-Datensatzes so rekodiert, dass sie ebenfalls nur noch drei Kategorien aufweist: die Antworten 'sehr gut' und 'gut' wurden zusammengefasst, ebenso 'sehr schlecht' und 'schlecht'.

In diese Untersuchung wurden außerdem Variablen mitaufgenommen, von denen man annimmt, dass sie die Lebenszufriedenheit beeinflussen. Dies sind im Einzelnen: das Geschlecht[3], das Alter[4], das Einkommen[5], das Vorhandensein von Kindern, Arbeitslosigkeit, Hochschulbildung, Familienstand[6], Wohnregion und Gemeindegröße. Hier wird zum Beispiel erwartet, dass sich ein höheres Einkommen positiv auf die Lebenszufriedenheit auswirkt[7], wohingegen sich Arbeitslosigkeit negativ auswirken sollte.[8]. Beim Familienstand wird angenommen, dass verheiratete oder zusammenlebende Menschen zufriedener sind als getrennt lebende oder geschiedene[9]. Durch den höheren Lebensstandard in Westdeutschland wird außerdem erwartet, dass Westdeutsche höhere durchschnittliche Werte bei der Lebenszufriedenheit aufweisen.

---

[3] vgl. Robinson, J. und Shaver, P.R. (1973): Measures of social psychological attitudes, MI: Institute for Social Research

[4] vgl. Blanchflower, D.G. und Oswald, A.J. (1996): the Rising Well-Being of the Young, paper presented at an NBER Conference on Disadvantaged Youth, North Carolina

[5] vgl. Frey, B.S. und Stutzer, A. (2001): Happiness and Economics. How the Economy and Institutions affect Human Well-Being, Princeton University Press, Princeton und Oxford

[6] vgl. Oswald, A.J. (1997): Happiness and exonomic performance, The Economic Journal 107, Seite 1815-1831

[7] vgl. Di Tella et al. (1999): How Do Macroeconomics Fluctuations Affect Happiness, Mimeo. Harvard Business School

[8] vgl. Winkelmann, L. und Winkelmann, R. (1998): Why Are the Unemployed So Unhappy? Evidence from Panel Data, Economica, 65, 1-15.

[9] vgl. Fernandez, R.M. und Kulik, J.C. (1981): A multilevel model of life satisfaction. Effects of individual characteristics and neighborhood composition, American Sociological Review, 46 (6:), Seite 840- 850

Das Konstrukt Lebenszufriedenheit gewinnt in den vergangenen Jahren in den Sozial- und Wirtschaftswissenschaften vermehrt an Bedeutung.[10] Lebenszufriedenheit ist eine latente Variable ( Y*), das heißt, dass sie nicht direkt beobachtbar ist. Man kann hier nur über Antworten auf eine Frage zur Lebenszufriedenheit (Y) auf das zugrunde liegende Merkmal Lebenszufriedenheit (Y*) schließen. Im ALLBUS ist die Frage nach der Lebenszufriedenheit auf folgende Weise gestellt:
"Wie glücklich sind Sie alles in allem im Leben?". Als Antwortmöglichkeit steht hierauf eine 7-stufige Skala zur Verfügung, auf der die Antwortmöglickeiten von 1 ('völlig glücklich') bis 7('völlig unglücklich') gegeben sind. Im SOEP lautet die Frage, die das Konstrukt Lebenszufriedenheit abbildet:
"Wie zufrieden sind Sie, alles in allem, mit Ihrem Leben?". Hier steht eine 11-stufige Skala mit den Antwortmöglichkeiten 0 ('völlig unzufrieden') bis 10 ('völlig zufrieden') zur Verfügung.
Wie man sieht, zielen beide Frage auf das Konstrukt Lebenszufriedenheit ab, unterscheiden sich jedoch hinsichtlich ihrer Skalierung und Operationalisierung. Um sie vergleichbar zu machen, werden beide Variablen in dichotome Variablen überführt. Diese neue Variable Y ist zwar weniger effizient als die Ursprungsvariablen, lässt dafür aber am Ende einen besseren Vergleich der beiden Datensätze zu.
Ausgehend von dieser dichotomen Response-Variablen wird hier ein Logit-Modell berechnet, das zur Klasse der verallgemeinerten linearen Modelle zählt.[11]
Die theoretische Herleitung dieses Modells erfolgt in Kapitel 2.
Der verwendetete Datensatz, die Herleitung der dichotomen response-Variablen und die Auswahl der erklärenden Variablen sind Gegenstand des dritten Kapitels.
Im vierten Kapitel werden einige vorausgehende Analysen vorgestellt, denen die Herleitung des zu untersuchenden Modells folgt.
Daran schließen im fünften Kapitel die Ergebnisse der Berechnungen des Logit-

[10] vgl. u.a.
Frey, B. S. und Stutzer, A.(2001)
Strack, F., Argyle, M. und Schwartz, N.(Hg.)(1991): Subjective well-being. An interdisciplinary perspective, Pergamon Press, London UK
Veenhoven, Ruut: Die Rückkehr der Ungleichheit in die moderne Gesellschaft? Die Verteilung der Lebenszufriedenheit in den EU-Ländern von 1973 bis 1996, in: Wolfgang Glatzer, Roland Habich und Karl-Ulrich Maier(Hg.)(2002): Sozialer Wandel und Gesellschaftliche Dauerbeobachtung. Festschrift für Wolfgang Zapf, Leske+Bundrich, Opladen Deutschland, Seite 273-294

[11] vgl. Rönz, B. (2001): Generalisierte Lineare Modelle, Berlin

Modells an.
Abschließend erfolgt ein Vergleich mit den Ergebnissen aus der Analyse mit den Daten des SOEP-Datensatzes, wobei die gefundenen Ergebnisse nochmals kritisch diskutiert werden.

# Kapitel 2

# Theoretischer Hintergrund

Im Rahmen der vorliegenden Arbeit wird als abhängige Variable die allgemeine Lebenszufriedenheit untersucht. Es wird angenommen, dass dieses Merkmal (Y*) metrisches Skalenniveau aufweist, jedoch nicht direkt beobachtbar ist. Es handelt sich also um eine latente Variable. In Abhängigkeit von dem nicht bekannten Wert einer Person hinsichtlich des latenten Merkmals Y* kann eine dichotome Variable Y beobachtet werden, die dann den Wert 1 annimmt, wenn der Wert der latenten kontinuierlichen Variable Y* größer ist als 0:

$$y_i = \begin{cases} 1 \text{ falls } y_i^* > 0 \\ 0 \text{ falls } y*_i \leq 0 \end{cases} \tag{2.1}$$

Zur Erklärung der latenten Variable durch exogene Größen X kann ein lineares Modell der Form

$$y_i^* = X_i^T \beta + u_i^* \tag{2.2}$$

aufgestellt werden.

In einem Modell zur Erklärung der beobachtbaren dichotomen Variable Y soll im folgenden für jede Beobachtung die Wahrscheinlichkeit $\pi_i$ modelliert werden, dass diese Variable den Wert 1 annimmt. Es ist also:

$$\begin{aligned} \pi_i &= P(Y_i = 1) = P(Y_i^* > 0) = P(u_i^* > -X_i^T \beta) = 1 - P(u_i^* \leq -X_i^T \beta) \\ &= 1 - F(-X_i^T \beta) = F(X_i^T \beta) \end{aligned} \tag{2.3}$$

Ein Logit-Modell ist dann spezifiziert, wenn als Funktion F die logistische Verteilungsfunktion gewählt wird.

Im folgenden Abschnitt soll gezeigt werden, warum das einfachere lineare Modell zur Modellierung der Wahrscheinlichkeit $\pi_i$ nicht angemessen ist.

## 2.1 Lineares Wahrscheinlichkeitsmodell

Die einfachste Herangehensweise an dieses Problem ist, analog zu der Untersuchung einer metrischen abhängigen Variablen ein lineares Modell aufzustellen[1]:

$$y_i = X_i^T \beta + u_i \tag{2.4}$$

Die grundlegenden Annahmen des OLS-Modells[2] sind jedoch eine lineare Beziehung zwischen den unabhängigen Variablen und der abhängigen Variablen und normalverteilte Störterme und homoskedastische Störterme. Dass diese Annahmen im Fall einer binären response-Variable nicht erfüllt werden, soll hier gezeigt werden:

1. Die als abhängige Variable modellierte Wahrscheinlichkeit $\pi_i$ dafür, dass ein bestimmtes Ereignis eingetreten ist ($Y_i = 1$), ist zwar kontinuierlich, jedoch nur zwischen 0 und 1. Im OLS-Modell liegen die über die Linearkombination

   $$E(y_i/X_i) = X_i^T \hat{\beta} \tag{2.5}$$

   geschätzten Werte, die den auf $X_i$ bedingten Erwartungswerten entsprechen, auf einer Geraden, können also theoretisch Werte zwischen $-\infty$ und $+\infty$ annehmen.

   Da Werte außerhalb des [0,1]-Intervalls jedoch nicht als Wahrscheinlichkeit interpretiert werden können, ist eine solche Verknüpfung zwischen den unabhängigen und der abhängigen Variable nicht sinnvoll.

[1] Vgl. Greene, William (2003):Economic Analysis, Prentice-Hall, New Jersey, Seite. 665 ff.
[2] Ordinary Least Squares, Kleinste-Quadrate-Schätzer

2. Die Störterme nehmen in Abhängigkeit von $Y_i$ folgende Werte an:

$$\epsilon_i = \begin{cases} y_i - \pi_i = 1 - \pi_i \text{ mit Wahrscheinlichkeit } \pi_i, & \text{wenn } y_i = 1 \\ y_i - \pi_i = -\pi_i \quad \text{mit Wahrscheinlichkeit } 1 - \pi_i, & \text{wenn } y_i = 0 \end{cases} \tag{2.6}$$

   Ist also die abhängige Variable $Y_k$ bernoulli-verteilt, so sind es auch die Störterme $\epsilon_i$. Somit ist die zentrale Annahme der normalverteilten Störterme des OLS-Modells verletzt.

3. Zwar ist der Erwartungswert des dichotomen Fehlerterms $\epsilon_i$ 0, doch ist die Varianz heteroskedastisch:

$$\begin{aligned} E(\epsilon_i) &= \sum_{i=1}^{N} \epsilon_i \cdot p(\epsilon_i) \\ &= (1 - \pi_i) \cdot \pi_i + (-\pi_i) \cdot (1 - \pi_i) \\ &= 0 \end{aligned} \tag{2.7}$$

$$\begin{aligned} V(\epsilon_i) &= \sum_{i=1}^{N} (\epsilon_i - E(\epsilon_i))^2 \cdot p(\epsilon_i) \\ &= (1 - \pi_i)^2 \cdot \pi_i + (-\pi_i)^2 \cdot (1 - \pi_i) \\ &= \pi_i \cdot (1 - \pi_i) \end{aligned} \tag{2.8}$$

   Es ist also auch die Annahme der homoskedastischen Störterme verletzt.

Aufgrund der hier aufgeführten Charakteristika ist es nicht angemessen, bei einer dichotomen response-Variablen ein lineares Modell zu rechnen, das heißt, dass in diesem Fall die geschätzten Werte der bedingten Wahrscheinlichkeit $\hat{\pi}_i$ nicht wie im klassischen Regressionsmodell über die Linearkombination $X_i^T \hat{\beta}$ berechnet werden können.

Dies führt über zur Konzeption des Logit-Modells,welches zur Klasse der verallgemeinerten linearen Modelle zählt.

## 2.2 Formulierung des Logit-Modells

Da die als abhängige Variable modellierte bedingte Wahrscheinlichkeit $\pi_i$ per Definition nur Werte im Intervall [0, 1] annehmen darf, die Linearkombination $X_i^T\beta$ jedoch theoretisch Werte zwischen $-\infty$ und $+\infty$ annehmen kann, ist es naheliegend, diese Linearkombination einer Transformation zu unterziehen. Diese Transformationsfunktion sollte die folgenden Eigenschaften aufweisen:

$$\begin{aligned} F(-\infty) = 0 \text{ und } \quad F(\infty) = 1, \\ f(x) = F'(x) > 0. \end{aligned} \tag{2.9}$$

Diese Eigenschaften sind bei Verteilungsfunktionen von Wahrscheinlichkeitsdichten gegeben.

Im Falle des Logit-Modells wird als Transformationsfunktion die logistische Verteilungsfunktion herangezogen:

$$F(\eta_i) = \frac{e^{\eta_i}}{1 + e^{\eta_i}} \tag{2.10}$$

sodass

$$F(\eta_i) = F(X_i^T\beta) = \frac{e^{X_i^T\beta}}{1 + e^{X_i^T\beta}} = \pi_i. \tag{2.11}$$

resultiert. Somit sind die bedingten zu schätzenden Wahrscheinlichkeiten $\pi_i$ eine nichtlineare Funktion der erklärenden Variablen, deren Werte aufgrund der Eigenschaften der logistischen Verteilungsfunktion immer in dem [0, 1]-Intervall liegen. Löst man obige Gleichung 2.11 nach $X_i^T\beta$ auf, erhält man:

$$X_i^T\beta = log\left(\frac{\pi_i}{1 - \pi_i}\right) = g(\pi_i). \tag{2.12}$$

Diese Funktion $log\left(\frac{\pi_i}{1-\pi_i}\right)$ ist die Inverse der logistischen Verteilungsfunktion und wird auch Linkfunktion genannt, da sie die Linearkombination $X_i^T\beta$ mit den zu schätzenden Werten der bedingten Wahrscheinlichkeit $\pi_i$ verknüpft. Sie wird auch als Logit-Funktion bezeichnet, wobei die Logits die logarithmierten Odds[3] sind.

[3] Odds sind das Verhältnis einer Erfolgswahrscheinlichkeit $\pi_i$ zur Nichterfolgswahrscheinlichkeit $1 - \pi_i$, auch Chance genannt, Bsp: Gewinnquoten

Diese Modellierung des Erwartungswertes der response-Variablen unter Zuhilfenahme einer Linkfunktion ist eine Komponente der verallgemeinerten Modelle.

Ein weiterer Aspekt dieser Modellklasse ist, dass die response-Variable eine Verteilung aufweist, die zur Familie der exponentiellen Verteilungen zählt. Dies ist im Fall einer Bernoulli-Verteilung[4] gegeben.

Da die resultierenden Gleichungssysteme für die $\hat{\beta}_k$ nichtlinear sind, werden sie mit Maximum-Likelihood-Verfahren iterativ gelöst.

Aufgrund des nichtlinearen Zusammenhangs zwischen der geschätzten bedingten Wahrscheinlichkeit $\pi_i$ und $X_i^T\beta$ (siehe Gleichung 2.11) auf Seite 9 sind die geschätzten Koeffizienten $\hat{\beta}_i$ nicht so einfach zu interpretieren wie im linearen Regressionsmodell.

Im Logit-Modell gibt der Koeffizient $\hat{\beta}_i$ an, wie sich die Logits verändern (siehe Gleichung 2.12 auf Seite 9), wenn sich die exogene Variable um eine Einheit erhöht, bzw. eine andere Kategorie als die Referenzkategorie annimmt. In Hinsicht auf die Erfolgswahrscheinlichkeit $\pi_i$ können die Koeffizienten $\hat{\beta}_i$ nur bezüglich der Richtung (Vorzeichen) interpretiert werden, in die die exogene Variable die Erfolgswahrscheinlichkeit beeinflusst, nicht aber hinsichtlich der Stärke des Zusammenhangs.

Hinsichtlich der Stärke des Zusammenhangs können jedoch die Odds Ratios (Verhältnis zweier Odds) interpretiert werden, was gleichbedeutend mit dem Wert $e^{\hat{\beta}_i}$ ist. So gibt dieser Wert den Faktor an, um welchen sich die Chance (dass die Variable Y den Wert 1 annimmt) erhöht, bzw. verringert, wenn eine Variable eine andere Kategorie als die Referenzkategorie annimmt.

[4] Verteilung einer binären Zufallsvariablen, die mit der Erfolgswahrscheinlichkeit des Parameters p den Wert 1 annimmt.

# Kapitel 3

# Daten und Methode

## 3.1 Der Datensatz

Der dieser Arbeit zugrunde liegende Datensatz ist der ALLBUS, die allgemeine Bevölkerungsumfrage der Sozialwissenschaften. Darin werden seit 1980 alle zwei Jahre in einer repräsentativen Querschnittsbefragung aktuelle Daten über Einstellungen, Verhaltensweisen und Sozialstruktur der Bevölkerung in der Bundesrepublik Deutschland erhoben. Der ALLBUS wird vom Zentrum für Umfragen, Methoden und Analysen (ZUMA) in Mannheim durchgeführt. Für diese Arbeit wurde die Befragung aus dem Jahr 2002 herangezogen. Der Stichprobenumfang des ALLBUS ist im Vergleich zum SOEP relativ gering, im Jahr 2002 betrug er nur n=2820. Die Frage zur Lebenszufriedenheit ("wie glücklich sind Sie alles in allem im Leben?"), die in dieser Analyse die abhängige Variable ist, wurde jedoch nur von n=1328 Personen beantwortet. Dies liegt daran, dass zusammen mit dem ALLBUS auch noch Daten für das ISSP ('International Social Survey Programme') erhoben werden und diese Frage nach der Lebenszufriedenheit nur an Personen gestellt wurde, die auch in der ISSP-Studie vertreten waren. Die folgende Analyse kann demnach nur mit einem Stichprobenumfang von n=1328 durchgeführt werden. Ob dieser doch recht geringe Stichprobenumfang problematisch für diese Arbeit ist, wird sich später noch herausstellen.

## 3.2 Die abhängige Variable

Lebenszufriedenheit ist eine latente Variable. Sie wird im ALLBUS über die Frage "wie glücklich sind sie alles in allem im Leben?" erfasst. Der Variablen liegt eine 7-stufige Skala zugrunde, die von 1 (völlig glücklich) bis 7 (völlig unglücklich) reicht. Da die Variable, die im SOEP das Konstrukt Lebenszufriedenheit misst, anders

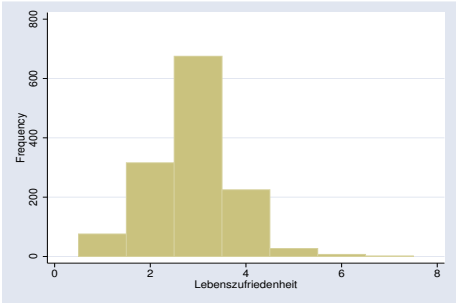


Abbildung 3.1: Histogramm Lebenszufriedenheit

skaliert ist -der Variablen liegt dort eine 11-stufige Skala zugrunde- kann man die Variablen in Originalform schlecht vergleichen. Es liegt also nahe, die Variablen so zu rekodieren, dass eine Vergleichbarkeit gewährleistet ist. Als Methode wurde dafür der Mediansplit gewählt, durch den die Ausgangsvariable in eine dichotome Variable übergeführt wird. Durch diese Dichotomisierung entsteht zwar ein Informationsverlust in den Daten, der hier jedoch in Kauf genommen wird, da eine Aufteilung in drei oder mehr Kategorien die Möglichkeit des Vorhandenseins von Nullzellen in der mulitvariaten Analyse steigert, wodurch die Schätzungen letztendlich verzerrt würden.

Der Median ist das 50 %-Quantil. Er ist anders als zum Beispiel der Mittelwert ein robuster Lageparameter, weil er beispielsweise relativ robust gegen Ausreißer ist. Die Ausgangsvariable wird also in eine dichotome Variable übergeführt, in dieser nimmt die eine Ausprägung all die Fälle auf, die oberhalb des Medians liegen, die andere Ausprägung alle, die unterhalb des Medians liegen.

$$y_i = \begin{cases} 1 & \text{,wenn} \quad z_i > z_{0.5} \quad \text{,d.h.} \quad z_i \in [1,2] \\ 0 & \text{,wenn} \quad z_i \leq z_{0.5} \quad \text{,d.h.} \quad z_i \in [3,7] \end{cases} \tag{3.1}$$

In diesem Fall liegt der Median bei 3. Dies bedeutet, dass allen Personen, die

in der Ausgangsvariablen die Werte 1 und 2 haben -also glücklicher als der Median sind-, in der neuen Variablen den Wert 1 zugewiesen bekommen, alle anderen den Wert 0. Durch diese Rekodierung der Ausgangsvariablen "wie glücklich sind Sie alles in allem im Leben?" in eine dichotome Variable ist diese jetzt auch mit dem Konstrukt der Variablen Lebenszufriedenheit aus dem SOEP vergleichbar. Diese dichotome Variable bildet in der folgenden Analyse die abhängige Variable.

## 3.3 Die erklärenden Variablen

Im Gegensatz zu der parallel erstellten Arbeit, die mit den Daten aus dem SOEP arbeitet, wurden in dieser Arbeit einige Einschränkungen vorgenommen: auf Grund der sehr viel geringeren Fallzahl wurden nicht alle Variablen, die im Modell mit dem SOEP-Daten vertreten sind, für dieses Modell herangezogen.
Nicht aufgenommen wurde eine Variable zur Nationalität, da im ALLBUS nur n=81 nicht deutsche Staatsbürger befragt wurden, von denen wiederum nur n=37 die Frage zur Lebenszufriedenheit gestellt wurde. Außerdem wird im ALLBUS keine Unterscheidung zwischen EU-Ausländern und Ausländern anderer Nationalität gemacht. Ebenfalls nicht aufgenommen wurde die Variable zur Art der Beschäftigung, also ob Vollzeit, Teilzeit oder im Nebenerwerb gearbeitet wird. Im Hinblick auf die geringe Grundgesamtheit und die daraus resultierende Gefahr zu vieler Nullzellen bei Aufnahme zu vieler Variablen wird diese Frage einfach durch die Variable zur Arbeitslosigkeit subsumiert, und eine Aufteilung der Erwerbstätigen nach der Art ihrer Erwerbstätigkeit entfällt. Auch eine Variable zum Vorhandensein eines Berufsabschlusses ist in diesem Modell nicht zu finden, dies jedoch aus einem anderen Grund: Eine Variable, die das Vorhandensein eines Berufsabschlusses misst, ist im ALLBUS-Datensatz in dieser Weise nicht vorhanden.
Noch ein wichtiger Unterschied im Gegensatz zum Modell mit den Daten aus dem SOEP ist an dieser Stelle anzumerken:
In dieses Modell geht an Stelle einer Variablen zur Einschätzung der allgemeinen Wirtschaftslage eine Variable zur Einschätzung der persönlichen Wirtschaftslage ein, da es nahe liegt, den Zusammenhang zwischen individueller Lebenszufriedenheit und der Einschätzung der Wirtschaftslage eher mit der individuellen Wirtschaftslage als mit der allgemeinen Wirtschaftslage herzustellen. Wie nachfolgende Grafik

zeigt, ähneln sich im ALLBUS die Variablen zur Einschätzung der allgemeinen und der persönlichen Wirtschaftslage. In beiden Fällen liegt der Median bei 3, wobei es natürlich Unterschiede in der Verteilung um den Median gibt: Die persönliche Wirtschaftslage wird eher besser als die allgemeine Wirtschaftslage eingeschätzt,aber dies soll hier nicht weiter stören. Die Aufnahme der Variablen zur Einschätzung der persönlichen Wirtschaftslage und der spätere Vergleich mit einer Variablen zur allgemeinen Lebenszufriedenheit aus dem SOEP scheint durch die Ähnlichkeit der beiden Variablen im ALLBUS somit durchaus angemessen.

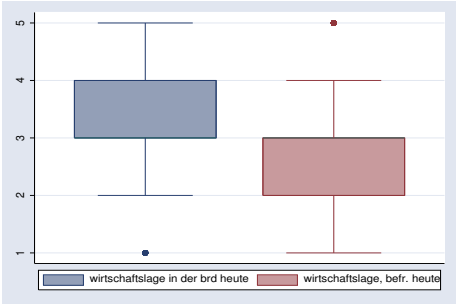


Abbildung 3.2: allgemeine und persönliche Wirtschaftslage

Als erklärende Variablen sollen letztendlich folgende Variablen in das Modell aufgenommen werden:

- das Geschlecht
- das Alter
- das persönliche Nettomonatseinkommen in Euro
- eine Dummy-Variable für Arbeitslosigkeit[1]
- eine Dummy-Variable für das Vorhandensein eines (Fach)-Hochschul-Abschlusses
- das Erhebungsgebiet aufgeteilt nach West- und Ostdeutschland

[1] zur Konstruktion der Variablen Arbeitslosigkeit: als Arbeitslose wurden solche Fälle gewertet, die nicht erwerbstätig sind und als Grund dafür den Status 'zur Zeit arbeitslos', 'Hausfrau/Hausmann' oder 'nicht berufstätig' nannten. Demnach werden hier andere Erwerbslose wie Schüler, Studenten, Rentner und Wehrpflichtige nicht zu den Arbeitslosen gezählt.

- die Größe der politischen Gemeinde mit den Kategorien ländlich[2], Verdichtungsansatz[3] und städtisch[4], wobei die letzte Kategorie als Referenzkategorie dient
- eine Dummy-Variable zum Vorhandensein von Kindern
- eine Variable zum Familienstand mit den Kategorien verheiratet/zusammen lebend, die als Referenzkategorie herangezogen wird, getrennt lebend/geschieden, verwitwet und ledig
- eine Variable zur Beurteilung der aktuellen persönlichen wirtschaftlichen Lage mit den Kategorien sehr gut/gut, teils/teils und schlecht/sehr schlecht. Hier wird als Referenzkategorie die Kategorie sehr gut/gut gewählt.

Abschließend soll hier noch auf zwei Unterschiede zwischen den Variablen des ALLBUS-Modells und des SOEP-Modells eingegangen werden:
Das Vorhandensein von Kindern bezieht sich im SOEP auf Kinder unter 21 Jahren, im ALLBUS gibt es hierzu keine Einschränkung. Bei der Variablen, die den Bildungsabschluss misst, finden sich auch Unterschiede in der Konstruktion: Im ALLBUS ist dies eine dichotome Variable, in der alle den Wert 1 annehmen, die einen Hochschul- oder Fachhochschulabschluss besitzen. Demnach gehen in die andere Kategorie all diejenigen ein, die keinen Hochschulabschluss besitzen, also auch Personen mit Berufsabschluss. Im SOEP wurden die Variablen zum Berufsabschluss und Bildungsabschluss zuerst als dreistufige Variable modelliert, also mit den Kategorien 'kein Abschluss', 'Berufsabschluss' und 'Hochschulabschluss'. Danach wurden daraus zwei dichotome Variablen gebildet. Demnach sind bei der Variablen zum Vorhandensein eines Hochschulabschlusses in der einen Kategorie alle Personen, die einen Hochschulabschluss haben und in der anderen, alle, die überhaupt keinen Abschluss haben.

[2] unter 2000 Einwohner
[3] zwischen 2000 und 20 000 Einwohner
[4] über 20 000 Einwohner

# Kapitel 4

# Das Modell

## 4.1 Bivariate Analysen

Um einen ersten Eindruck von den Daten zu bekommen, wurden zunächst eine bivariate Analyse zwischen den einzelnen gewählten erklärenden Variablen und der Response-Variablen durchgeführt.[1] Bei diesen Analysen zeigt sich, dass -zumindest im bivariaten Bereich- der vermutete Zusammenhang zwischen Einschätzung der Wirtschaftslage und Lebenszufriedenheit besteht. Als Test wurde hierzu der $\chi^2$-Unabhängigkeitstest verwendet, der mit p=0.000 zeigt, dass es Abhängigkeiten zwischen den beiden untersuchten Variablen gibt. Der $\chi^2$-Test sagt allerdings nichts über die Richtung dieser Abhängigkeit aus. Diesen Zusammenhang verdeutlicht auch das folgende Balkendiagramm:

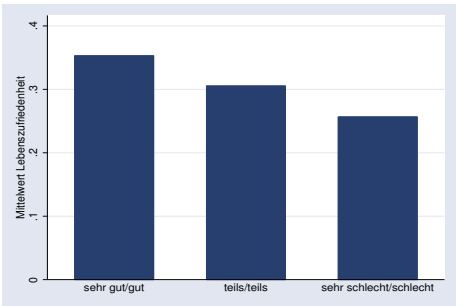


Abbildung 4.1: Wirtschaftslage und Lebenszufriedenheit

[1] siehe Anhang Tabelle 7.1 und Tabelle 7.2

Zwischen den einzelnen sozio-demographischen Faktoren und der Lebenszufriedenheit finden sich in den bivariaten Analysen wenig Zusammenhänge.
Für die bivariaten Analysen wurde das Alter und das Einkommen klassiert, um nicht gegen die Approximationsbedingungen des $\chi^2$-Unabhängigkeitstests zu verstoßen[2]. Nur bei Alter und Familienstand hat der $\chi^2$-Unabhängigkeitstests eine Signifikanz von p=0.000. In allen anderen Fällen gibt der Test eine Signifikanz von über 0.05, das heißt dass die beobachteten Merkmale nicht abhängig voneinander sind. Zum Verdeutlichen dieser bivariaten Zusammenhänge und um einen besseren Eindruck von den Verteilungen der zugrunde liegenden Variablen zu bekommen, wurden einige graphische Darstellungen modelliert. Wie man erkennen kann, unterscheiden sich

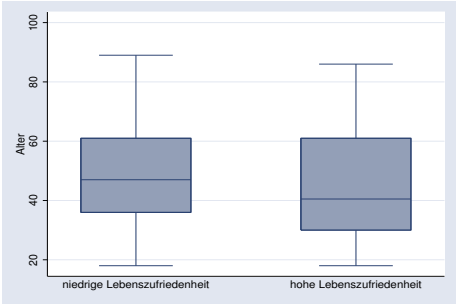


Abbildung 4.2: Alter und Lebenszufriedenheit

die Boxplots für niedrige und hohe Lebenszufriedenheit hinsichtlich des Alters.
Bezüglich des Einkommens, für das bei der bivariaten Analyse kein Zusammenhang mit der Lebenszufriedenheit gefunden wurde, unterscheiden sich diese Boxplots nicht sehr voneinander. Um den gefundenen Zusammenhang zwischen Familienstand und Lebenszufriedenheit zu verdeutlichen, wurde ein Balkendiagramm angefertigt, auf dem der Familienstand auf der X-Achse und der Mittelwert der Variable Lebenszufriedenheit für die einzelnen Kategorien auf der Y-Achse abgetragen wurde.

[2] Approximationsbedingungen:
1. erwartete Häufigkeit in jeder Zelle größer als 0
2. nicht mehr als 20% der Zellen dürfen eine erwartete Häufigkeit kleiner als 5 aufweisen

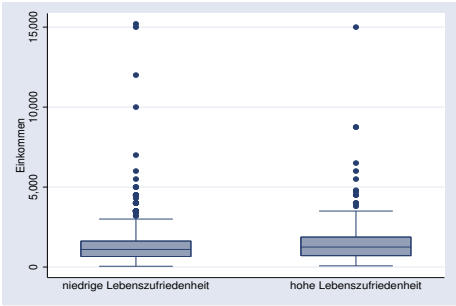


Abbildung 4.3: Einkommen und Lebenszufriedenheit

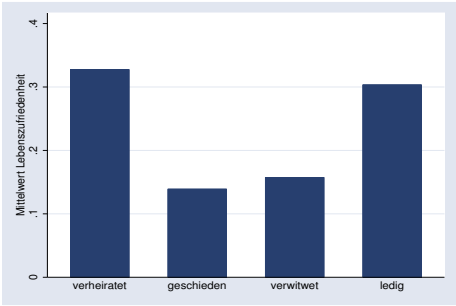


Abbildung 4.4: Familienstand und Lebenszufriedenheit

## 4.2 Modellwahl

Nach den Ergebnissen der bivariaten Analysen stellt sich die Frage, inwieweit die Aufnahme aller Variablen in das Modell gerechtfertigt ist. Als Vorgehensweise wurde die Backward-Selektion gewählt. Hierfür wurde zunächst ein Logit-Modell mit allen Variablen, die in die Analyse aufgenommen werden sollen, berechnet. Danach wurde die Variable, die die geringste Signifikanz aufwies (in diesem Fall ist dies das Einkommen), aus diesem Modell herausgenommen und es wurde erneut ein Logit-Modell gerechnet. Nun wurde mittels eines likelihood-ratio-Tests geprüft, ob sich das zweite Modell im ersten wiederfindet. Wäre dies der Fall, ist es nicht notwendig, das Modell mit allen Variablen zu rechnen, sondern es würde sich anbieten, diese Variablen herauszunehmen. Daraufhin könnte man die Prozedur so oft wiederholen, bis das kleinere Modell nicht mehr im größeren wiederzufinden ist. Da hier die Annahme, dass sich das zweite Modell im ersten finden lässt, jedoch mit p=0.000 abgelehnt wurde, ist es angebracht, alle Variablen im Modell zu belassen.

## 4.3 Das resultierende Modell

Die wenigen gefundenen Zusammenhänge bei den bivariaten Analysen lassen vermuten, dass es auch im multivariaten Bereich wenige Zusammenhänge zwischen denen zur Erklärung herangezogenen Variablen und der Lebenszufriedenheit gibt. Das Ergebnis des likelihood-ratio-Tests legt jedoch nahe, dass alle gewählten Variablen in das Modell mit aufgenommen werden sollen. Selbst bei anderslautenden Ergebnissen des likelihood-ratio-Tests sollte man, um den Vergleich mit der parallel angefertigten Arbeit zu gewährleisten, dazu tendieren, dennoch alle in Kapitel 3.3 aufgeführten Variablen in die Analyse aufzunehmen.
Grundsätzlich sind alle aufgenommenen Variablen grundlegende soziodemographische Faktoren, die in ein solches Modell mitaufgenommen werden sollten.
Demnach ergibt sich folgendes zu schätzendes Modell:

$$P(z_i > z_{0.5}/x_i^T) = P(y_i = 1/x_i^T) = F(x_i^T\beta) = \frac{e^{x_i^T\beta}}{1 + e^{x_i^T\beta}} = \pi_i, \tag{4.1}$$

wobei der Vektor $x_i^T$ für jede Beobachtung die Werte der oben genannten exogenen Variablen enthält. Die Ergebnisse dieses Modells und deren Interpretation sind Gegenstand des folgenden Kapitels.

# Kapitel 5

# Ergebnisse

## 5.1 Die Effekte der exogenen Variablen

Einleitend soll zunächst auf einige Probleme des ALLBUS-Datensatzes eingegangen werden, welche sich sowohl in den folgenden Ergebnissen als auch schon in der bivariaten Analyse der Variablen teilweise widerspiegeln. Der ALLBUS-Datensatz ist -wie schon in den Kapitel 3.1. und 3.3. kurz erwähnt- relativ klein. Schon im Vorfeld liegt es daher nahe zu vermuten, dass der ALLBUS-Datensatz weniger signifikante Ergebnisse liefert als ein größerer Datensatz wie zum Beispiel der SOEP. Bedingt durch die geringe Grundgesamtheit ist auch nicht auszuschließen, dass er zum Teil ein verzerrtes Bild liefert, weil sich zum Beispiel Ausreißer verstärkt in den Ergebnissen widerspiegeln als bei einer Analyse mit einem größeren Datensatz.

Mit den Ergebnissen aus den bivariaten Analysen überrascht auch der Blick auf das Logit-Modell nicht:

Die zwei Kategorien zur Einschätzung der Wirtschaftslage, die in das Modell mitaufgenommen worden sind, unterscheiden sich im Hinblick auf die Lebenszufriedenheit signifikant von der Referenzkategorie (p=0.000 bzw. p=0.002) Bei einer Einschätzung der persönlichen Wirtschaftslage mit 'teils/teils' im Vergleich mit der Referenzkategorie 'sehr gut/gut' senkt sich die Wahrscheinlichkeit überdurchschnittlich lebenszufrieden zu sein um den Faktor 0.58. Ebenso senkt sich die Wahrscheinlichkeit, überdurchschnittlich lebenszufrieden zu sein bei der Kategorie 'schlecht/sehr

Tabelle 5.1: Koeffizienten des Logit-Modells

| **Variable** | **Coef.** | **Std. Err.** | **z** | **P>** $\|z\|$ |
|---|---|---|---|---|
| teils/teils | -.5486 | .1574 | -3.49 | 0.000 |
| schlecht/sehr schlecht | -.8608 | .2803 | -3.07 | 0.002 |
| Alter | -.0131 | .0056 | -2.34 | 0.019 |
| Einkommen | .00001 | .00006 | 0.24 | 0.812 |
| arbeitslos | -.4868 | .2478 | -1.96 | 0.049 |
| Hochschulabschluss | -.14208 | .1884 | -0.75 | 0.451 |
| Geschlecht | -.1142 | .1506 | -0.76 | 0.448 |
| West/Ost | .1817 | .1621 | 1.12 | 0.262 |
| ländlich | .2513 | .2145 | 1.17 | 0.241 |
| Verdichtungsansatz | .0857 | .1567 | 0.55 | 0.584 |
| Kinder | -.0383 | .2179 | -0.18 | 0.860 |
| getrennt lebend/geschieden | -1.0652 | .3281 | -3.25 | 0.001 |
| verwitwet | -.7658 | .36269 | -2.11 | 0.035 |
| ledig | -.3283 | .26073 | -1.26 | 0.208 |
| Konstante | .1278 | .4146 | 0.31 | 0.758 |

schlecht' im Vergleich zur Referenzkategorie 'gut/sehr gut' um den Faktor 0.422. Anders ausgedrückt:
Die Wahrscheinlichkeit, dass jemand, der seine persönliche Wirtschaftslage als schlecht oder sehr schlecht ansieht, überdurchschnittlich lebenszufrieden ist, ist weniger als halb so hoch wie die Wahrscheinlichkeit, dass jemand, der seine persönliche Wirtschaftslage als sehr gut oder gut einschätzt, überdurchschnittlich lebenszufrieden ist. Außer den Variablen zur Wirtschaftslage ist auch das Alter signifikant. Hier sinkt bei zunehmendem Alter die Wahrscheinlichkeit, dass jemand überdurchschnittlich lebenszufrieden ist. Dieser Effekt ist jedoch recht gering, die Wahrscheinlichkeit, überdurchschnittlich lebenszufrieden zu sein, senkt sich pro Jahr um den Faktor 0.987. Da das Alter nicht kategorisiert in das Modell aufgenommen wurde, verwundert dieser geringe Effekt nicht: Man kann nicht wirklich davon ausgehen, dass sich pro Lebensjahr große sichtbare Effekte ergeben. Hier zählt die Tendenz des Effekts, die besagt, dass die Wahrscheinlichkeit, dass ältere Personen überdurchschnittlich lebenszufrieden sind, geringer ist als die Wahrscheinlichkeit, dass jüngere dies sind. Ein Effekt der Hochschulbildung auf die zu untersuchende Variable konnte in diesem Modell nicht festgestellt werden (p=0.451). Demnach hat eine bessere formale

Tabelle 5.2: Odds Ratios des Logit-Modells

| **Variable** | **OR** | **Std. Err.** | **z** | **P**$> \|z\|$ |
|---|---|---|---|---|
| teils/teils | .5778 | .0909 | -3.49 | 0.000 |
| schlecht/sehr schlecht | .4228 | .1185 | -3.07 | 0.002 |
| Alter | .987 | .0055 | -2.34 | 0.019 |
| Einkommen | 1.0001 | .00006 | 0.24 | 0.812 |
| arbeitslos | .6146 | .1523 | -1.96 | 0.049 |
| Hochschulabschluss | .8676 | .1635 | -0.75 | 0.451 |
| Geschlecht | .8921 | .1344 | -0.76 | 0.448 |
| West/Ost | 1.1992 | .1944 | 1.12 | 0.262 |
| ländlich | 1.2857 | .2758 | 1.17 | 0.241 |
| Verdichtungsans. | 1.0895 | .1707 | 0.55 | 0.584 |
| Kinder | .9624 | .2097 | -0.18 | 0.860 |
| getrennt lebend/geschieden | .3446 | .11307 | -3.25 | 0.001 |
| verwitwet | .4649 | .1686 | -2.11 | 0.035 |
| ledig | .7202 | .1878 | -1.26 | 0.208 |

Bildung keinen Einfluss auf die persönliche Lebenszufriedenheit. Anzunehmen wäre, dass sich hier positive und negative Effekte des höheren Bildungsniveaus aufheben: Einerseits führt bessere formale Bildung‘in der Regel zu einem besseren sozialen Status, einem höheren Einkommen und dadurch zu einem ‘sorgenfreieren’ Leben. Andererseits neigen Personen mit höherem Bildungsniveau auch dazu, ihre Umwelt und ihre Lebensumstände differenzierter und kritischer wahrzunehmen als Personen mit niedrigerem Bildungsniveau -sie schauen über ihren eigenen Tellerrand hinaus-, weswegen sie nicht so leicht zufrieden zu stellen sind. Objektiv könnten sie demnach glücklicher sein, subjektiv braucht es aber mehr, um sie wirklich zufrieden zu stellen.

Signifikante Effekte zeigen sich bei den Variablen zum Familienstand: Hier wurde als Referenzkategorie die Kategorie ‘verheiratet/zusammen lebend’ gewählt. Es zeigt sich, dass Personen, die getrennt leben beziehungsweise geschieden sind, weniger lebenszufrieden sind als solche, die verheiratet sind (p=0.001). Auch verwitwete Personen sind weniger lebenszufrieden als verheiratete(p=0.035). Bei Ledigen kann kein signifikanter Unterschied zu Verheirateten hinsichtlich der Lebenszufriedenheit in diesem Modell festgestellt werden (p=0.208). Die Wahrscheinlichkeit, als geschiedene oder getrennt lebende Person hoch lebenszufrieden zu sein, senkt sich

im Vergleich zu verheirateten Personen um den Faktor 0.345. Das bedeutet, dass die Wahrscheinlichkeit, als geschiedene oder getrennt lebende Person hoch zufrieden zu sein, nur ungefähr ein Drittel so hoch ist wie bei verheirateten oder zusammenlebenden Personen. Bei verwitweten Personen ist diese Wahrscheinlichkeit im Vergleich zu verheirateten Personen auch nur ungefähr halb so hoch: Hier senkt sich die Wahrscheinlichkeit, hoch zufrieden zu sein, im Vergleich zu verheirateten Personen um den Faktor 0.465.
Demnach trägt ein stabiles familiäres Umfeld zur Lebenszufriedenheit des Einzelnen bei. Personen, die ihren Partner durch Scheidung oder Tod verloren haben, fehlt dieses stabile Netzwerk. Bei Ledigen ist zum einen zu vermuten, dass hier auch solche Personen als ledig gezählt werden, die in einer festen, aber nicht formal anerkannten Partnerschaft leben. Zum anderen geht gerade bei jüngeren Personen die Tendenz in die Richtung, dass ein soziales Netzwerk außerhalb der Familie mehr und mehr das Netzwerk Familie ersetzt und daher nicht unbedingt der feste Partner als stabilisierender Faktor benötigt wird, um lebenszufrieden zu sein.

Beim Betrachten der Variablen Geschlecht zeigt sich, dass es keine Unterschiede hinsichtlich der Lebenszufriedenheit bei Männern und Frauen gibt(p=0.448).

Ebenfalls keine Unterschiede zeigen sich zwischen Personen mit Kindern und kinderlosen Personen(p=0.86). Im Gegensatz zu den bivariaten Analysen ist die Arbeitslosigkeit im Logit-Modell signifikant. Die Tatsache, dass jemand arbeitslos ist, senkt die Wahrscheinlichkeit, hoch lebenszufrieden zu sein, um den Faktor 0.615: Die Wahrscheinlichkeit, hoch zufrieden zu sein, ist also bei einem Arbeitslosen nur knapp zwei Drittel so groß wie bei einem Erwerbstätigen.[1] Mit Arbeitslosigkeit hängen viele verschiedene Dinge zusammen, die sicherlich dazu beitragen, dass Arbeitslose im Vergleich zu arbeitenden Personen weniger lebenszufrieden sind: Vielfältige Belastungen wie zum Beispiel die Angst vor Verlust des sozialen Status, das Gefühl, nichts wert zu sein, die eingeschränkten finanziellen Mittel und die seelische Belastung durch die unsichere Situation und durch Zukunftsängste sind sicherlich wichtige Randfaktoren, die zu der geringeren Lebenszufriedenheit beitragen.

[1] Vgl. Winkelmann und Winkelmann (1998).

Interessant ist in diesem Zusammenhang, dass das Einkommen keinen Einfluss auf die Lebenszufriedenheit hat (p=0.812). Ein höheres Einkommen führt also nicht automatisch zu einer höheren Lebenszufriedenheit, oder dem Volksmund folgend: Geld allein macht wohl doch nicht glücklich.[2] Keine Unterschiede zeigen sich auch beim Betrachten der Wohnregionen: weder beim West-Ost-Vergleich noch bei den Variablen zur Gemeindegröße zeigen sich signifikante Effekte.[3] Dieses Ergebnis ist zumindest hinsichtlich des nicht feststellbaren Unterschieds zwischen West- und Ostdeutschen interessant: Ist das Klischee vom 'Jammerossi' doch nur ein Klischee?

## 5.2 Zur Güte der Anpassung

Zum Abschluss der Analyse dieses Logit-Modells wurde nun noch die Güte der Anpassung untersucht. Dies geschieht mittels des Hosmer-Lemeshow Goodness-of-fit Test[4]. Die zugrunde liegende Hypothese lautet, dass das angegebene Modell ein passendes Modell für die zugrunde liegenden Daten darstellt. Hierfür werden die nach den durch das Modell hervorgesagten Wahrscheinlichkeiten geordneten Daten in etwa gleichgroße Gruppen aufgeteilt. Über diese Gruppen werden die Abweichungen zwischen den vorhergesagten Wahrscheinlichkeiten und den beobachteten Erfolgen berechnet. Der resultierende Wert ist $\chi^2$-verteilt.[5] Diese $\chi^2$-Statistik führt hier zu einem P-Wert von 0.6244, was besagt, dass die Nullhypothese, dass das zugrunde liegende Modell das angemessene Modell ist, nicht abgelehnt werden kann.
Dies rechtfertigt die Annahme, dass die Auswahl der erklärenden Variablen angemessen war.

---

[2] Vgl. Easterlin, Richard A. (1974): Does Economic Growth Improve the Human Lot? Some Empirical Evidence, in Paul A. David and Melvin W. Reder hg.): Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz, Academic Press,New York und London, Seite 89-125
Diener, Ed und Seligman, Martin E.P. (2004): Beyond Money. Toward an economy of Well-Being, American Psychological Society Journal, Volume 5 Number 1

[3] West: p=0.262, ländlich: p=0.241, Verdichtungsansatz: p=0.584

[4] Pearsons Goodness-of-fit Tests ist wegen der Aufnahme metrischer Variablen hier nicht geeignget
vgl. Hosmer, D.W. und Lemeshow, S. (1989): Applied Logistic Regression, Wiley, New York

[5] Vgl. Kohler, Ulrich und Kreuter, Frauke (2001): Datenanalyse mit Stata, Oldenbourg, München/Wien, Seite.276

# Kapitel 6

# Zusammenfassung und Fazit

## 6.1 Zusammenfassung der Ergebnisse im ALLBUS-Modell

Gibt es nun einen Zusammenhang zwischen der Perzeption der Wirtschaftslage und der Lebenszufriedenheit?
Anhand der vorliegenden Ergebnisse kann man dies vermuten. Es ergaben sich eindeutige signifikante Effekte der Einschätzung der Wirtschaftslage auf die persönliche Lebenszufriedenheit: Bei denjenigen Personen, die ihre persönliche wirtschaftliche Lage als 'gut' oder 'sehr gut' einschätzen, ist die Wahrscheinlichkeit, dass sie mit ihrem Leben zufrieden sind, höher als bei Personen, die diese als 'teils/teils', 'schlecht' oder 'sehr schlecht' einstufen. Wie schon eingangs erwähnt, bleibt hier natürlich die Frage unbeantwortet, ob sich dieser Zusammenhang eventuell auch in die andere Richtung finden lassen könnte, ob also eine höhere Lebenszufriedenheit zu einer positiveren Einschätzung der Wirtschaftslage führt. Dies an dieser Stelle zu erörtern, würde jedoch den Rahmen dieser Arbeit sprengen.

Der Einfluss der anderen in die Untersuchung miteingeschlossenen Variablen blieb weitaus geringer, als man vermuten würde: Lediglich drei der aufgenommenen Variablen zeigten signifikante Effekt hinsichtlich der Lebenszufriedenheit. Nur das Alter, der Familienstand und der Fakt, ob jemand arbeitslos ist oder nicht, zeigten signifikante Effekte. Naheliegende Einflüsse wie Unterschiede zwischen West- und Ostdeutschen oder die Höhe des Einkommens konnten in diesem Modell nicht

nachgewiesen werden[1]. Hier muss jedoch nochmals auf das Hauptproblem des Datensatzes hingewiesen werden, die sehr geringe Fallzahl: In die Untersuchung gingen nur n=420 Ostdeutsche mit ein, Westdeutsche hatten immerhin eine Fallzahl von n=908. Diese geringen Fallzahlen führen zu einer niedrigen Zellbelegung, die wiederum zu Verzerrungen führt. Dadurch bleiben eventuell signifikante Effekte unaufgedeckt.

## 6.2 Vergleich der Ergebnisse der beiden Untersuchungen

Ziel dieser Arbeit war unter anderem, herauszufinden, ob eine Analyse derselben Fragestellung anhand unterschiedlicher Datensätze vergleichbare Ergebnisse liefert. In diesem Fall wurde die Fragestellung nach dem Zusammenhang zwischen Einschätzung der Wirtschaftslage und Lebenszufriedenheit mittels Daten aus dem ALLBUS und des SOEP untersucht.
Schon in der Operationalisierung ergaben sich einige Probleme, da sich die Art der Fragestellung und die Skalierung in den beiden Datensätzen zum Teil erheblich voneinander unterscheiden. Mittels geeigneter Methoden wurden die Variablen so weit vergleichbar gemacht, dass es möglich erschien, nach den einzelnen Analysen einen Vergleich der Ergebnisse vornehmen zu können.
Die Response-Variable wurde mittels eines Mediansplits in eine dichotome Variable übergeführt, die abhängigen Variablen wurden so gewählt und kodiert, dass sie eine größtmögliche Ähnlichkeit aufwiesen. Unterschiede gibt es in der Variablen zur Einschätzung der Wirtschaftslage, die im Modell mit den ALLBUS-Daten als Einschätzung der persönlichen Wirtschaftslage verwendet wird, im SOEP-Modell als allgemeine Einschätzung. Man sollte trotz allem davon ausgehen, dass diese beiden Variablen in der Tendenz das Gleiche messen. Auch unterschiedlich erfasst wurde die Variable zum Vorhandensein von Kindern: Im SOEP-Modell wird mit dieser Variablen erfasst, ob Kinder unter 18 Jahren vorhanden sind, im ALLBUS-Modell misst diese Variable nur das Vorhandensein von Kindern im allgemeinen, egal ob drei oder 30 Jahre alt. Im SOEP-Modell geht zusätzlich zum Vorhandensein eines Hochschul-

[1] Testweise wurden noch Berechnungen mit Einkommensquintilen bzw. dem Einkommen als dichotome Variable statt des unklassierten Einkommens (mittels Mediansplit)durchgeführt: auch dies lieferte keine signifikanten Effekte.

abschlusses noch die Information über das Vorhandensein eines Berufsabschlusses ein: Diese Kategorie fehlt im ALLBUS-Modell. Die Variable zum Hochschulabschluss wurde außerdem noch unterschiedlich konstruiert.[2]
Die vorliegenden Untersuchungen zeigen, dass die Ergebnisse in der Tendenz in die gleiche Richtung weisen. Die Einschätzung der Wirtschaftslage hat in beiden Fällen signifikante Effekte auf die Lebenszufriedenheit, die sogar ähnlich hoch sind. Auch die Variable Alter wirkt sich in beiden Modellen signifikant auf die Lebenszufriedenheit aus. Auch hier sind die Richtung und die Höhe des Effektes gleich.
Bei der Arbeitslosigkeit zeigen sich sowohl im ALLBUS-Modell als auch im SOEP-Modell signifikante Effekte, die in die gleiche Richtung weisen. Bei der Untersuchung des Einflusses von Familienstand auf die Lebenszufriedenheit zeigen sich in beiden Modellen Effekte, die jedoch zum Teil unterschiedlich sind: In beiden Modellen sind getrennt lebende oder geschiedene Personen weniger lebenszufrieden als verheiratete. Im ALLBUS-Modell kann man denselben Effekt auch bei verwitweten Menschen beobachten; im SOEP-Modell zeigt sich dieser Effekt nicht. Dafür gibt es im SOEP-Modell signifikante Unterschiede hinsichtlich der Lebenszufriedenheit zwischen verheirateten und ledigen Menschen; dieser Effekt ist wiederum nicht im ALLBUS-Modell zu erkennen.
Hier ist jedoch zu erwähnen, dass sich alle Effekte, die in den Kategorien der Variablen zum Familienstand gefunden wurden, ob signifikant oder nicht, zumindest in der Richtung ähneln.

Die Gemeindegröße und das Vorhandensein von Kindern hat in beiden Modellen keinen Einfluss auf die Lebenszufriedenheit.

Unterschiede gibt es beim West-Ost-Vergleich, der Hochschulbildung und beim Einkommen: Hier konnten im Modell, das aus den SOEP-Daten errechnet wurde, signifikante Effekte gefunden werden, die sich im Modell mit Daten aus dem ALLBUS nicht zeigten. Effekte hatten im SOEP-Modell auch die Variablen, die in das ALLBUS-Modell nicht aufgenommen wurden: die Nationalität, der Berufsabschluss und der Status der Erwerbstätigkeit.
Auch hier ist zu bemerken, dass sich -mit Ausnahme der Variablen zur Hochschulbildung- die Effekte in ALLBUS und SOEP so weit gleichen, dass sie in die gleiche Richtung

[2] siehe Kapitel 3.3

laufen. Die Unterschiede bei der Variablen zur Hochschulbildung kann man durch die unterschiedliche Bauweise der Variablen in den Modellen erklären: Im ALLBUS-Modell hat das Vorhandensein eines Hochschulabschlusses einen negativen, nicht signifikanten Effekt auf die Lebenszufriedenheit, im SOEP-Modell findet man hier einen positiven signifikanten Effekt. Den Unterschied in der Richtung des Effektes kann man folgendermaßen erklären:
Im SOEP-Modell gehen in die eine Kategorie Personen mit Hochschulabschluss ein, in die Gegenkategorie solche ohne jeglichen Abschluss. Im ALLBUS-Modell finden sich in der Gegenkategorie all jene Personen, die keinen Hochschulabschluss haben, also auch solche mit Berufsabschluss.

Auffallend ist im Vergleich der beiden Modelle, dass das Modell mit der größeren Grundgesamtheit, also das´mit den Daten aus dem SOEP, weitaus mehr signifikante Effekte aufweist als das kleinere Modell. Wie schon häufiger erwähnt, liegt es also nahe, zu vermuten, dass einige Effekte, die das kleinere Modell nicht aufdeckt, deshalb nicht aufgedeckt werden, weil der Umfang des Datensatzes für solch eine komplexe Fragestellung etwas zu klein ist.
Dem Problem der unterschiedlichen Operationalisierung und Skalierung der einzelnen Variablen kann man mit geeigneten Rekodierungsinstrumenten beikommen. So kann die Vergleichbarkeit der Ergebnisse gewährleistet werden.

Grundsätzlich liefern die beiden Personenbefragungen recht ähnliche Ergebnisse. Deshalb ist die Frage, ob mit beiden Datensätzen eigentlich dasselbe gemessen wird, im Großen und Ganzen mit einem ‘ja’ zu beantworten.

Für eine Fragestellung mit relatitv vielen Einflussgrößen scheint jedoch der SOEP der geeignetere Datensatz zu sein, da die große Fallzahl viele der beim ALLBUS auftretenden Probleme gar nicht erst aufkommen lässt.

# Kapitel 7

# Anhang

Tabelle 7.1: Bivariate Analysen 1

| **Variable** | **Lebenszuf.=0** | **Lebenszuf.=1** | **gesamt** |
|---|---|---|---|
| Einschätzung der Wirtschaftslage: | | | |
| sehr gut/gut | 421 | 241 | 662 |
| teils/teils | 388 | 127 | 515 |
| schlecht/sehr schlecht | 127 | 24 | 151 |
| Total | 936 | 392 | 1328 |
| Pearson chi2(2) = | 34.4045 | Pr = 0.000 | |
| Geschlecht: | | | |
| Frau | 476 | 203 | 679 |
| Mann | 460 | 189 | 649 |
| Total | 936 | 392 | 1328 |
| Pearson chi2(1) = | 0.0958 | Pr = 0.757 | |
| Alter klassiert: | | | |
| 18-25 | 97 | 67 | 164 |
| 26-35 | 133 | 82 | 215 |
| 36-45 | 209 | 79 | 288 |
| 46-55 | 176 | 39 | 215 |
| 56-65 | 163 | 53 | 216 |
| älter als 65 | 158 | 72 | 1230 |
| Total | 935 | 392 | 1328 |
| Pearson chi2(5) = | 34.7218 | Pr = 0.000 | |
| Einkommensquintile: | | | |
| 1. Quintil | 151 | 56 | 207 |
| 2. Quintil | 158 | 50 | 208 |
| 3. Quintil | 156 | 57 | 213 |
| 4. Quintil | 142 | 66 | 208 |
| 5. Quintil | 151 | 74 | 225 |
| Total | 758 | 303 | 1061 |
| Pearson chi2(4) = | 5.7442 | Pr = 0.219 | |

Tabelle 7.2: Bivariate Analysen 2

| **Variable** | **Lebenszuf.=0** | **Lebenszuf.=1** | **gesamt** |
|---|---|---|---|
| Arbeitslosigkeit: | | | |
| Arbeit | 468 | 206 | 674 |
| keine Arbeit | 467 | 186 | 653 |
| Total | 935 | 392 | 1327 |
| Pearson chi2(1) = | 0.6893 | Pr = 0.406 | |
| Hochschulabschluss: | | | |
| kein Hochschulabschluss | 756 | 321 | 1077 |
| Hochschulabschluss | 180 | 71 | 251 |
| Total | 936 | 392 | 1328 |
| Pearson chi2(1) = | 0.2255 | Pr = 0.635 | |
| West/Ost: | | | |
| Ost | 311 | 109 | 420 |
| West | 625 | 283 | 908 |
| Total | 936 | 392 | 1328 |
| Pearson chi2(1) = | 3.7539 | Pr = 0.053 | |
| Kinder: | | | |
| keine Kinder | 269 | 122 | 391 |
| Kinder | 664 | 266 | 930 |
| Total | 933 | 388 | 1321 |
| Pearson chi2(1) = | 0.8969 | Pr = 0.344 | |
| Gemeindegröße: | | | |
| Klein | 131 | 61 | 192 |
| Mittel | 346 | 151 | 497 |
| Groß | 459 | 180 | 639 |
| Total | 936 | 392 | 1328 |
| Pearson chi2(2) = | 1.2057 | Pr = 0.547 | |
| Familienstand: | | | |
| Verheiratet zusammenlebend | 536 | 261 | 797 |
| geschieden/getrennt lebend | 93 | 15 | 108 |
| Witwer | 75 | 14 | 89 |
| ledig | 232 | 101 | 333 |
| Total | 936 | 391 | 1327 |
| Pearson chi2(3) = | 24.9385 | Pr = 0.000 | |

# Literaturverzeichnis


[1] Blanchflower, D.G. and Oswald, A.J.: *The Rising Well-Being of the Young* paper presented at an NBER Conference on Disadvantaged Youth, North Carolina, December 1996

[2] David, Paul A. und Reder, Melvin W.(Hg.): *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz.*
Academic Press, New York und London 1974

[3] Diener, Ed und Suh, Eunkokk M.(Hg.): *Culture and Subjective Well-being*
The MIT Press, London 2003

[4] Diener, Ed und Seligman, Martin E.P: *Beyond Money. Toward an Economy of Well-Being*
American Psychological Society Journal, Volume 5 Number 1 2004

[5] Di Tella, Rafael, MacCulloch, Robert J. und Andrew J. Oswald: *How Do Macroeconomic Fluctuations Affect Happiness?*
Mimeo, Harvard Business School, 1999

[6] Easterlin, Richard A.(Hg.): *Happiness in Economics*
Edward Elgar Publishing, Cheltenham 2002

[7] Fernandez, R.M. und Kulik, J.C.: *A multilevel model of life satisfaction: Effects of individual characteristics and neighborhood composition*
American Sociological Review, 46 (6:), Seite 840- 850, 1981

[8] Frey, Bernd S. und Stutzer, Alois: *Happiness and Economics. How the Economy and Institutions affect Human Well-Being*
Princeton University Press, Princeton und Oxford 2001

[9] Glatzer, Wolfgang und Zapf, Wolfgang (Hg.): *Lebensqualität in der Bundesrepublik. Objektive Lebensbedingungen und sujektives Wohlbefinden*
Campus Verlag, Frankfurt/New York 1984

[10] Glatzer, Wolfgang, Habich, Roland und Klarl-Ulrich Maier(Hg.): *Sozialer Wandel und Gesellschaftliche Dauerbeobachtung. Festschrift für Wolfgang Zapf*
Leske+Bundrich, Opladen Deutschland 2002

[11] Greene,William H.: *Econometric Analysis*
Prentice-Hall, New Jersey 2003

[12] Horowitz, Joel L. und Savin, N. Eugene: *Binary response models: Logits, Probits and Semiparametrics*
Journal of Economic Perspektives 15(4), 2001

[13] Hosmer, D.W. und Lemeshow, S.: *Applied Logistic Regression*
Wiley, New York 1998

[14] Kohler, Ulrich und Kreuter, Frauke : *Datenanalyse mit STATA*
Oldenbourg, München/Wien 2001

[15] Nussbaum, Martha und Sen, Amatya(Hg.): *The Quality of life*
Clarendon Press, Oxford UK 1993

[16] Oswald, A.J.: *Happiness and exonomic performance*
The Economic Journal 107, Seite 1815-1831

[17] Robinson, J.P. und Shaver, P.R. (Hg.): *Measures of social psychological attitudes*
Ann Arbor: Institute for Social Research, University of Michigan 1973

[18] Rönz, Bernd : *Generalisierte lineare Modelle*
Berlin, 2001

[19] Ronning, Gerd: *Mikroökonometrie*
Springer, Berlin 1991

[20] Strack, Fritz, Argyle, Michael und Norbert Schwarz: *Subjective Well-being: An interdisciplinary perspecitve*
Pergamon Press, New York USA 1991

[21] Winkelmann,L. and R. Winkelmann: *Why Are the Unemployed So Unhappy? Evidence from Panel Data*
Economica, 65 (257), 1-15, 1998