

Aus dem Institut für Agrarökonomie  
der Christian-Albrechts-Universität zu Kiel

# METHODISCHE ANWENDUNGEN AUF AGRARMÄRKTEN

Dissertation  
zur Erlangung des Doktorgrades  
der Agrar- und Ernährungswissenschaftlichen Fakultät  
der Christian-Albrechts-Universität zu Kiel

vorgelegt von  
Dipl. Vw. Meike Wocken, M.Sc.  
aus Bielefeld  
Kiel, 2012

---

Dekan: Prof. Dr. Rainer Horn

1. Berichterstatter: Prof. Dr. Jens-Peter Loy

2. Berichterstatter: Prof. Dr. Martin Schellhorn

Tag der mündlichen Prüfung: 30. Januar 2013

Gedruckt mit Genehmigung der  
Agrar- und Ernährungswissenschaftlichen Fakultät  
der Christian-Albrechts-Universität zu Kiel

Diese Dissertation kann als elektronisches Medium  
über den Internetauftritt der Universitätsbibliothek Kiel  
([www.ub.uni-kiel.de](http://www.ub.uni-kiel.de); [eldiss.uni-kiel.de](http://eldiss.uni-kiel.de))  
aus dem Internet geladen werden.

# Inhaltsverzeichnis

Inhaltsverzeichnis . . . . .	I
<b>1 Einleitung und Zusammenfassung</b>	<b>1</b>
<b>2 Validation study of consumer scanner and retail scanner data</b>	<b>13</b>
2.1 Introduction . . . . .	14
2.2 Measurement error and its impact . . . . .	16
2.3 Data and data preparation . . . . .	18
2.4 Examine measurement error . . . . .	22
2.4.1 Additive structure of measurement error . . . . .	25
2.4.2 Correlation of true price and the measurement error . . . . .	26
2.4.3 Quantifying biases . . . . .	28
2.5 Discussion of results . . . . .	31
2.6 Conclusion . . . . .	35
<b>3 Tobit regression to estimate impact of EU market intervention in the German butter market</b>	<b>39</b>
3.1 Introduction . . . . .	40
3.2 German Butter Market . . . . .	41
3.3 Tobit Model . . . . .	43
3.4 Results from Maximum Likelihood Estimation . . . . .	47
3.5 Prediction of market development . . . . .	51
3.6 Test for distributional assumption . . . . .	54
3.7 Conclusions . . . . .	56
<b>4 Beurteilung der Kompensation auf dem EU-Milchmarkt auf Basis eines Tobit-Ansatzes</b>	<b>63</b>
4.1 Einleitung . . . . .	64
4.2 Entwicklung Butter- und Magermilchpulvermarkt . . . . .	65
4.3 Berechnung Rohstoffwert und Kompensationsgrad . . . . .	67

4.4	Tobit Modellierung für Butter und Magermilchpulver . . . . .	68
4.5	Ergebnisse Tobit Modell . . . . .	71
4.6	Prognose Exporterstattung . . . . .	73
4.7	Bestimmung der Kompensation . . . . .	74
4.8	Diskussion der Ergebnisse und Fazit . . . . .	76
<b>5</b>	<b>Einsatz eines Prozessmanagementzyklus zur Optimierung von eLearning:</b>	
	<b>Praxisbeispiel „eLearning QMM“</b>	<b>81</b>
5.1	Einleitung . . . . .	82
5.2	Handlungsschritte Prozessmanagement im eLearning . . . . .	82
5.3	Praxisbeispiel „eLearning QMM“ . . . . .	84
5.4	Nutzen, Aufwand und weitere Informationen . . . . .	89
<b>6</b>	<b>Evaluation of eLearning - A study of Undergraduate Agricultural Economics</b>	
	<b>course</b>	<b>97</b>
6.1	Introduction . . . . .	98
6.2	Data and Methods . . . . .	99
6.2.1	Course description . . . . .	99
6.2.2	Data description . . . . .	100
6.2.3	Methods . . . . .	102
6.3	Results . . . . .	103
6.4	Conclusion . . . . .	107
<b>7</b>	<b>Quantitative Untersuchung zur Effizienz von eLearning</b>	<b>115</b>
7.1	Einleitung . . . . .	116
7.2	Daten und Hypothesen . . . . .	117
7.3	Methoden . . . . .	119
7.4	Ergebnisse . . . . .	120
7.5	Fazit . . . . .	121
<b>8</b>	<b>Tobit Modell</b>	<b>129</b>
8.1	Modellformulierung . . . . .	130
8.2	Interpretation des Modells . . . . .	132
8.3	Probleme der OLS Schätzung bei zensierten Daten . . . . .	135
8.4	Schätzverfahren . . . . .	139
8.4.1	Maximum Likelihood . . . . .	139
8.4.2	Heckmann 2-Schritt Verfahren . . . . .	141
8.4.3	Censored Least Absolute Deviations . . . . .	142

8.5	Fehlspezifikationen des Modells . . . . .	143
8.6	Likelihood Ratio Test . . . . .	145
<b>9</b>	<b>Schlussbetrachtung</b>	<b>149</b>
<b>10</b>	<b>Summary</b>	<b>159</b>



# Kapitel 1

## Einleitung und Zusammenfassung

Die zweite Jahrestagung der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaus (GeWiSoLa) 1961 in Hohenheim stellt einen wichtigen Meilenstein in der Entwicklung der agrarökonomischen Forschung in der Bundesrepublik Deutschland dar. (Reisch, 1967a, S. 9) Die Tagung zu dem Thema „Bedeutung und Anwendung ökonometrischer Methoden in der agrarwirtschaftlichen Forschung“ befasste sich mit den Möglichkeiten und Grenzen der ökonometrischen Methoden. (Hanau, 1963, S. 492) Das Ergebnis der Tagung ist die grundsätzliche Befürwortung und die Etablierung der ökonometrischen Methoden zu einem festen Bestandteil der Agrarökonomie. (Reisch, 1967a, S. 9)

Die Ökonometrie bietet Methoden zur empirischen Überprüfung der Theorien der (Agrar-)Ökonomie. (Petrick, 2004, S. 8) Gemäß dem heute üblichen deduktiven Vorgehen wird gefordert, dass theoretische Überlegungen mit empirischen Beobachtungen in Einklang stehen. Ergibt sich mittels empirischen Tests kein Widerspruch zu aufgestellten Hypothesen, so wird die Theorie vorläufig unter den gegebenen Bedingungen als gültig angesehen. Stimmt die Empirie nicht mit der Theorie überein, so ist die Theorie falsifiziert. Die Prämissen sind dann zu überarbeiten und weiterzuentwickeln. (Koester, 1992, S. 5ff.) Es ist daraus zu folgern, „*daß jeder wissenschaftliche Fortschritt das Ergebnis theoretischer und empirischer Forschung darstellt.*“ (Koester, 1992, S. 7)

Die Ökonometrie liefert damit große Beiträge zur Weiterentwicklung der theoretischen Erklärung der Realität. Doch führt sie nicht dazu, dass die Agrarökonomie eine harte Wissenschaft wird. (Petrick, 2004, S. 10f.) Empirische Beobachtungen sind raum- und zeitgebunden. Daher ist eine nicht falsifizierte Theorie kein allgemeingültiges Gesetz. (Koester, 1992, S. 6) Weitere kritische Punkte sind unter anderem, dass Modellspezifikationen bei ökonometrischen Analysen als vollständig und korrekt angenommen werden, dass statistische Signifikanzniveaus zur Bestätigung signifikanter Zusammen-

## *Einleitung und Zusammenfassung*

hänge auf allgemeiner Konvention beruhen, dass die Beschaffung relevanter Daten zum Teil nicht möglich ist, oder dass die Daten zum Teil fehlerbehaftet sind. (Petrick, 2004, S. 10f.) Trotz der Limitierung sind ökonometrische Modelle ein mächtiges Instrument. Die Grenzen einer empirischen Untersuchung sind dabei jeweils klar zu kommunizieren. (Petrick, 2004, S. 14)

Seit der zweiten GeWiSoLa Jahrestagung 1961 bestehen Bestrebungen in der Agrarökonomie, die quantitativen Methoden weiterzuentwickeln. Bereits auf der fünften GeWiSoLa Jahrestagung 1966 konnte eine Vielzahl an Erweiterungen präsentiert werden. (Reisch, 1967b, S. 455f.) Im Laufe der Zeit haben AgrarökonomInnen zentrale Entwicklungen in der angewandten Ökonometrie zur Analyse der Nachfrage, des Angebots und der Preisbildung auf Agrarmärkten bestimmt. Weiterentwicklungen der empirischen Untersuchungen von unvollkommenen Märkten der Agrar- und Ernährungswirtschaft, der politischen Ökonomie der Agrarprotektion und allgemeiner Gleichgewichtsmodelle kommen ebenso aus der Agrarökonomie. (Herrmann, 2005, S. 297f.) Dieser Methodenbeitrag zur Ökonomie „*ist offenbar für das Selbstverständnis der Agrarökonomie sehr wichtig.*“ (Herrmann, 2005, S. 297)

Ein weiterer wichtiger Einsatzbereich quantitativer Methoden ist die Politikevaluierung. Eine ex post Analyse von Politikmaßnahmen zur Beurteilung der Wirkung ist notwendig, um damit eine weitergehende effiziente (Agrar-)Politikplanung zu erreichen und zu kontrollieren. Die empirische Politikevaluierung spielt daher in der EU und in weiteren internationalen Institutionen eine immer bedeutendere Rolle, einhergehend mit steigenden Ansprüchen an die methodischen Ansätze. (Henning und Michalek, 2008, S. 233) Die Weiterentwicklung, Erweiterung und Anwendung der quantitativen Methoden ist daher auch in der Zukunft eine wichtige Aufgabe innerhalb der Agrarökonomie.

Die vorliegende, kumulative Arbeit beinhaltet sechs eigenständige Beiträge. Die Beiträge in den Kapiteln 2 bis 4 sind quantitative Untersuchungen auf Agrar- und Ernährungsmärkten. Der Beitrag in Kapitel 2 fokussiert die Datenqualität des deutschen Haushaltspanels *ConsumerScan* der GfK Panel Services Deutschland. Die Preisangaben werden hinsichtlich Abweichungen zu Preisangaben aus dem Lebensmitteleinzelhandel untersucht. Gegenüber einer vorhergehenden Untersuchung des Haushaltspanel von Nielsen durch Einav, Leibtag und Nevo (2010) wird dabei die eigentliche, vom Panelteilnehmer eingetragene Preisangabe berücksichtigt, anstatt einer durch den Panelanbieter eingefügten Preisvariable. Zusätzlich wird die Auswirkung fehlerhafter Preisangaben in linearen Regressionsmodellen beleuchtet.



In dem Beitrag in Kapitel 3 wird der Einfluss der EU-Intervention auf Höhe und Volatilität der Marktpreise am Beispiel des deutschen Buttermarktes untersucht. Für die Untersuchung wird ein heteroskedastisches Tobit Modell verwendet, welches, im Gegensatz zu einer ähnlichen Untersuchung des amerikanischen Buttermarktes (Chavas und Kim, 2006), auch die Entwicklung des Weltmarktpreises berücksichtigt. Diese Modellspezifikation findet zusätzlich in dem Beitrag in Kapitel 4 Anwendung. Dort wird das Modell jeweils für den deutschen Butter- und Magermilchpulvermarkt geschätzt. Mithilfe der Schätzergebnisse werden die Marktpreisentwicklungen ohne Reform der Interventionspolitik prognostiziert. Die Prognosen werden genutzt, um daraus den Preisrückgang durch die Interventionspreissenkung zu ermitteln. Damit wird anschließend der Grad der Kompensation durch die Zahlung der Milchprämie berechnet. Im Gegensatz zu einer vorhergehenden Studie von Loy und Steinhagen (2009), berücksichtigt die hier verwendete Methode bei der Prognose neben der Weltmarktpreisentwicklung zusätzlich die stabilisierende Wirkung der Interventionspolitik auf dem inländischen Markt.

Wie bereits einleitend erwähnt, hat die Anwendung quantitativer Methoden Grenzen und Schwachpunkte. Zur Unterbindung einer „*Zahlen- und Methodengläubigkeit*“ ist eine sorgfältige Ausbildung und die Vermittlung einer kritischen Grundeinstellung notwendig. Dies ist vor allem eine Aufgabe der Hochschule. (Hanau, 1963, S. 492; Reisch, 1967b, S. 455) Die Beiträge in den Kapitel 5 bis 7 beschäftigen sich mit der Vermittlung quantitativer Methoden im Studium der Agrarwissenschaften. Dabei steht die Einführung von „Electronic Learning“, kurz eLearning, im Mittelpunkt. eLearning ist eine Lehrform, die durch neue Medien unterstützt und ermöglicht wird. Die Abgrenzung des Begriffs eLearning birgt Schwierigkeiten, da es nicht das eine eLearning gibt. Im Laufe der Zeit hat sich das Verständnis des Begriffes von einem stark technisch orientierten zu einem vermehrt didaktischen gewandelt. Die Wandlung geht einher mit vielfältigen Formen von eLearning in der Lehre, welche zwischen rein virtueller Lehre und Präsenzlehre variieren. Dabei ist ein Link allein zu online verfügbaren Dokumenten kein eLearning, es kommt auf die Aufbereitung des Materials an. (Revermann, 2006, S. 25)

Für das erfolgreiche Weiterbestehen der Agrarökonomie, gerade im Hinblick auf die Mittelvergabe gemäß der Studentenzahl, ist zusätzlich zu dem Forschungserfolg wichtig, für Studenten attraktiv zu sein. (Herrmann, 2005, S. 299) eLearning bietet in Zeiten steigender Studentenzahlen und sinkender Budgets Lösungsmöglichkeiten in der Studentenbetreuung. Der sinnvolle Einsatz neuer Medien kann zusätzlich die Attraktivität eines Studiengangs steigern. Die Entwicklung der quantitativen Methoden geht einher mit einer Zunahme an statistischen und ökonometrischen Softwarepaketen, die

## *Einleitung und Zusammenfassung*

die Rechenarbeit durch den Computer übernehmen. Die Vermittlung der Theorie und insbesondere der Anwendung am Computer kann durch die Vermeidung von Medienbrüchen profitieren und die Anwenderfreundlichkeit für die Studenten steigern. Doch den potenziellen Vorzügen stehen auch Nachteile gegenüber. Zum einen können Kognitions- und Kommunikationsmechanismen durch die Abwesenheit von Lehrenden und Kommilitonen in Mitleidenschaft gezogen werden. (Revermann, 2006, S. 6) Zum anderen können eLearning-Szenarien aber auch vermehrt Kosten verursachen, die den Nutzen weit übersteigen. (Schwarz, 2001, S. 5; Twigg, 2003, S. 28)

Im Rahmen eines universitätsweiten Projektes mit dem Ziel der nachhaltigen Integration von eLearning in der Präsenzlehre wurde an der Agrarfakultät der Christian-Albrechts-Universität zu Kiel zur Unterstützung einer Vorlesung, welche die ökonometrische Analyse von Daten zum Inhalt hat, zum ersten Mal die Open-Source Lernplattform OLAT eingesetzt. Der Beitrag in Kapitel 5 der vorliegenden Arbeit stellt anhand dieses Projektes einen Prozessmanagementzyklus zur effizienten Umsetzung von eLearning vor. Es werden dabei Aufwand und Nutzen eines solchen Vorgehens diskutiert. Die Evaluationsphase des Prozessmanagements im ersten Zyklusdurchlauf ist Inhalt des Beitrags in Kapitel 6. Hauptaugenmerk liegt dabei auf der Entwicklung und Nutzung eines Konstrukts zur Messung des Erfolgs von eLearning im Rahmen der Lernerfahrung der Studenten. Der Beitrag in Kapitel 7 betrachtet die Effizienz von eLearning und untersucht den Effekt des eLearning-Einsatzes auf den Betreuungsaufwand.

In Kapitel 8 wird die Theorie der Tobit Modelle umfassend dargestellt. Neben der Modellformulierung werden unter anderem auch Verfahren zur Schätzung erläutert. Die Tobit Modelle in Kapitel 3 und 4 werden mittels Maximum Likelihood Methode geschätzt. Daher wird in Kapitel 8 der auf der Maximum Likelihood Methode basierende und in Kapitel 3 und 4 genutzte Likelihood Ratio Test zur Überprüfung von Hypothesen erklärt.

Die einzelnen Beiträge der kumulativen Dissertation werden im Folgenden kurz zusammengefasst. Es werden dabei die Ziele der jeweiligen Untersuchungen, die Methoden und die wichtigsten Ergebnisse dargestellt. Eine kritische Würdigung und Ansätze für Erweiterungen der Untersuchungen werden in der abschließende Betrachtung der Beiträge in Kapitel 9 aufgegriffen und diskutiert.

## Validation study of consumer scanner and retail scanner data

Studien, die Daten von Haushaltspanels als Grundlage benutzen, um u.a. das Kaufverhalten von Konsumenten zu analysieren, beruhen meistens auf der Annahme fehlerfreier, unverzerrter Daten. Kritische Betrachtungen zeigen aber, dass diese Annahme nicht immer gerechtfertigt ist. Das erste Ziel der Untersuchung ist daher, die Preisangaben der Panelteilnehmer des *Consumer Scan* Haushaltspanels der GfK Panel Services Deutschland hinsichtlich fehlerhafter Angaben zu untersuchen. Das zweite Ziel ist die Klassifizierung des möglichen Messfehlers, sowie die Quantifizierung resultierender Verzerrungen in linearen Regressionsanalysen durch die Benutzung der möglicherweise fehlerbehafteten Preisbeobachtungen.

Die benutzten Daten werden im Beitrag auch als Konsumenten Scannerdaten bezeichnet, da die teilnehmenden Haushalte mittels Handscanner jeden gekauften Artikel einscannen und den gezahlten Preis, sowie die Handelskette, zu der die aufgesuchte Einkaufsstätte gehört, hinzufügen. Zur Validierung der Preisangaben wird ein Scannerdatensatz aus dem Lebensmitteleinzelhandel (LEH) verwendet. Dieser Datensatz enthält Beobachtungen über die einzelnen Verkäufe auf der Ebene der individuellen Einkaufsstätten. Aufgrund der automatischen Erfassung der Beobachtungen an den Kassen während des Bezahlvorgangs wird für diesen Datensatz Fehlerfreiheit unterstellt. Ein direktes Zusammenfügen der beiden Datensätze ist wegen fehlender detaillierter Angabe der Einkaufsstätte im Haushaltspanel nicht möglich. Daher werden nur Artikel in die Untersuchung einbezogen, für die im LEH Datensatz eine einheitliche Preispolitik in allen Geschäften einer Kette vorliegt. Dieser einheitliche Preis eines Artikels in einer Kalenderwoche wird als der wahre Preis für die jeweilige Woche und Kette unterstellt. Damit ist die Information über die besuchte Einkaufsstätte im Detail vernachlässigbar.

44,87% der untersuchten Preisangaben im Haushaltspanel sind fehlerhaft. Die Annahme des additiven, normalverteilten Messfehlers kann für den gesamten Datensatz nicht abgelehnt werden. Der Fehler in den Beobachtungen ist korreliert mit dem Niveau des wahren Preises, d.h. dem einheitlichen Preis innerhalb der Kette. Aufgrund der Korrelation muss die Annahme des klassischen Messfehlers verworfen werden. Zusätzlich sind statistisch signifikante Effekte des Alters und des Einkommens der Haushalte auf den Messfehler in den Preisangaben zu beobachten. Die berechnete Reliability Ratio ist statistisch signifikant kleiner 1 für jede Produktkategorie, mit Ausnahme von H-Milch. Daraus folgt, dass die Schätzer in linearen Regressionsmodellen verzerrt sind, wenn der fehlerbehaftete Preis als exogene Variable in das Modell eingeht. Die Schätzer sind auch

## *Einleitung und Zusammenfassung*

verzerrt, wenn der Preis mit Messfehler u.a. auf Alter und Einkommen regressiert wird. Durch den nicht-klassischen Messfehler ist die Struktur der resultierenden Verzerrung in beiden Fällen komplex.

Durch die restriktive Auswahl der untersuchten Artikel, sind vermehrt Artikel mit geringen Preisänderungen in die Untersuchung aufgenommen worden. Daher ist der entdeckte Fehler in den Daten als untere Grenze des Messfehlers aufzufassen. Im Allgemeinen ist zu erwarten, dass Messfehler, aber auch resultierende Verzerrungen, stärker auftreten.

### **Tobit regression to estimate impact of EU market intervention in the German butter market**

Mit der Agenda 2000 und der Halbzeitbewertung 2003 begann die Liberalisierung des Milchmarktes im Rahmen der Reform der gemeinsamen Agrarpolitik. Dazu gehörte die Absenkung der Interventionspreise auf dem Butter- und Magermilchpulvermarkt. Der Beitrag untersucht die Effekte der europäischen Interventionspolitik auf dem europäischen Buttermarkt am Beispiel Deutschlands. Ein Verständnis der Effekte ist notwendig, für einen effizienten Gebrauch der reformierten Preisstützung zur möglichen Eindämmung von Preisvolatilitäten in der Zukunft. Um den Markt für Butter analysieren zu können, wird ein Modell benötigt, welches die Zensierung der abhängigen Variablen durch den politisch festgelegten Interventionspreis berücksichtigt. Das Tobit Modell berücksichtigt die Linkszensierung durch den Interventionspreis. Ein heteroskedastisches Tobit Modell wird mit den monatlichen Beobachtungen des deutschen Buttermarktes für den Zeitraum 1979-2012 geschätzt. Die Spezifikation des Tobit Modells beruht auf einem Fehlerkorrekturmodell und berücksichtigt die Nicht-Stationariät und Kointegration der inländischen Preisreihe und des Exportpreises (Weltmarktpreis plus EU Exportsubvention). Das heteroskedastische Tobit Modell wird mittels Maximum Likelihood Methode geschätzt.

Die Ergebnisse deuten darauf hin, dass die Preisstützung zur Verringerung von Instabilitäten auf dem deutschen Buttermarkt beiträgt. Mit dem geschätzten Modell wird die Marktentwicklung ohne Absenkung der Interventionspreise prognostiziert. Die Ergebnisse zeigen einen positiven Effekt hinsichtlich des erwarteten Preisniveaus, obwohl der durchschnittlich erwartete Marktpreis oberhalb des (nicht gesenkten) Interventionspreises liegt und daher keine Marktintervention stattfindet. Das bedeutet, dass auch ohne Aufkäufe eine stützende Wirkung auf dem Markt durch die Interventionspolitik festzustellen ist.

## **Beurteilung der Kompensation auf dem EU-Milchmarkt auf Basis eines Tobit-Ansatzes**

Vor dem Hintergrund der Marktliberalisierung geht dieser Beitrag der Frage nach, ob die Zahlung der sogenannten Milchprämie ausreichend ist, um den Rückgang des Milchauszahlungspreises aufgrund der Reduktion der Interventionspreise auf den Märkten für Butter und Magermilchpulver zu kompensieren. Die Frage wird unter Anwendung des bereits im Beitrag *Tobit regression to estimate impact of EU market intervention in the German butter market* entwickelten Tobit Modells beantwortet. Mithilfe des Modells wird die Auswirkung der Senkung der Preisstütze auf den Märkten für Butter und Magermilchpulver abgeschätzt. Es werden die Marktentwicklungen ohne Reform prognostiziert. Aus den Prognosen heraus wird ein Rohstoffwert zur Approximation des Milchauszahlungspreises berechnet. Der Rohstoffwert dient als Referenzpreis zur Abschätzung des Kompensationsgrads.

Der als Referenzpreis prognostizierte Rohstoffwert für den deutschen Markt weist durch die Wirkung der Intervention gedämpfte Preisausschläge auf. Es zeigt sich die stabilisierende Wirkung der Interventionspolitik. Daher ist zu vermuten, dass im Fall stark abfallender Weltmarktpreise das deutsche Preisniveau auf dem Milchmarkt durch Interventionspolitik stabil gehalten werden könnte. Im gesamten Zeitraum 2004 bis 2008 liegt keine Überkompensation vor. Der Grad der Kompensation in den einzelnen Jahren variiert mit der Entwicklung des Weltmarktes. Um die Senkung der Interventionspreise vollständig zu kompensieren, müsste die Milchprämie um 50% steigen.

## **Einsatz eines Prozessmanagementzyklus zur Optimierung von eLearning: Praxisbeispiel „eLearning QMM“**

Zur Schaffung einer zielgerichteten eLearning-Umgebung für eine Lehrveranstaltung, muss der eLearning-Einsatz geplant, umgesetzt, evaluiert und verbessert werden. Es werden dabei die unterschiedlichen Inputs, u.a. die Bedürfnisse der Studenten und Dozenten, der zu vermittelnde Lehrstoff, Materialien, Hard- und Software, in den Output eLearning-Umgebung transformiert. Dieser wiederkehrende Prozess der Transformation ist durch einen Prozessmanagementzyklus zu unterstützen. In dem Beitrag wird dafür ein an den eLearning-Prozess von Čukušić u. a. (2010) angelehntes Prozessmanagement vorgeschlagen, welches geeignet ist, in mehreren Zyklusdurchläufen nacheinander angewendet zu werden. Der praktische Einsatz dieser Managementmethode wird anhand des Praxisprojektes „eLearning QMM“ skizziert. Das Projekt „eLearning QMM“ wurde im

## *Einleitung und Zusammenfassung*

Rahmen der Förderung der nachhaltigen Integration von eLearning in der Präsenzlehre an der Christian-Albrechts-Universität zu Kiel in den Sommersemestern 2010 und 2011 durchgeführt. Dabei wurde eLearning zur Unterstützung des Moduls „Quantitative Methoden der Marktanalyse“ (QMM) eingeführt. Inhalt des Moduls QMM ist die Theorie und Anwendung von statistischen und ökonometrischen Methoden.

Insgesamt hat das Modul QMM durch den Einsatz des Prozessmanagements profitiert. Die Wiederverwendbarkeit der eLearning-Inhalte gewinnt durch die zyklische Betrachtung. Es findet eine Standardisierung der sich wiederholenden Zyklusdurchläufe statt, was die Effizienz der Durchführung steigert. Wiederkehrende Fehler werden vermieden. Durch die Transparenz der Planung und Umsetzung resultiert ein Kommunikationsvorteil, der Kooperationen, z.B. mit dem technischen Support, vereinfacht. Die Durchführbarkeit des Prozessmanagements ist dadurch beschränkt, dass nicht immer alle Ziele quantifizierbar und mithilfe kritischer Erfolgsfaktoren quantitativ kontrollierbar sind. Auch bestehen an Hochschulen Schwierigkeiten, die Budgetierung hinsichtlich Effizienz im Blick zu behalten, da an Hochschulen häufig Kostenzuweisungen im Rahmen zentraler Einrichtungen, z.B. technischem Support, unzureichend sind.

### **Evaluation of eLearning - A study of Undergraduate Agricultural Economics course**

Integration von eLearning in die Lehre gewinnt immer mehr an Bedeutung. Im Rahmen des Projektes „eLearning QMM“ wurde zur Unterstützung einer Vorlesung, welche die ökonometrische Analyse von Daten zum Inhalt hat, im Sommersemester 2010 zum ersten Mal die Open-Source Lernplattform OLAT eingesetzt. Die Lernplattform bietet unterschiedliche Instrumente, um didaktische Lehr- und Lernformen online zu unterstützen. Um den Einsatz der Lernplattform zu evaluieren, wurde eine schriftliche Befragung der an der Vorlesung teilnehmenden Studenten am Ende des Semesters durchgeführt. Das Ergebnis der Untersuchung dient als Grundlage der Optimierung des weiteren Einsatzes der Inhalte der Lernplattform. Da eine hohe Punktzahl in einem Test/Abschlussprüfung nicht zwangsläufig mit einer hohen Zufriedenheit mit der Lernumgebung einhergehen muss, wird ein Konstrukt zur Quantifizierung der Lernerfahrung entwickelt und im Rahmen der schriftlichen Befragung in der vorliegenden Untersuchung abgefragt. Zusätzlich werden Fragen zur Charakterisierung der Studenten gestellt. Außerdem ist von den Studenten ein Multiple Choice Test über die Inhalte der Vorlesung auszufüllen.

Neben einer Analyse der Nutzungshäufigkeiten werden die Daten mittels linearer Regression und Hauptkomponentenanalyse untersucht. Die Anpassung eines linearen Modells

zeigt einen positiven statistisch signifikanten Zusammenhang zwischen der angegebenen Häufigkeit der Nutzung der Lernplattform und der erreichten Punktzahl im Multiple Choice Test. Der positive Zusammenhang lässt auf einen erfolgreichen Einsatz der Lernplattform schließen, muss aber kritisch aufgrund der weiteren Ergebnisse gesehen werden. Es fand vermehrt eine passiv-konsumierende Nutzung der Inhalte statt, anstatt einer aktiven kognitiven Auseinandersetzung. Auch sind die didaktischen Funktionen Motivation zur Auseinandersetzung mit dem Stoffgebiet, verbesserte Einschätzung des eigenen Lernerfolgs und Organisation der Inhalte durch Vernetzung von Vorlesung, Lernplattform und Übung nicht wie gewünscht erfüllt worden. Die Hauptkomponentenanalyse bestätigt, dass eine hohe Punktzahl in dem Multiple Choice Test nicht mit einer hohen Zufriedenheit mit dem eLearning einhergeht, sondern mit kontinuierlichem Lernen und der extrinsischen Motivation, eine gute Note erzielen zu wollen. Eine hohe Zufriedenheit mit dem eLearning Angebot ist hingegen zum einen bei Studenten mit ausgeprägtem Interesse an der Ökonometrie zu finden, zum anderen bei Studenten, die sich als passive Lerntypen einschätzen.

Da das eLearning sich auch immer an den Anforderungen der Zielgruppe orientieren soll, ist aufgrund des Ergebnisses, dass sowohl ökonomisch interessierte Studenten als auch passiv arbeitende Studenten zufrieden mit dem Angebot der Lernplattform sind, die Zielsetzung des eLearning Einsatzes zu überdenken. Neben der Unterstützung eines aktiven Lernverhaltens, sollte es - gerade bei einer Pflichtvorlesung - auch ein Ziel sein, die Studenten, die den Inhalten der Veranstaltung passiv gegenüberstehen, zufrieden zustellen. Daher werden die eLearning Inhalte zum einen durch Elemente zur Aktivierung des Lernverhaltens optimiert. Zum anderen wird zur Unterstützung passiver Studenten ohne ökonomisches Interesse die Wissensorganisation und -strukturierung auf der Lernplattform ausgebaut.

### **Quantitative Untersuchung zur Effizienz von eLearning**

Die Effizienz und Effektivität von eLearning in der Hochschullehre wird häufig nicht in Frage gestellt. Es wird aber in einigen Veröffentlichungen die Vermutung geäußert, dass vermehrt Studenten einen Nutzen durch den Einsatz von eLearning haben, aber nicht Lehrende. Der Beitrag untersucht die Auswirkung des eLearning Einsatzes auf den Betreuungsaufwand am Beispiel der Betreuung von studentischen Projektarbeiten über ein online Learn Content Management System (LCMS, auch als online Lernplattform bezeichnet).

Die individuelle Betreuung, welche zum größten Teil über elektronische Nachrichten statt-

## *Einleitung und Zusammenfassung*

findet, ist ein Haupttreiber des Betreuungsaufwands. Zur Abschätzung des Aufwands wird daher approximativ die Anzahl der vom Betreuer geschriebenen Nachrichten je Student herangezogen. Das eLearning Szenario wird mit einem Szenario ohne eLearning Einsatz verglichen. Zur Feststellung allgemeiner Gruppenunterschiede werden Wilcoxon-Rangsummentests durchgeführt. Zusätzlich wird der Effekt der Anzahl der geschriebenen Nachrichten auf die Note mittels Logit-Modell untersucht. Dafür werden die erreichten Noten in der Projektarbeit durch den Median in gute und schlechte Leistungen eingeteilt.

Weder für die erreichten Noten, noch für die Anzahl der Nachrichten werden statistisch signifikante Gruppenunterschiede festgestellt. Es ist auch kein statistisch signifikanter Zusammenhang zwischen der Anzahl der geschriebenen Nachrichten und der erreichten Note zu erkennen. Es ist zu vermuten, dass es latente Gruppen innerhalb der Studenten gibt. Diese zu identifizieren könnte helfen, die Betreuung zielgerichteter zu gestalten. Insgesamt ist als Ergebnis festzuhalten, dass eLearning nicht per se zu einer Verringerung des Betreuungsaufwands führt.



# Literatur

- Chavas, Jean-Paul und Kwansoo Kim (2006). »An econometric analysis of the effects of market liberalization on price dynamics and price volatility«. In: *Empirical Economics* 31, S. 65–82.
- Ćukusić, Maja u. a. (2010). »E-Learning process management and the e-learning performance: Results of a European empirical study«. In: *Computers&Education* 55, S. 554–565.
- Einav, Liran, Ephraim Leibtag und Aviv Nevo (2010). »Recording discrepancies in Nielsen Homescan data: Are the present and do they matter?«. In: *Quantitative Marketing and Economics* 8, S. 207–239.
- Hanau, Arthur (1963). »Rückblick und Ausblick«. In: *Bedeutung und Anwendung ökonomischer Methoden in der agrarwirtschaftlichen Forschung*. Hrsg. von Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaus e.V. Hilstrup: Landwirtschaftsverlag, S. 491–493.
- Henning, Christian H.C.A. und Jerzy Michalek (2008). »Ökonometrische Methoden der Politikevaluation: Meilensteine für eine sinnvolle Agrarpolitik der 2. Säule oder akademische Fingerübung?«. In: *Agrarwirtschaft* 57, S. 232–243.
- Herrmann, Roland (2005). »Gibt es keinen Methodenbeitrag der Agrarökonomie mehr?«. In: *Agrarwirtschaft* 54, S. 297–299.
- Koester, Ulrich (1992). *Grundzüge der landwirtschaftlichen Marktlehre*. 2. Aufl. München: Vahlen.
- Loy, Jens-Peter und Carsten Steinhagen (2009). »Preissenkung und Kompensation auf dem EU-Milchmarkt«. In: *Agrarwirtschaft - Zeitschrift für Betriebswirtschaft, Marktforschung und Agrarpolitik* 58 (5/6), S. 259–268.
- Petrick, Martin (2004). *Can econometric analysis make (agricultural) economics a hard science?* Discussion Paper 69. Halle (Saale): IAMO - Institut für Agrarentwicklung in Mittel- und Osteuropa.

## Literatur

- Reisch, Erwin (1967a). »Einführung«. In: *Quantitative Methoden in den Wirtschafts- und Sozialwissenschaften des Landbaus*. Hrsg. von Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaus e.V. München: Bayrischer Landwirtschaftsverlag, S. 9–10.
- (1967b). »Zusammenfassung«. In: *Quantitative Methoden in den Wirtschafts- und Sozialwissenschaften des Landbaus*. Hrsg. von Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaus e.V. München: Bayrischer Landwirtschaftsverlag, S. 455–458.
- Revermann, Christoph (2006). *eLearning in Forschung, Lehre und Weiterbildung in Deutschland, Sachstandsbericht zum Monitoring eLearning*. Arbeitsbericht 107. Besucht am 07.09.2010. Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag (TAB). URL: <http://www.tab-beim-bundestag.de/de/pdf/publikationen/berichte/TAB-Arbeitsbericht-ab107.pdf>.
- Schwarz, Christine (2001). »eLearning und Bildungspolitik: Von der Nachhaltigkeit hoher Erwartungen«. In: *Virtueller Campus. Szenarien - Strategien - Studium*. Hrsg. von E. Wagner und M. Kindt. Waxmann, S. 374–384.
- Twigg, Carol A. (2003). »Improving learning and reducing costs: new models for online learning«. In: *EDUCAUSE Review Magazine* 38, S. 28–38.

# Kapitel 2

## VALIDATION STUDY OF CONSUMER SCANNER AND RETAIL SCANNER DATA

Autor:

Meike Wocken

Zur Publikation eingereicht bei  
International Journal of Statistics & Economics  
(vormals Bulletin of Statistics & Economics)

## Abstract

*The aim of this study is, first, to analyse consumer scanner panel data with regard to measurement error in price records and, second, to classify potential biases. Studies, which use such data, commonly rely on the assumption of unbiased data. In the used consumer scanner data participating households record each bought article, corresponding price, and attended chain after purchase. For validation, retail scanner data are used which contain information about sales on store-level. Matching between both datasets is impossible on store-level. Therefore, only articles with uniform pricing across all stores of a chain and particular calendar week are analysed. With it, the visited store in detail is negligible. Information about the chain is sufficient. In total, I find an additive, normal distributed measurement error in consumer price data, which is correlated with the price level. The assumption of a classical measurement error has to be rejected. Additional, I observe statistically significant effects of age and income on the measurement error. The calculated reliability ratio is statistically significant less than 1 for each product category, except UHT milk. Hence, in linear regression models estimates are biased, whether the mismeasured price is an explanatory variable or it is the dependent variable regressed on age or income. Due to the non-classical measurement error the bias is rather complex. As a result of the restrictive choice of analysed articles, detected measurement error is only a lower limit. In general the measurement error and the resulting bias are expected to be larger in size.*

## 2.1 Introduction

This study analyses consumer scanner and retail scanner data with regard to measurement error in price observations. Retail scanner data provide weekly information about sales on store-level. The universal product code (UPC) is automatically recorded at cash-point with price information and number of sales. Consumer scanner data are generated after purchase by households participating in the consumer panel. The UPC of each bought article is scanned, accompanied by added information about price and amount. It results in longitudinal data for every panellist.

Academics as well as commercials use such scanner data to analyse e.g. consumers'

## 2.1 Introduction

buying behaviour. (Bucklin and Gupta, 1999; Goodstein, 1994; Adamowicz and Swait, 2011) Such studies commonly rely on the assumption of unbiased data. Critical examinations of scanner data show that this assumption is not always warranted. Goodstein (1994) examined retail scanner data. He discovered discrepancies between prices marked on shelves, where purchase decisions are made, and the recorded prices at check-out. Zhen et al. (2009) compared household expenditures in the Nielsen Homescan panel with household expenditures from the Bureau of Labor Statistic's Consumer Expenditure Diary Survey. They find differences which indicate that e.g. higher-income households as well as households who purchase higher quantities of food might neglect some purchases. Lusk and Brooks (2011) recently found evidence for sample selection and participation biases in consumer scanner data. Einav, Leibtag, and Nevo (2010) had a critical look at the Nielsen Homescan consumer panel data. For validation they used loyalty card records from a retailer. Observed discrepancies in half of the analysed prices are explained, at least in part, due to the way Nielsen imputes store-level prices when available. (Einav, Leibtag, and Nevo, 2010) These findings raise the question whether misreported prices also occur if the actual price recorded by households is declared. Further, it is necessary to enhance the study to classify the measurement error and potential bias in more detail.

To characterize potential measurement errors German household scanner data *ConsumerScan* from GfK Panel Services Germany are used. For validation retail scanner data from Madakom GmbH are applied. Both datasets are available for the years 2000 and 2001. This form of validation data is the main modification to the study from Einav, Leibtag, and Nevo, who used information from loyalty cards. Access to validation data only rarely exists, although it is needed to detect reporting errors. (Zhen et al., 2009) Using retail scanner data is one possible approach to construct validation data. This method is limited to exposure only price discrepancies. It cannot uncover reporting errors in purchased quantities.

In the next section I describe the impact of recording errors in the context of the simple linear regression model. The section that follows provides a description of data and preparing validation data. In section 4, I examine the detected measurement error with regard to structure, correlation, and resulting bias. Findings are discussed in section 5. Finally, concluding remarks are given.

## 2.2 Measurement error and its impact

At first I have to define the term measurement error in this study. Measurement error means a distinction between price observation in consumer scanner data and price observation in retail scanner data. Consistent with Einav, Leibtag, and Nevo (2010) retail scanner data are assumed to be true. Because of automatic recording at the cash point it is presumed that these data have high quality. Therefore measurement errors in consumers' data need to be interpreted as discrepancies from retailers' data. This interpretation should be kept in mind. Certainly, measurement errors can also occur in retail scanner data. In the consumer scanner dataset used for this study panellists add the number of bought units of a specific product, total expenditure on this product, and whether there had been promotion activity by hand to the scanned universal product code (UPC). At best it is done with the aid of the receipt. Otherwise the panellist needs the ability of an accurate price recall. Various studies supply evidence that in general exact price knowledge is low. (Evanschitzky, Kenning, and Vogel, 2004) The abandonment of a cross-check using the receipts is a possible source of measurement error. Therefore, measurement error is more likely to occur in consumers' data. For this reason I also refer to the measurement error as recording error.

To get insight into the impact of a potential recording error, let us consider two variables  $X$  and  $Y$ . They are assumed to be latent. The corresponding observed variables  $D$  and  $W$  contain an additional measurement error. Given  $x_i$  and  $y_i$  it is

$$D_i = y_i + q_i \quad \text{and} \quad W_i = x_i + u_i.$$

$q_i$  and  $u_i$  have conditional mean 0. The measurement errors  $(u_i, q_i)$  have to be independent over  $i$ . Further assumptions are (Buonaccorsi, 2010, pp. 77)

$$\text{Var}(u_i|x_i) = \sigma_{u_i}^2, \quad \text{Var}(q_i|y_i) = \sigma_{q_i}^2, \quad \text{Cov}(u_i, q_i|x_i, y_i) = \sigma_{uq_i}.$$

One further assumption is the "*classical error assumption*", e.g. measurement errors are independent or at least uncorrelated with the true value. Many researcher assume the classical error, although it is often done due to "*convenience rather*

## 2.2 Measurement error and its impact

than conviction". (Bound, Brown, and Mathiowetz, 2001, p. 3708) In validation studies it has been little empirical evidence of classical measurement error. The non-classical error is often more likely. (Hyslop and Imbens, 2001) The advantage of the classical measurement error is that its impact is well understood. In the simple linear regression, with fixed  $x_i$ ,  $Y_i = D_i - q_i$  and  $x_i = W_i - u_i$  can be substituted into  $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ . It results in

$$D_i = \beta_0 + \beta_1 W_i + \epsilon_i^*$$

with  $\epsilon_i^* = -\beta_1 u_i + \epsilon_i + q_i$ . Due to  $Cov(W_i, \epsilon_i^*) = \sigma_{uqi} - \beta_1 \sigma_u^2$  and  $E(\epsilon_i^* | W_i = w_i) = -\beta_1(w_i - x_i)$  standard regression assumptions are violated. In the case of normal additive measurement error and constant measurement error variances/covariance the bias can be stated. As shown in Buonaccorsi (2010) it leads to

$$D_i | w_i = \gamma_0 + \gamma_1 w_i + \delta_i.$$

$\delta_i$  is normal distributed  $N(0, \sigma_\delta^2)$ .  $\gamma_0, \gamma_1, \sigma_\delta^2$  are the naive estimators and result in

$$\gamma_0 = \beta_0 + \frac{E(X)}{\sigma_X^2 + \sigma_u^2} (\beta_1 \sigma_u^2 - \sigma_{uq}) \quad (2.1)$$

$$\gamma_1 = \left(1 - \frac{\sigma_u^2}{\sigma_X^2 + \sigma_u^2}\right) \beta_1 + \frac{\sigma_{uq}}{\sigma_X^2 + \sigma_u^2} \quad (2.2)$$

$$\sigma_\delta^2 = \sigma^2 + \sigma_q^2 + \beta_1^2 \sigma_x^2 - \frac{(\beta_1 \sigma_x^2 + \sigma_{uq})^2}{\sigma_x^2 + \sigma_u^2}. \quad (2.3)$$

With  $\sigma_{uqi} \neq 0$ , the bias in the estimates can go in both directions. If measurement error is only in the dependent variable, e.g.  $u_i = 0$  and  $\sigma_{uqi} = 0$ , consistent estimates are gained. Standard errors of the estimators are only inflated. If  $q_i = 0$  and  $\sigma_{uqi} = 0$ , that means no error in the response, the slope estimate is

$$\gamma_1 = \left(1 - \frac{\sigma_u^2}{\sigma_X^2 + \sigma_u^2}\right) \beta_1.$$

One way to quantify the classical measurement error in this case is the reliability ratio  $\kappa = \left(1 - \frac{\sigma_u^2}{\sigma_X^2 + \sigma_u^2}\right) = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_u^2}$ . Because  $0 < \kappa < 1$ , measurement error leads to attenuation ( $|\gamma_1| < |\beta_1|$ ) in the slope estimator. (Buonaccorsi, 2010, p. 80)

## Validation study of consumer scanner and retail scanner data

If this ratio  $\kappa$  of true and observed variance is known and the assumption of a classical measurement error holds it is possible to construct unbiased estimators in regression analysis. (Fuller, 1987, p. 5)

In the more general case of dependence between measurement error and the true value the reliability ratio is no longer appropriate. If  $Cov(X_i, u_i) \neq 0$  in the attenuation bias  $(1 - \kappa)$  an additional covariance term is appended. The generalization of the reliability ratio is

$$\kappa_A = \frac{Cov(W_i, X_i)}{Var(W_i)}.$$

With this,  $\kappa_A$  is the slope estimate from a simple linear regression of the latent variable on the observed, mismeasured variable.  $\kappa_A$  states "*the proportion of a change in the observed variable that translates into a change in the true, latent variable.*" (Bound and Krueger, 1991)

The stated biases are only exact under normal structural model and constant variance/covariance of the normal measurement error. Otherwise, the expressions are approximate considerations. (Buonaccorsi, 2010, pp. 81) In Buonaccorsi (2010) analysis is enhanced to a multiple linear model. One important result in the more complex model is that measurement error in one of the variables causes bias in all estimated coefficients, whether they are measured with error or not. (Buonaccorsi, 2010, pp. 109) It shows the importance of dealing with the problem of measurement error.

## 2.3 Data and data preparation

In this study German retail scanner and consumer scanner data are used. Both data sets contain observations from the years 2000 and 2001. The universal product code (UPC) identifies each item. Consumer scanner data are *Consumer Scan* from GfK Panel Services Germany. In this dataset households record the total amount (EUR) spend on a specific UPC in one purchase trip. They add the number of bought units of each item. From this, I calculate the price per unit. Retail scanner data are from Madakom GmbH. In this dataset each observation consists of the scanned UPC, the paid price (DM), and the total number of bought



### 2.3 Data and data preparation

items. Madakom GmbH offers additional information about the number of units belonging to a particular item. With this, I calculate the price for a single unit and convert it to EUR.

To be able to compare prices from consumer scanner data with observed prices from retail scanner data, I need the true retail price for each UPC. Because I only know the name of the chain but cannot identify the individual store, matching between the visited store and the recording from cash-point is impossible. To balance this shortage, in the retail scanner data I try to find chains with uniform price policy. Uniform price policy means the price needs to be identical across all stores of a chain and particular calendar week. The uniform price policy makes sure that it is negligible which store in detail was visited by the consumer due to the uniform price in each store. To identify uniform price policy in retail scanner data I define two requirements:

1. across all reporting stores of a particular chain the prices, which belong to a UPC with respect to calendar week, are uniform for at least 99 of the 104 observed calendar weeks (95%), and
2. for all observations with uniform price policy, the median of the number of reporting stores is 8.

Requirement 2 of the definition is necessary to avoid uniform pricing due to insufficient number of reporting stores. If in extreme only one store reports, it ends in uniform prices. If I find such a uniform price for a particular UPC with regard to chain and calendar week, with a high degree of certainty the uniform price is valid for all observations for the particular UPC, calendar week, and chain within Germany. To clarify the definition various situations are shown in the Figures 2.1, 2.2, 2.3, and 2.4. Each figure represents observations of an individual UPC at one chain. The x-axis represents the time-axis. Every dot indicates a price record from a particular store belonging to the chosen chain. Prices can be read of the left-hand axis. The number of reporting stores is plotted using a blue solid line. The blue solid line is only drawn in case all reporting stores report one uniform price. If there is no uniform price, the blue solid line disappears. On the right hand side the number of reporting stores can be read of.

Figure 2.1 displays an example of uniform price policy which complies with the

Validation study of consumer scanner and retail scanner data

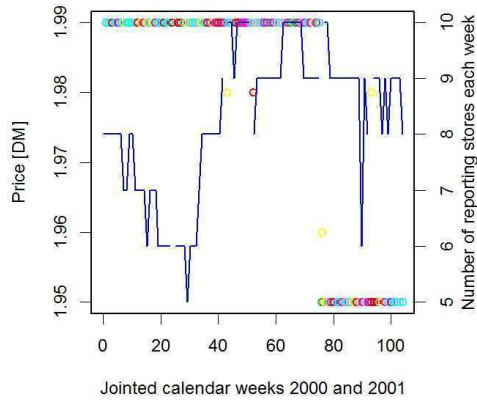


Figure 2.1: Uniform price policy regarding applied definition.

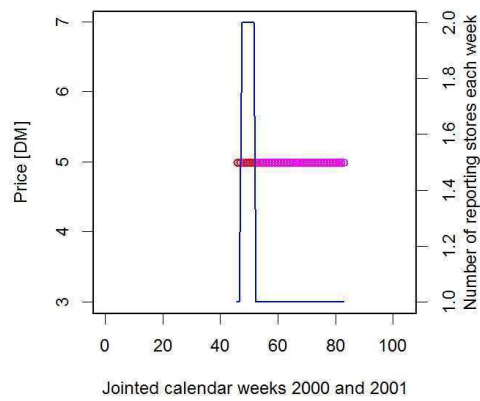


Figure 2.2: No uniform price policy regarding applied definition.

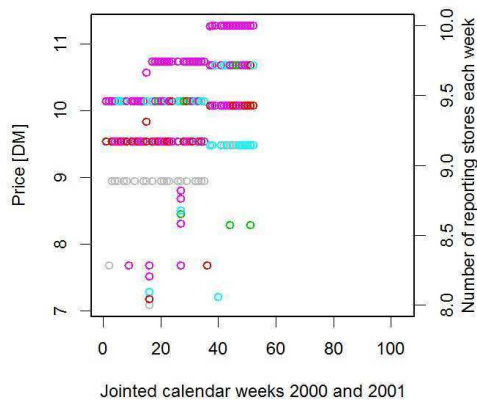


Figure 2.3: No uniform price policy regarding applied definition.

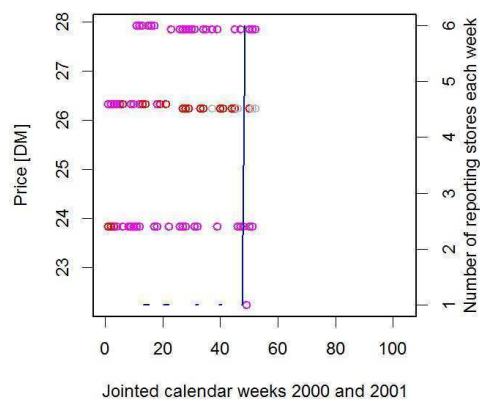


Figure 2.4: No uniform price policy regarding applied definition.

### 2.3 Data and data preparation

applied definition. The blue solid line indicates weeks with uniform prices and the number of reporting stores, which can be read of on the right-hand axis. Each single price observation is shown by a coloured point. I select such uniform retail scanner prices to compare with consumer scanner prices. In Figure 2.2 again each point indicates a price observation. All price observations are located at 5 DM. At first view it looks like uniform price policy for this particular UPC and chain. The blue solid line shows on the right-hand axis that in maximum two stores reported. With only two reporting stores, uniform pricing might coincidence. Therefore, with regard to the given definition, it is no uniform price policy. Figure 2.3 also shows no uniform pricing. It is indicated by the absence of the blue solid line. Finally, Figure 2.4 shows the case where uniform pricing only appears if just one store reports per calendar week (with exception of sixth reporting stores at the 48th calendar week). Again, it is no uniform pricing regarding to the used definition.

To be able to identify uniform pricing I need chains with a high number of stores in the dataset. Therefore, I choose the following 5 chains for further analysis:

- Kaiser's Markt (17 stores);
- Plus (16 stores);
- Real (12 stores);
- Rewe (12 stores);
- Tengelmannmarkt (9 stores).

Remaining chains only have three or less stores within the dataset. For comparison I can only choose product categories which occur in both datasets. These are

1. UHT milk;
2. UHT cacao;
3. cream cheese;
4. pizza;
5. cornflakes;
6. orange juice.

Additional analysis is done for the following product categories which only appear

## Validation study of consumer scanner and retail scanner data

Product group	Real	Rewe	Kaisers	Plus	Tengelmann	Total
UHT milk	0	0	0.026	0	0	0.007
UHT cacao	0	0	0	0	0	0
Cream cheese	0.007	0.043	0.030	0.069	0	0.022
Pizza	0.016	0.035	0	0.060	0	0.018
Cornflakes	0.154	0	0	0	0	0.061
Orange juice	0.020	0	0.082	0.080	0.044	0.04
Total	0.016	0.025	0.03	0.056	0.011	

Table 2.1: Selected UPCs with uniform pricing relative to total number of UPCs (in retail scanner data).

in retail scanner data: 7. long grain rice; 8. jam; 9. chocolate creme; 10. fruit-cocktail can; 11. potato crisps. I conduct this enhancement to get a better understanding of the used constraints with respect to uniform prices.

Table 2.1 shows the ratios of selected UPCs with uniform pricing to the total number of UPCs with respect to product category and chain in the retail scanner data.<sup>1</sup> The applied strict definition involves huge reduction of analysed UPCs. Largest share of uniform price UPCs can be found at the chain Plus which is a discount supermarket chain. In the product categories largest shares appear for cornflakes, potato crisps, and chocolat creme. Especially for these products, it might be explained by joint advertising efforts. In the following, only the small share of items fulfilling definition of uniform pricing is used. I assume the uniform prices, filtered by definition, can be imputed to be the true price for a particular UPC within calendar week and chain. In the next section I compare the true price with prices reported in consumer scanner data.

## 2.4 Examine measurement error

To be able to examine the measurement error I consider retail scanner data to be true data. It is done in accordance to Einav, Leibtag, and Nevo (2010). The

---

<sup>1</sup>Additional analysed product groups have the following ratios in total: long grain rice 0.039, jam 0.027, chocolate creme 0.041, fruit-cocktail can 0, potato crisps 0.046.

## 2.4 Examine measurement error

term retail scanner price refers to the uniform prices filtered in data preparation. Further, I assume the following relationship between the observed consumer scanner price and the retail scanner price:

$$\begin{aligned}
 \mu_{t(i)} &:= \text{uniform price from retail scanner data at time } t \\
 t &\in 1, \dots, 104 \quad (t \text{ is a function of the observation } i) \\
 p_i &:= \text{observed price in consumer scanner data} \\
 i &\in 1, \dots, N \quad (N \text{ is number of observations in consumer scanner data.}) \\
 p_i &= \mu_{t(i)} + u_i. \tag{2.4}
 \end{aligned}$$

where the measurement error  $u_i$  is a  $(0, \sigma_u)$  distributed random variable. Table 2.2 gives mean and standard deviation of the measurement error with respect to category and chain. Sign and magnitude of the measurement error varies across chains and categories. 44.87% of the analysed observations possess  $u_i \neq 0$ . This finding is consistent with Einav, Leibtag, and Nevo (2010). In their study half of price observations contain a measurement error. To be able to compare the overall error with results from Einav, Leibtag, and Nevo the relative frequency of  $\log(\text{consumer scanner price}/\text{retail scanner price})$ , which is equal to the difference of log prices, is presented in Figure 2.5.

Similar to Einav, Leibtag, and Nevo (2010), the average difference of log prices is with 0.0375 close to zero as well as most of the probability mass. In contrast to Einav, Leibtag, and Nevo (2010), I do not conclude from this figure that the measurement error is close to a *classical error*. As mentioned above, the classical error is not only the assumption that "*even when the price does not match, the differences are small*". (Einav, Leibtag, and Nevo, 2010) I go one step further to analyse the measurement error in consumer scanner prices. Detailed classification of the measurement error helps to understand the resulting bias which might occur in studies that do not correct for recording errors. Therefore, in the following subsections at first I test the additive and normal structure of the error. Second, I have a closer look at the variance-covariance-matrix of true price and the measurement error. Finally, I try to quantify potential biases.

Validation study of consumer scanner and retail scanner data

	Real	Rewe	Kaisers	Plus	Tengelmann	Total
Pizza	0.0213 (0.4024)	0.0818 (0.4904)		0.1300 (0.4261)		0.0681 (0.4300)
UHT milk			-0.0002 (0.0153)			-0.0002 (0.0153)
Orange juice	-0.0087 (0.1169)		0.0042 (0.1529)	-0.0014 (0.0310)	-0.0635 (0.1503)	-0.0125 (0.0973)
Cream cheese			-0.0041 (0.0607)			-0.0041 (0.0607)
Cornflakes	0.2007 (0.4755)					0.2007 (0.4755)
Total	0.0171 (0.3495)	0.0818 (0.4904)	-0.0001 (0.0567)	0.0725 (0.3269)	-0.0635 (0.1503)	0.0395 (0.3464)

Table 2.2: Mean of the absolute measurement error  $u_i = p_i - \mu_{t(i)}$  [EUR] with standard deviation in parentheses.

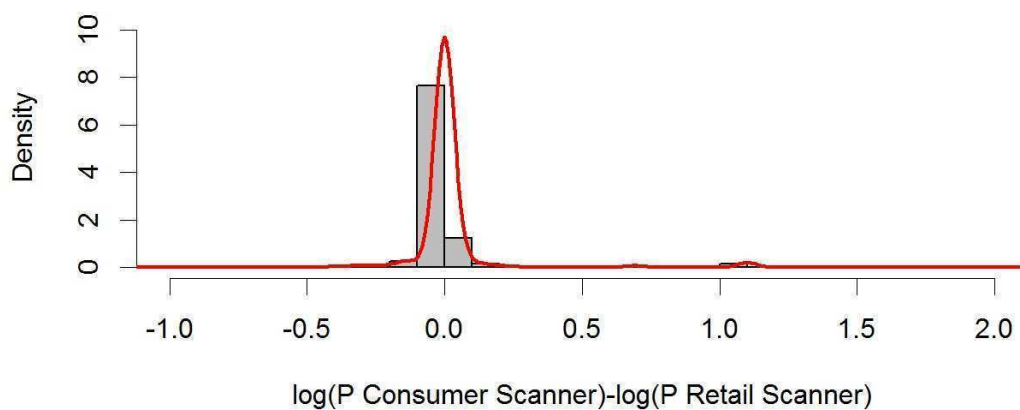


Figure 2.5: Probability density and estimated kernel density of the  $\log(\text{consumer scanner price}) - \log(\text{retail scanner price})$ . Mean and standard deviation are 0.0375 and 0.2881.

### 2.4.1 Additive structure of measurement error

As noted earlier the biases expressions (2.1), (2.2), and (2.3) are only exact in case of an additive, normal measurement error having constant variance. To test this characteristic, I use a  $\chi^2$  distributed test statistic. For this I assume that the observed price in consumer scanner data consists of the true price plus some normally distributed error. Again, the true price is equal the uniform price given in retailers' data.

$$p_i = \mu_{t(i)} + N(0, \sigma) \Rightarrow \frac{p_i - \mu_{t(i)}}{\sigma} \sim N(0, 1) \quad (2.5)$$

$$\hat{\sigma} = \sqrt{\frac{1}{N-1} \cdot \sum_{i=1}^N \left[ (p_i - \mu_{t(i)}) - \overline{(p_i - \mu_{t(i)})} \right]^2}$$

After applying standardization (2.5) to observation  $i$  it is now possible summing standardized discrepancies between observed and true prices across all observations  $N$ . With the assumption of independent observations I gain a  $\chi^2$ -distributed test statistic

$$\Rightarrow \sum_{i=1}^N \left( \frac{p_i - \mu_{t(i)}}{\sigma} \right)^2 \sim \chi^2(N). \quad (2.6)$$

Table 2.3 presents the calculated results. First column states the number  $n$  of observations from consumer scanner data matching a uniform price in retail scanner data. Second column gives calculated test statistic. The 95% quantile from the  $\chi^2$  distribution with  $n$  degrees of freedom is written in the third column. Last column presents the corresponding p-values. Test statistic is calculated for each product category, each chain, and in total. The assumption of an additive, normal measurement error has to be rejected for Plus and Tengelmann. Also, the assumption must be refused for pizza and cornflakes. With this, resulting bias cannot exactly be stated in simple linear regression. If there is a simple linear regression having an error prone price variable from either pizza, cornflakes, Plus, or Tengelmann on the right hand side, stated biases in (2.1), (2.2), and (2.3) are

*Validation study of consumer scanner and retail scanner data*

only approximates. Aggregated data on the overall level achieve the test statistic 23701.69 ( $N = 23399$ ) and p-value 0.0813. In total, the assumption of an additive, normal distributed measurement error cannot be rejected.

	$n$	$\hat{\sigma}$	test statistic	95% quantile from $\chi^2(n)$ distribution	p-value
Real	10125	0.3495	10148.2154	10360.1983	0.4334
Rewe	2560	0.4904	2630.3116	2678.8222	0.1628
Kaisers	1280	0.0567	1279.0031	1364.3452	0.5026
Plus	8380	0.3269	8791.6849	8594.0744	0.0009
Tengelmann	1054	0.1503	1241.5011	1130.6399	0.0001
Pizza	14265	0.43	14621.658	14543.9617	0.0179
UHT milk	995	0.0153	994.099	1069.4953	0.5021
Orange juice	7752	0.0973	7878.2331	7957.9398	0.1554
Cream cheese	140	0.0607	139.6335	168.613	0.4928
Cornflakes	247	0.4755	290.0198	284.6599	0.0312
Total	23399	0.3464	23701.6897	23755.962	0.0813

Table 2.3:  $\chi^2$  test statistic with regard to product category and chain. Null hypothesis is an additive, normal distributed measurement error.

## 2.4.2 Correlation of true price and the measurement error

For the analysis of the correlation between true price and measurement error I calculate Pearson's product moment and Spearman's rank correlation coefficient. Due to the result from the  $\chi^2$ -distributed test statistic, Pearson's coefficient is not appropriate for Plus and Tengelmann as well as pizza and cornflakes. Table 2.4 states the correlation coefficient, the two-sided p-value for the null hypothesis correlation coefficient is equal zero, and the 95% confidence interval for Pearson's correlation coefficient. The independence of measurement error and true price cannot be rejected for Kaisers, Tengelmann, and the product category cream cheese. Overall, the independence of measurement error and true price has to be rejected. The assumption of the classical measurement error does not hold.



## 2.4 Examine measurement error

	$Corr(u_i, \mu_{t(i)})$	p-value	95% confidence interval
Real	-0.3575	0.0000	[-0.3744 , -0.3404]
Rewe	-0.7045	0.0000	[-0.7235 , -0.6844]
Kaisers	-0.0345	0.2180	[-0.0891 , 0.0204]
Plus	0.3140	0.0000	
Tengelmann	-0.0383	0.2136	
Pizza	-0.1257	0.0000	
UHT milk	-0.0758	0.0167	[-0.1373 , -0.0138]
Orange juice	-0.2283	0.0000	[-0.2492 , -0.2070]
Cream cheese	-0.0502	0.5559	[-0.2143 , 0.1167]
Cornflakes	-0.4815	0.0000	
Total	-0.2854	0.0000	[-0.2971 , -0.2735]

Table 2.4: Correlation coefficient of measurement error and true price.

Pearson's product moment correlation coefficient is stated in first column. In case of missing confidence interval Spearman's rank correlation coefficient is calculated. P-value refers to a two-sided test basing on the calculated correlation coefficient. Null hypothesis is correlation coefficient equal zero. The interval states the 95% confidence interval for the correlation coefficient derived from Pearson's coefficient.

### **2.4.3 Quantifying biases**

In the previous subsections in total there is evidence for an additive, normal measurement error, which does not match the assumptions of a classical measurement error. I.e., the condition of independence between measurement error and true value is not fulfilled. At the beginning of this section, this finding should not give reason for concern: An error prone variable on the left-hand-side of a linear regression leads to biased standard errors and consistent estimates. Only if the measurement error depends on external regressors, estimated coefficients are biased. Bound and Krueger, 1991 To quantify the potential biases, I compare regression results from regression of consumer scanner price and regression of true price on various sozio-demographics (categorical data), chain, and product category. It is done in accordance to Einav, Leibtag, and Nevo (2010). They find differences in estimates of a price regression using either Nielsen Homescan price or retailer price. Also, in their study households reporting more accurately have differences in their characteristics to households reporting less accurately. Therefore, it is reasonable that they achieve different results. To be able to quantify the bias, I use a third regression. The magnitude of the bias is equal to the estimated coefficients from the regression of measurement error on the right-hand-side-variables. (Bound and Krueger, 1991) Table 2.5 presents the results for the used sozio-demographics, chains, and product categories. Table 2.6 states corresponding goodness-of-fit measures. It should be clear, with third regression the first two regressions are redundant. They are only presented in reference to Einav, Leibtag, and Nevo (2010).

Table 2.7 presents results of further investigation. First column gives the number of matched observations in this particular chain or category. Second column shows calculated reliability ratio  $\kappa$ . Only in case of a normal classical measurement error and a simple linear regression, size and direction of the bias is exact indicated by  $\kappa$ . If a non-classical measurement error exist a general statement about the implied bias is not possible. Therefore, in column 3 the calculated alternative reliability ratio  $\kappa_A$  is written. It reports the conversion from a change in the mismeasured variable into the true variable. In case of a high reliability,  $\kappa_A$  would be close to 1. Estimation is done using OLS technique including standard error and 95% confidence intervals for  $\kappa_A$ .

## 2.4 Examine measurement error

Dependent variable	Consumer Scanner Price		True Price		Measurement Error	
(Intercept)	1.7724	(0.0778)***	1.6018	(0.0897)***	0.1706	(0.0674)*
AGE 19 or younger	-0.1569	(0.0454)***	-0.2650	(0.0524)***	0.1081	(0.0394)**
AGE 20-24	-0.0723	(0.0188)***	-0.0933	(0.0217)***	0.0210	(0.0163)
AGE 25-29	-0.0126	(0.0168)	-0.0304	(0.0194)	0.0178	(0.0146)
AGE 30-34	0.0084	(0.0168)	-0.0284	(0.0194)	0.0368	(0.0146)*
AGE 35-39	-0.0130	(0.0168)	-0.0409	(0.0194)*	0.0279	(0.0146)
AGE 40-44	-0.0210	(0.0170)	-0.0624	(0.0196)**	0.0414	(0.0147)**
AGE 45-49	-0.0155	(0.0167)	-0.0335	(0.0192)	0.0180	(0.0144)
AGE 50-54	-0.0119	(0.0173)	-0.0284	(0.0200)	0.0165	(0.0150)
AGE 55-59	0.1120	(0.0174)***	0.0938	(0.0201)***	0.0182	(0.0151)
AGE 60-64	0.0297	(0.0164)	0.0270	(0.0189)	0.0026	(0.0142)
AGE 65-69	0.0837	(0.0180)***	0.0385	(0.0208)	0.0452	(0.0156)**
Employee	0.1159	(0.0700)	0.0844	(0.0807)	0.0315	(0.0607)
Worker	0.1145	(0.0702)	0.0834	(0.0809)	0.0311	(0.0608)
Civil servant	0.1479	(0.0705)*	0.1154	(0.0812)	0.0325	(0.0611)
Freelancer	0.0794	(0.0703)	0.0329	(0.0811)	0.0465	(0.0610)
Self-employed businessmann	0.1382	(0.0713)	0.0618	(0.0822)	0.0764	(0.0618)
Liberal profession	0.1832	(0.0747)*	0.1672	(0.0862)	0.0160	(0.0648)
Income 499 EUR or less	0.0299	(0.0351)	-0.0264	(0.0404)	0.0564	(0.0304)
I. 500-749 EUR	-0.0694	(0.0228)**	-0.0361	(0.0263)	-0.0333	(0.0198)
I. 750 - 999 EUR	-0.0940	(0.0225)***	-0.0627	(0.0259)*	-0.0313	(0.0195)
I. 1000 - 1249 EUR	-0.0893	(0.0195)***	-0.0441	(0.0225)	-0.0452	(0.0169)**
I. 1250 - 1499 EUR	-0.0994	(0.0192)***	-0.0564	(0.0221)*	-0.0431	(0.0166)**
I. 1500 - 1749 EUR	-0.0888	(0.0186)***	-0.0587	(0.0214)**	-0.0301	(0.0161)
I. 1750 - 1999 EUR	-0.0716	(0.0185)***	-0.0490	(0.0213)*	-0.0227	(0.0160)
I. 2000 - 2249 EUR	-0.0766	(0.0183)***	-0.0400	(0.0211)	-0.0366	(0.0159)*
I. 2250 - 2449 EUR	-0.0417	(0.0189)*	-0.0300	(0.0218)	-0.0117	(0.0164)
I. 2500 - 2749 EUR	-0.0460	(0.0190)*	0.0004	(0.0219)	-0.0464	(0.0165)**
I. 2750 - 2999 EUR	-0.1031	(0.0205)***	-0.0620	(0.0236)**	-0.0411	(0.0177)*
I. 3000 - 3249 EUR	-0.0532	(0.0183)**	-0.0084	(0.0211)	-0.0448	(0.0159)**
I. 3250 - 3499 EUR	0.0155	(0.0297)	0.0528	(0.0343)	-0.0373	(0.0258)
I. 3500 - 3749 EUR	-0.0825	(0.0279)**	-0.0504	(0.0321)	-0.0322	(0.0242)

Table 2.5 continued on next page.

Validation study of consumer scanner and retail scanner data

Dependent variable	Consumer Scanner Price		True Price		Measurement Error	
Rewe	0.2424	(0.0090)***	0.1977	(0.0104)***	0.0447	(0.0078)***
Kaisers	0.2336	(0.0334)***	0.1872	(0.0385)***	0.0463	(0.0289)
Plus	-0.0377	(0.0061)***	-0.1073	(0.0070)***	0.0695	(0.0053)***
Tengelmann	0.2120	(0.0136)***	0.2352	(0.0157)***	-0.02324	(0.0118)*
pizza	-0.0747	(0.0256)**	0.0872	(0.0295)**	-0.1619	(0.0222)***
UHT milk	-1.5753	(0.0436)***	-1.3276	(0.0503)***	-0.2476	(0.0378)***
orange juice	-1.1246	(0.0259)***	-0.8833	(0.0298)***	-0.2413	(0.0224)***
cream cheese	-1.1226	(0.0535)***	-0.8687	(0.0617)***	-0.2539	(0.0463)***

Table 2.5: OLS regressions of consumer scanner price, true price, and measurement error on age, professional category, monthly net income per household, chain, and product category. Exogenous variables are categorical. Omitted categories are 70 and older (age), farmer (professional category), 3750 EUR and more (net income per household), Real (chain), and cornflakes (product category). All matched observations (23399) are used. Standard errors are in parentheses. Asterisks indicate statistical significance (\*  $p < .05$ , \*\*  $p < .01$ , and \*\*\*  $p < .001$ ).

Dependent variable	Consumer Scanner Price	True Price	Measurement Error
Residual standard error	0.3943	0.4547	0.3418
$R^2$	0.6526	0.5523	0.0279
$\bar{R}^2$	0.652	0.5515	0.0263
F(39,23359)	1125	738.8	17.19
p-value(F)	< 2.2e-16	< 2.2e-16	< 2.2e-16

Table 2.6: Goodness of fit measures from previous OLS regression Table 2.5.

	$\kappa$ Classical Measurement Error	$\kappa_A$ Non-Classical Measurement Error	95% Confidence interval for $\kappa_A$
Real	0.7880	0.9069 (0.0054)	[0.8964 , 0.9174]
Rewe	0.5283	0.5954 (0.0236)	[0.5491 , 0.6417]
Kaisers	0.9516	0.9584 (0.0061)	[0.9464 , 0.9703]
Plus	0.7867	0.8399 (0.0052)	[0.8297 , 0.8501]
Tengelmann	0.8402	0.8363 (0.0112)	[0.8144 , 0.8582]
Pizza	0.6287	0.7823 (0.0073)	[0.7680 , 0.7967]
UHT milk	0.9844	0.9936 (0.0040)	[0.9858 , 1.0015]
Orange juice	0.8917	0.9565 (0.0040)	[0.9486 , 0.9643]
Cream cheese	0.7914	0.8038 (0.0360)	[0.7326 , 0.8750]
Cornflakes	0.5786	0.6679 (0.0568)	[0.5560 , 0.7797]
Total	0.7935	0.8817 (0.0033)	[0.8752 , 0.8881]

Table 2.7: Reliability ratios.

## 2.5 Discussion of results

Overall, recording errors exist in 44.87% of the 23399 analysed observations. This number is similar to findings in Einav, Leibtag, and Nevo (2010). To assess the impact of the measurement error, I analyse the structure of it. The assumption of an additive normal measurement error has to be rejected for Plus, Tengelman, pizza, and cornflakes. Table 2.3 shows the estimated standard deviation  $\hat{\sigma}$  of the particular recording errors. The standard deviation is unequal to zero for Plus, Tengelman, pizza, and cornflakes. Although the null hypotheses, that the price consists of the true price plus an additive normal distributed recording error, of the  $\chi^2$ -distributed test statistic has to be rejected for these chains and product categories respectively, with nonzero variance I still conclude a measurement error. Made assumptions for  $\chi^2$ -distributed test statistic are very restrictive. It is likely that measurement errors own another structure with regard to distribution. In future one could test for a distribution with heavier tails. In this study detailed knowledge of distribution is not of value, because the bias expressions (2.1), (2.2), and (2.3) are only exact for normal distributed measurement errors.

### *Validation study of consumer scanner and retail scanner data*

Considering all observations in total it yields no rejection of the assumed additive normal recording error. Calculating correlation coefficient shows that the assumption of independence between error and true value has to be rejected in nearly all product categories and chains. Except for Kaisers, Tengelmann, and cream cheese it still holds. First results may be summarised in the statement that the classical measurement error assumption does not comply with findings.

The applied price regressions show the impact of the error prone variable on the left-hand-side (Table 2.5). Comparing retail scanner price and consumer scanner price estimates, it results in different estimates. But only four estimates (three in sozio-demographics and one from pizza) have different signs in these regressions. In contrast to Einav, Leibtag, and Nevo (2010), if a regressor is significant in one regression, it is more often the consumer scanner data which produce significant results. Differences in significance may arise from major variance in the consumer scanner price compared to retail scanner price. To explain the differences in estimated coefficients, I use the correlation of measurement error with used sozio-demographics, chain, and product category. Using an error prone variable on the left-hand-side of regression it results in biased estimates if the measurement error is correlated with exogenous variables. (Bound and Krueger, 1991) Table 2.5 provides the estimated effects of age, professional category, net income per household, chain, and product category on the measurement. At first I look at the age. There are positive effects for all age classes on the measurement error. The omitted category is the age 70 and older. Only four estimates are statistically significant. The class 19 and younger owns the smallest significant estimate. Compared to the oldest class, the significant estimates increase with increasing age. Next, the impact of various professional categories on recording error is in focus. All estimated coefficients of the professional categories are positive. Omitted category is farmer. The estimates are not statistically significant. Further, the coefficient for the lowest net income per household category (499 EUR and less) is positive whereas the other estimates of the income have a negative sign. It suggests that the measurement error is smaller in comparison to the omitted category 3750 EUR and more, except for the lowest income class. But this positive coefficient is not statistically significant. For the income categories only 6 of the 14 estimated coefficients are statistically significant. There is no distinct tendency, only the insight that highest income category has worst measurement error. Table 2.5

	Homescan price $\hat{\beta}_{homescan}$	Retailer price $\hat{\beta}_{true}$	$\hat{\beta}_{me} = \hat{\beta}_{homescan} - \hat{\beta}_{true}$
Household			
income	0.014	0.094	-0.08
age female	-1.247	-1.713	0.466
age male	-0.395	-1.342	0.347

Table 2.8: Results from Einav, Leibtag, and Nevo (2010) using all matched items in their sample.

shows that in both price regressions the impact of various chains and product categories is very large. It is naturally due to the price determination basing on product characteristics. This also explains the highly significant impact of chain as well as product category regarding the measurement error. The magnitude of the recording error depend on the price level, as also shown by the correlation coefficients in Table 2.4.

Einav, Leibtag, and Nevo (2010) use sozio-demographics to a larger extent, not only with the help of categorical data. Therefore, direct comparison is difficult. For some insight I have a closer look on the estimated effects of age and household income. Using results from Einav, Leibtag, and Nevo (2010), I calculate the difference ( $\hat{\beta}_{me}$ ) between estimates from homescan price regression ( $\hat{\beta}_{homescan}$ ) and estimates from retailer price (true price) regression ( $\hat{\beta}_{true}$ ). The difference indicate the bias, which is equal to the impact on the measurement error. Table 2.8 shows the results. They are roughly in line with the ones I get. The recording error increases with increasing income. There is a tendency, that younger age-groups produce less measurement error.

As shown in Table 2.6, explaining the measurement error regression estimates are jointly statistically significant. Overall, with the small  $R^2$  from explaining measurement error and only three estimates having different signs in consumer scanner price regression and true price regression, I conclude the bias from recording error in left-hand-side-variable is small in this application.

To assess the effect of recording error in a right-hand-side-variable I calculate the reliability ratio  $\kappa$ . In case of classical measurement error it states the attenuation

### *Validation study of consumer scanner and retail scanner data*

bias in simple linear regression with measurement error only in the exogenous variable. As mentioned above the assumption of classical measurement error is mostly not reliable, therefore the alternative measure  $\kappa_A$  is additionally used. The alternative measure  $\kappa_A$  quantifies the error transferred into the true observation. Both measures tend to two different things. But, for non-classical measurement error the reliability ratio  $\kappa$  has no meaningful interpretation. As seen in Table 2.7, just using  $\kappa$  despite the non-classical-measurement error, the size of reliability is underestimated. Only for Tengelmann I find  $\kappa > \kappa_A$ .

Calculated confidence intervals show that hypothesis  $\kappa_A = 1$  has to be rejected for all chains and categories except UHT milk. I conclude from this, a regression using UHT milk prices to explain a dependent variable might face no bias. In contrast, in multiple regression model all other analysed prices will induce biases. Bias will take place not only in estimated coefficient for the error prone price-variable. It appears in all estimates. Therefore, the overall impact of measurement error must not be neglected.

The used strategy to construct validation data faces various potential problems. The restrictive definition is applied to be able to neglect information about the visited store in detail, because I only know the name of the chain visited by the consumer. The strategy secures a high probability that the uniform price for each chosen UPC in retail scanner data is the true price within any store of the particular chain regarding UPC and calendar week. The first problem is the assumption retailers' data are true data. These records may also contain inaccuracies. There is evidence for imprecision by the way Madakom matches additional information for each UPC. Due to variations in the article descriptions I assume (but could not verify) Madakom may add the information rather by hand than automatically. The second problem is the representativeness of the analysed sample. As Table 2.1 shows, the restrictive definition heavily reduces the dataset. Only a small share of different items within each product category respective chain is analysed. Pricing strategy varies between supermarket and discounter. Discounter rather constantly offers low prices. (Cleeren et al., 2010) From there, largest share of UPCs is analysed from Plus, the sole discounter in dataset. Third, due to the restrictive nature of the definition chosen UPCs do not vary in price that much. Therefore, in general the measurement error might be larger in size. I



can only detect the lower limit of measurement error. Fourth, in retailers' records the sales promotion variable shows, individual stores execute sales promotions independent of each other without pursuing uniform strategy. These promotion sales indicate slight uncertainty in the applied proceeding. But, found results are similar to findings from Einav, Leibtag, and Nevo (2010). It suggests that made assumptions are not critical for general conclusions from this study. The delimited representativeness of validation data is much more a problem for constructing a correction method for resulting biases. Due to the restrictive validation dataset it is not possible to use the correction method like Einav, Leibtag, and Nevo (2010) apply.

## 2.6 Conclusion

This study investigates the measurement error in consumer scanner dataset. I show the existence of measurement error in the data. Findings from this study confirm results from Einav, Leibtag, and Nevo (2010). Misreported prices also occur if the actual price recorded by household is used instead of store-level prices. In total I find an additive normal distributed recording error. The assumption of classical measurement error does not hold, because the independence between error and true value has to be rejected. An immediate consequence of this finding is that potential bias in regression analysis is rather complex.

In contrast to the existing study from Einav, Leibtag, and Nevo (2010), comparing Nielsen Homescan data with loyalty card records, I use for validation purpose retailer scanner records. Retail data are more readily available. To construct validation data from retail data I use a restrictive definition to find chains with uniform price policy for a selection of UPCs. Due to this method only a small share of articles is analysed. These commonly are articles with minor price changes. The representativeness of the analysed sample is seriously affected. For articles with augmented price changes I expect an increased appearance of recording errors. Therefore, the detected measurement error is only a lower limit. Although in this study bias is rather small, in general the resulting bias should not be underestimated.

I find statistically significant effects from age and income on the recording error.

### *Validation study of consumer scanner and retail scanner data*

The professional category seems to be statistically irrelevant for recording correctness. The accuracy in reporting can be improved by cross-checks, e.g. using the receipts.

Though, validation method applied in this study is very restrictive, I believe it is a good method to get at least an impression of the lower limit of the measurement error. Validation data like loyalty card data are not as easy accessible. Overall, the impact of recording errors in such scanner data should not be neglected for convenience.

# Bibliography

- Adamowicz, Wiktor and Joffre D. Swait (2011). “Discrete Choice Theory and Modelling”. In: *The Oxford handbook of the economics of food consumption and policy*. Ed. by Jayson L. Lusk, Jutta Roosen, and Jason Shogren. Oxford, UK: Oxford University Press, pp. 119–151.
- Bound, John, Charles Brown, and Nancy Mathiowetz (2001). “Measurement error in survey data”. In: *Handbook of Econometrics*. Ed. by J.J. Heckman and E.E. Leamer. Vol. 5. Handbook of Econometrics. Elsevier. Chap. 59, pp. 3705–3843.
- Bound, John and Alan B. Krueger (1991). “The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?” In: *Journal of Labor Economics* 9, pp. 1–24.
- Bucklin, Randolph E. and Sunil Gupta (1999). “Commercial Use of UPC Scanner Data: Industry and Academic Perspectives”. In: *Marketing Science* 18, pp. 247–273.
- Buonaccorsi, John P. (2010). *Measurement Error: Models, Methods, and Applications*. Boca Raton, Florida: CRC Press.
- Cleeren, Kathleen et al. (2010). “Intra-and Interformat Competition Among Discounters and Supermarkets”. In: *Marketing Science* 29.3, pp. 456–473.
- Einav, Liran, Ephraim Leibtag, and Aviv Nevo (2010). “Recording discrepancies in Nielsen Homescan data: Are the present and do they matter?” In: *Quantitative Marketing and Economics* 8, pp. 207–239.
- Evanschitzky, Heiner, Peter Kenning, and Verena Vogel (2004). “Consumer price knowledge in the German retail market”. In: *Journal of Product & Brand Management* 13, pp. 390–405.
- Fuller, Wayne A. (1987). *Measurement Error Models*. New York: Wiley.
- Goodstein, Ronald C. (1994). “UPC Scanner Pricing Systems: Are They Accurate?” In: *The Journal of Marketing* 58, pp. 20–30.

## *Bibliography*

- Hyslop, Dean R. and Guido W. Imbens (2001). “Bias From Classical and Other Forms of Measurement Error”. In: *Journal of Business & Economic Statistics* 19, pp. 475–481.
- Lusk, Jayson L. and Kathleen Brooks (2011). “Who participates in Household Scanning Panels?” In: *American Journal of Agricultural Economics* 93, pp. 226–240.
- Zhen, Chen et al. (2009). “Understanding Differences in Self-Reported Expenditures between Household Scanner Data and Diary Survey Data: A Comparison of Homescan and Consumer Expenditure Survey”. In: *Review of Agricultural Economics* 31, pp. 470–492.

# Kapitel 3

## TOBIT REGRESSION TO ESTIMATE IMPACT OF EU MARKET INTERVENTION IN THE GERMAN BUTTER MARKET

Autoren:

Meike Wocken, Thomas Kneib

Eine erste Version des Beitrags („Tobit regression to estimate impact of EU market intervention in dairy sector“) ist als Seminarbeitrag veröffentlicht 123rd EAAE Seminar 2012, PRICE VOLATILITY AND FARM INCOME STABILISATION: Modelling Outcomes and Assessing Market and Policy Based Responses, 23.-24. Februar 2012, Dublin, Irland.

Quelle: <http://purl.umn.edu/122528>(besucht am 29.11.2012)

Tobit regression to estimate impact of EU market intervention

## Abstract

*This study examines the effect of European intervention politics in the European butter market in the context of market liberalization using the example of Germany. A heteroscedastic Tobit model is estimated using German butter market data from 1979-2012. Model specification is basing on an error correction model. It is done to account for the non-stationary nature and cointegration of the export price series (world market price plus export subsidies) and German butter price series. We get similar results to Chavas and Kim (2006) who analysed the U.S. butter market. There is evidence that price support reduces price instability in the German butter market. With the estimated model, we predict market development without lowering the intervention price since 2004. It indicates a positive effect towards the expected butter price, even though the mean market price is higher than the intervention level and no intervention purchases take place. Maximum likelihood method is conducted for estimation. For comparative purpose censored least absolute deviations estimation is applied as well. Discrepancies in results may come from a misspecified likelihood.*

## 3.1 Introduction

Common Agricultural Policy (CAP) was implemented to achieve the objectives of price stabilization, farm income support, and food security in the European Community. (European Court of Auditors, 2009, p. 12) In CAP different support mechanisms have been developed. One of the main instruments is intervention agency purchasing for various agricultural commodities, e.g. grains, sugar, butter, and skimmed milk powder (SMP). These purchases prevent market prices from falling below specified intervention prices. (Bureau of Agricultural Economics, 1985, p. 31) The dairy sector was unconcerned by the radical change in CAP by the MacSherry reform in 1992. With Agenda 2000 (Benjamin, Gohin, and Guyomard, 1999) and Mid-Term Review 2003 the reform process towards market liberalization continues and decreases intervention prices for butter and SMP. Because for other agricultural commodities support reduction already took place in the 1990s, the European dairy market reform is still in progress. Reform's impact is not yet fully conceivable.

### 3.2 German Butter Market

Basing on the work of Chavas and Kim (2006) analysing the U.S. butter market, this study examines price dynamics while lowering intervention prices using the example of the German butter market. The results of the study indicate how recent CAP reform influences butter prices. Understanding the effects of intervention politics helps to make efficient use of this instrument to manage price volatility in future. This is needed due to EU Commission's decision to retain intervention purchases and public stock holding as safety net after health check. (European Court of Auditors, 2009, p. 42) In contrast to Chavas and Kim (2006) we include world market price to explain price developments on the German butter market. Due to EU's position in international trade, supported by export subsidies, world market is a non-negligible market opportunity.

The paper is organized as follows. Section 2 contains description of the data and the German butter market. In section 3 we propose a Tobit model to analyse the butter price. Section 4 presents the results of the maximum likelihood (ML) estimation including some numeric simulations. In section 5 we predict the market development without lowering the intervention price. In the proximate section we test the specification of the error distribution in the ML estimation using censored least absolute deviations (CLAD) estimation method. The paper ends with a conclusion.

## 3.2 German Butter Market

The development of the German butter market price (EUR/kg) from July 1979 through May 2012 can be seen in Figure 3.6. From the beginning until 1991, intervention agencies in each member state bought in butter at intervention price without restrictions. In 1991 intervention is linked to market price. If market price falls below 92% of intervention price for two weeks, agencies buy butter into intervention at 90% of intervention price. (Regulation (EEC) No 1634/91) For analysis monthly purchase prices (EUR/kg) are used, which consist of intervention prices and from 1991 onwards of 90% of intervention prices. Beside the German butter stock (1000 t) the world market price (EUR/kg) is used for explanation in the Tobit model. Here, butter stock consists of stored quantities from private storage aid plus quantities from intervention purchases. EU, New Zealand and

*Tobit regression to estimate impact of EU market intervention*

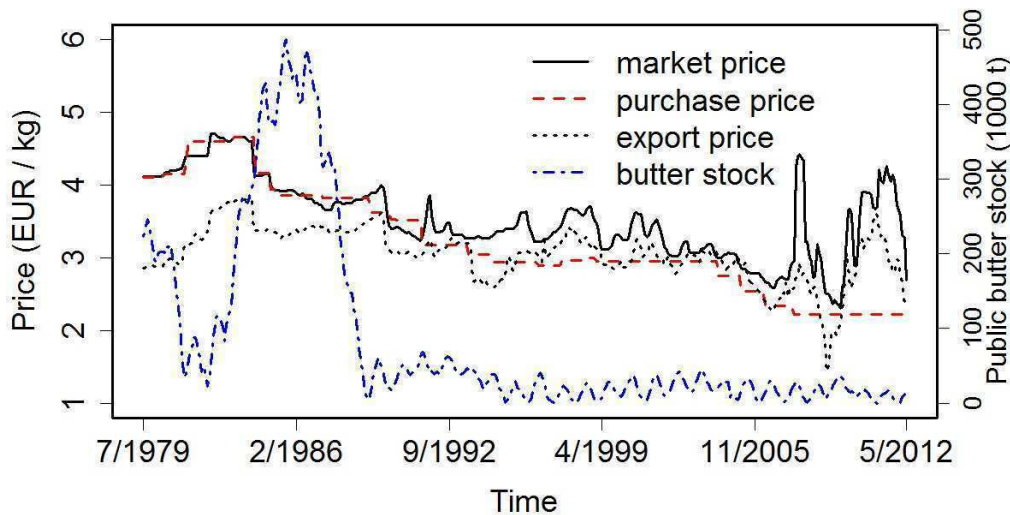


Figure 3.6: Market price, purchase price, public butter stock, and export price in German butter market July 1979 - May 2012.

Australia are the main exporters regarding butter. In the EU the Netherlands are the most important exporter. (Hockmann and Vöneki, 2009, p. 137f.) Until the beginning 1990s the Netherlands exported larger butter quantities compared to New Zealand. For this reason and due to data availability, world market price is defined to be Netherlands wholesale price minus export refund from 07/1979 to 08/1993. Data are provided by Eurostat (2012). After 08/1993 world market price is given by the Oceania export price reported by United States Department of Agriculture (2012). From German market's perspective, for further analysis the relevant price is the so called export price which is the sum of world market price and export refund. It reflects the actual perceived price in exports. The export price is additionally pictured in Figure 3.6.

Butter stock, purchase price, and German market price are accessed via the monthly reports from Federal Office for Agriculture and Food (1979-2012). Export refund is from Central Marketing and Price Reporting Unit for Products of Agriculture, Forestry and the Food Industry (ZMP) (1979-2008) and Agriculture and Board (2012). Purchase price and public butter stock are displayed in Figure 3.6 as well. It reveals two different regimes: one with high stock quantities and



### 3.3 Tobit Model

Total $n = 395$ 7/1979 - 5/2012				Before reform $n_1 = 297$ 7/1979 - 3/2003				After reform $n_2 = 98$ 4/2003 - 5/2012			
mean	sd	min	max	mean	sd	min	max	mean	sd	min	max
Market Price $p_t$											
3.54	0.53	2.3	4.7	3.67	0.45	3.01	4.7	3.16	0.58	2.3	4.42
Public butter stock $ST_t$											
86.76	125.55	0	487.7	110.38	136.73	0	487.7	15.16	9.79	0	37.2
World market price $p_t^w$											
1.68	0.57	0.86	3.61	1.5	0.39	0.86	2.49	2.2	0.71	1.28	3.61
Export refund $r_t$											
1.37	0.61	0	2.12	1.67	0.24	0.95	2.12	0.48	0.53	0	1.59

Table 3.9: Descriptive statistics including the average value (mean), standard deviation (sd), minimum value (min), and maximum value (max) in the respective data.

high purchase price (government regime), and the other with low stock quantities and low purchase price (market regime). Summary statistics of the data are presented in Table 3.9. To get a first impression of the impact of the reform, descriptive statistics are additionally calculated for the data splitted at April 2003. Comparison of statistics between before and after the reform yields some insights. Table 3.9 indicates that butter stock and export refund, on average, tend to be larger before April 2003 which might be explained by augmented policy interventions. The mean level of the German butter market price decreases after the reform whereas world market's average price level increases. Both price series face an increased volatility.

### 3.3 Tobit Model

To analyse the butter price a model is required which takes account of the fact that the dependent variable of the model is limited by intervention prices. Using Ordinary Least Squares (OLS) estimation yields biased estimates due to violation of the linearity assumption. (Amemiya, 1984) The Tobit model (3.7), suggested by Tobin (1958) accounts for left censoring at purchase price  $s_t$ . In the model

*Tobit regression to estimate impact of EU market intervention*

$p_t^*$  is a latent variable representing market outcome without censoring, whereas dependent variable  $p_t$  and regressors  $\mathbf{X}_t$  are observed for each  $t$ . Tobin (1958) assumed disturbances  $u_t$  to be normal distributed  $N(0, \sigma_t)$ , yielding the model

$$p_t^* = f(\mathbf{X}_t, \beta) + u_t \quad (3.7)$$

$$p_t = \begin{cases} p_t^*, & \text{if } p_t^* > s_t, \\ s_t, & \text{if } p_t^* \leq s_t, \end{cases} \quad t = 1, \dots, T.$$

Define the censoring indicator  $D_t = 1$  if  $p_t^* > s_t$ . The expected value and variance of the censored price variable  $p_t$  can be derived as follows:

$$E(p_t) = Prob(D_t = 0) \cdot s_t$$

$$+ Prob(D_t = 1) \cdot [f(\mathbf{X}_t, \beta) + E(u_t | u_t > s_t - f(\mathbf{X}_t, \beta))]$$

$$= [1 - \Phi(h_t)] \cdot f(\mathbf{X}_t, \beta) + \sigma_t \cdot \phi(h_t) + \Phi(h_t) \cdot s_t \quad (3.8)$$

$$Var(p_t) = \sigma_t^2 \cdot [1 - \Phi(h_t) + h_t \cdot \phi(h_t) + h_t^2 \cdot \Phi(h_t) - [h_t \cdot \Phi(h_t) + \phi(h_t)]^2], \quad (3.9)$$

where  $h_t = (s_t - f(\mathbf{X}_t, \beta))/\sigma_t$ , and  $\phi$  and  $\Phi$  are standard normal density and cumulative distribution function. (Chavas and Kim, 2006; Amemiya, 1984; Maddala, 1983; Tobin, 1958) The expected price  $p_t$  is a weighted average of  $y_t^*$  and  $s_t$ . The weights are given by the probability of government intervention.

As seen in Figure 3.6 the German butter market consist of different regimes. We presume period beginning in the mid 1990s to be a *market regime*. Purchase price declines and market price is detached from it. Before mid 1990s we suppose a *government regime*, where market price development is dominated by the change in purchase price  $s_t$ . This finding is similar to Chavas and Kim (2006), who also find such regimes on the Amercian butter market. In the *market regime*, where market price  $p_t$  is equal to the latent price  $p_t^*$ , export price  $p_t^e$  might influence market price  $p_t$ . To analyse this relation we need to be careful. Both variables are non-stationary in *market regime*. Results of unit root tests can be seen in Table 3.10. To carry out the tests, we define *market regime* to start in July 1991 due to the uncoupled market and purchase prices (see Figure 3.6). The first

### 3.3 Tobit Model

		ADF	PP	KPSS
	lag	1	5	5
$N = 251$	cv	-2.58/ - 1.95/ - 1.62	-4.00/ - 3.43/ - 3.14	0.74/0.46/0.35
German market				
	price $p_t$	-0.7256	-3.7355	0.4037
	Export price			
	$p_t^e = p_t^w + r_t$	-0.6872	-2.4791	0.9161

Table 3.10: Unit root tests in *market regime* July 1991 to May 2012. ADF: Augmented Dickey-Fuller test; PP: Phillips-Perron test (Z-tau statistic); KPSS: Kwiatkowski et al. (1992); cv: critical value at 1%, 5%, and 10% level. Source: Own calculation using the `urca`-package for the statistical software environment R.

differences of both series results in stationary time series. A regression model for non-stationary variables give spurious results in general. The problem is dealing with non-stationary disturbances. An error correction model (ECM) can be applied to overcome the problem. (Engle and Granger, 1987) To assure that the ECM is a valid framework for analysis, cointegration has to be verified. Johansen trace statistic and Johansen trace test procedure by Lütkepohl, Saikkonen, and Trenkler (2004), which allows a level shift, show a cointegrating relationship between  $p_t$  and  $p_t^e$ . Test procedure by Lütkepohl, Saikkonen, and Trenkler (2004) estimates a breakpoint at May 2007. Table 3.11 shows the results. To examine the direction of the relationship, a Granger causality test is applied. The test confirms theoretical considerations, that the export price influences German market price. Results agree with findings in Loy and Steinhagen (2009). They investigated an vector error correction model to analyse the impact of export price on market price for butter and skimmed milk powder using data from 1995 to 2004.

Using only the ECM neglects the *government regime*. The Tobit model allows regime switching. Also, the non-linearity of market price expectation should be kept in mind. Therefore, a combination of both approaches is used. With regard to Engle and Granger (1987), we start with a general ECM

*Tobit regression to estimate impact of EU market intervention*

rank	Johansen trace			Lütkepohl test		
	test statistic	5%	1%	test statistic	5%	1%
$r \leq 1$	6.65	9.24	12.97	8.63	6.79	10.04
$r = 0$	33.82	19.96	24.60	46.24	15.83	19.85

Table 3.11: Cointegration rank test in *market regime* July 1991 to May 2012: Johansen trace statistic and Johansen trace test procedure by Lütkepohl, Saikkonen, and Trenkler (2004). Source: Own calculation using the `urca`-package for the statistical software environment R.

$$\Delta p_t^* = c + \gamma \Delta p_t^e + \alpha (p_{t-1} - \theta_0 - \theta_1 p_{t-1}^e) + \sum_{i=1}^L \beta_{t-i} \Delta p_{t-i}^e + \sum_{i=1}^L \xi_{t-i} \Delta p_{t-i} + u_t, \quad (3.10)$$

where  $\Delta$  denotes the first difference operator. Equation (3.10) is equivalent to

$$p_t^* = (c - \alpha \cdot \theta_0) + \gamma \Delta p_t^e + (1 + \alpha) p_{t-1} - (\alpha \cdot \theta_1) p_{t-1}^e + \sum_{i=1}^L \beta_{t-i} \Delta p_{t-i}^e + \sum_{i=1}^L \xi_{t-i} \Delta p_{t-i} + u_t. \quad (3.11)$$

$p_t^*$  is still the latent market price which is only observable in case  $p_t^* > s_t$  (*market regime*). We define  $f(\mathbf{X}_t, \beta)$  in the Tobit model (3.7) to be equal to the right hand side of equation (3.11).  $\gamma$  captures the short-run effect between  $p_t^e$  and  $p_t^*$ .  $\alpha$  gives the speed of adjustment at which the model returns to its equilibrium.  $\theta_1$  indicates long run relationship between both variables. To test the significance of  $\hat{\alpha}$  and  $\hat{\theta}_1$  a Likelihood Ratio test is needed. For the other estimators, standard errors can be calculated from the variance-covariance matrix.

According to Tobin (1958), we assume disturbances  $u_t$  to be normal distributed  $N(0, \sigma_t)$ . Chavas and Kim (2006) detected heteroscedasticity in their data. To account for potential heteroscedasticity in our data, standard deviation  $\sigma_t$  is changing over time. Variables affecting  $\sigma_t$  are investigated to answer the question whether variance increase is due to lowering intervention prices or other factors. Following Chavas and Kim (2006), our standard deviation specification is  $\sigma_t = \exp(\gamma' \mathbf{Z}_t)$ . Beside time trend  $TT$  to control for long term changes in  $\sigma_t$  the

### 3.4 Results from Maximum Likelihood Estimation

difference of price to purchase level ( $p_{t-1} - s_{t-1}$ ) is included in  $\mathbf{Z}_t$ . If the difference is low, it is likely to correspond to the *government regime*. If the difference is high, it is likely to correspond to the *market regime*. Similar to Chavas and Kim (2006) it detects changes in market instability not captured by censoring. Also, the effect of the difference on standard deviation should vary with stock level. Therefore an interaction term ( $(p_{t-1} - s_{t-1}) \cdot ST_{t-1}$ ) is introduced to reflect latent volatility shifters. Finally, the lagged stock variable  $ST_{t-1}$  is part of  $\mathbf{Z}_t$ . The standard deviation is specified as follows

$$\sigma_t = \exp(\gamma_0 + \gamma_1 ST_{t-1} + \gamma_2(p_{t-1} - s_{t-1}) + \gamma_3((p_{t-1} - s_{t-1}) \cdot ST_{t-1}) + \gamma_4 TT). \quad (3.12)$$

## 3.4 Results from Maximum Likelihood Estimation

The ML estimator is strongly consistent and asymptotically normal if the standard assumptions hold. Disregarded heteroskedasticity and an incorrect normal assumption imply inconsistency for the ML estimator. (Arabmazar and Schmidt, 1981; Arabmazar and Schmidt, 1982; Hurd, 1979) Serial correlation appears not to be a major problem. (Amemiya, 1984) Although the specified model accounts for heteroscedasticity, there is no warranty of correct specification. The maximization is done by using Nelder-Mead algorithm implemented in the package `maxLik` for the statistical software environment R (for details see Henningsen and Toomet (2011)). To choose  $L$  (lag order in equation (3.11)), the Schwarz criterion is applied. For the estimated model specification  $L = 2$  is selected. Table 3.12 reports the ML parameter estimates of the dynamic Tobit model.

At first, the overall significance of the model is tested. The p-value of the corresponding Likelihood Ratio test statistic ( $1192.351 \sim \chi^2(11)$ ) is zero. The overall model is statistically significant. To assess the impact of omitting the long run relationship from the model another Likelihood Ratio test is conducted. To test the hypothesis  $\theta_1 = 0$  the restricted specification of  $f(\mathbf{X}_t, \beta)$  in (3.7) is

Tobit regression to estimate impact of EU market intervention

	Value	Std. error	p-value
(Intercept)	-0.0320	0.0387	0.410
$\Delta p_t^e$	0.3524	0.0385	0.000 ***
$p_{t-1}$	0.9911	0.0084	0.000 ***
$p_{t-1}^e$	0.0202	0.0196	0.301
$\Delta p_{t-1}^e$	0.0837	0.0430	0.052 .
$\Delta p_{t-2}^e$	0.0348	0.0489	0.477
$\Delta p_{t-1}$	0.3860	0.0583	0.000 ***
$\Delta p_{t-2}$	-0.0908	0.0525	0.084 .
(Intercept)	-4.2696	0.1547	0.000 ***
$ST_{t-1}$	-0.0013	0.0013	0.337
$p_{t-1} - s_{t-1}$	0.6142	0.1380	0.000 ***
$(p_{t-1} - s_{t-1}) \cdot ST_{t-1}$	0.0327	0.0070	0.000 ***
$TT$	0.0050	0.0006	0.000 ***
Log likelihood		352.8532	

Table 3.12: ML estimates for dynamic model. Asterisks indicate statistical significance.

$$p_t^* = (c - \alpha \cdot \theta_0) + \gamma \Delta p_t^e + (1 + \alpha)p_{t-1} + \beta_{t-1} \Delta p_{t-1}^e + \beta_{t-2} \Delta p_{t-2}^e + \xi_{t-1} \Delta p_{t-1} + \xi_{t-2} \Delta p_{t-2} + u_t. \quad (3.13)$$

The test yields a test statistic of  $77.753 \sim \chi^2(1)$  (p-value = 0). The long run effect is statistically significant. To test that  $\alpha$ , the speed of adjustment, is equal to zero we estimate the restricted specification of  $f(\mathbf{X}_t, \beta)$  in (3.7) as follows:

$$\begin{aligned} p_t^* &= c + \gamma \Delta p_t^e + \beta_{t-1} \Delta p_{t-1}^e + \beta_{t-2} \Delta p_{t-2}^e + \xi_{t-1} (p_{t-1} - p_{t-2}) + p_{t-1} + \xi_{t-2} \Delta p_{t-2} + u_t \\ &= c + \gamma \Delta p_t^e + \beta_{t-1} \Delta p_{t-1}^e + \beta_{t-2} \Delta p_{t-2}^e + (\xi_{t-1} + 1)p_{t-1} + (-\xi_{t-1})p_{t-2} + \xi_{t-2} \Delta p_{t-2} + u_t. \end{aligned} \quad (3.14)$$

The Likelihood-Ratio statistic is  $72.04 \sim \chi^2(1)$ . The p-value is 0. The adjustment rate is statistically significant. Next we test the assumption of heteroscedasticity. The hypothesis that standard deviation  $\sigma_t$  is constant has to be rejected ( $369.4034 \sim \chi^2(4)$ , p-value = 0). Beside the second lag of first differences of world

### 3.4 Results from Maximum Likelihood Estimation

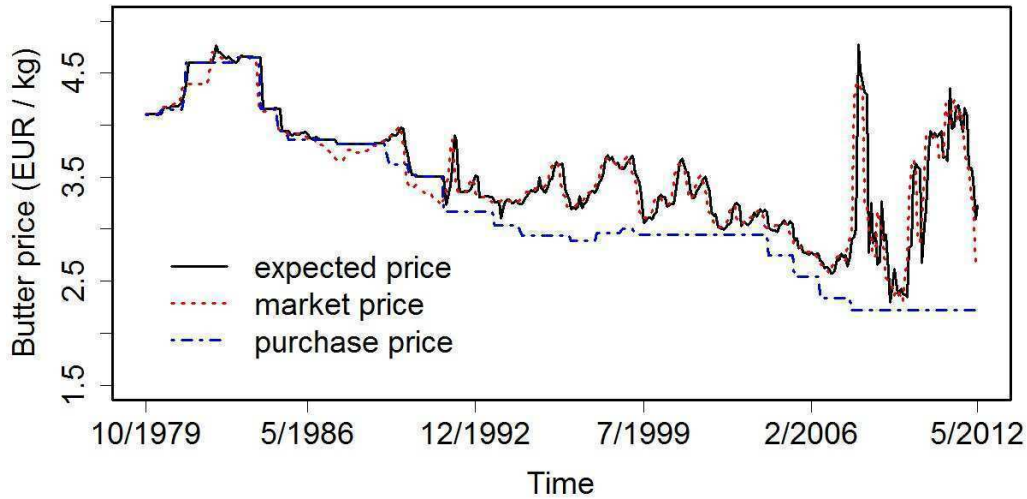


Figure 3.7: Expected price, market price, and purchase price in German butter market Oct 1979 - May 2012.

market price series and the lagged butter stock in the standard deviation, the estimated parameters are statistically different from zero at least at 10% level.

In *market regime* the speed of adjustment rate is  $\hat{\alpha} = -0.0089$ . The convergence towards the long run equilibrium seems to be long-lasting. The value is smaller compared to the result in Loy and Steinhagen (2009) ( $-0.112$ ). The estimated long run relationship between German market and export price series is positive ( $\hat{\theta}_1 = 2.2697$ ) and larger comparing to Loy and Steinhagen (2009). Here, simultaneous adjustment rate of changes in the export price ( $\hat{\gamma}$ ) is 0.3524. First and second lagged first differences of export price series have positive effects. Also, the first lag of first differences of the German market price has a positive effect, whereas the second lag has a negative effect. These results reflect price dynamics in the butter market.

Withing the estimates of the standard deviation we find a negative effect of lagged stock meaning stocks reduce price volatility. The effect of the difference between market price and purchase price on standard deviation is positive. As the difference increases (as in the *market regime*), standard deviation is estimated to increase. In contrast to Chavas and Kim (2006) the estimator  $\gamma_3$  of the interaction

Tobit regression to estimate impact of EU market intervention

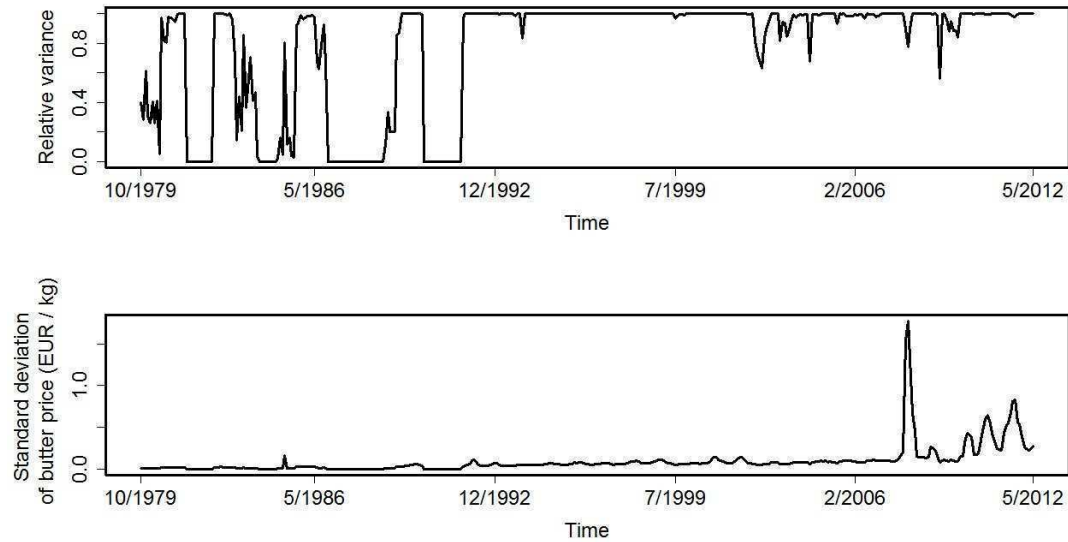


Figure 3.8: Relative variance  $Var(p_t)/\sigma_t^2$  (upper graph); Estimated standard deviation (lower graph).

term has positive sign. The stock effect to reduce volatility declines under *market regime*. This result provides support for structural changes in market due to different regimes. The time trend has a positive sign.

Performance evaluation of the model is done comparing expected price from (3.8) with observed prices (see Figure 3.7). From the beginning of the 1990s until mid 2007 the estimated model has high explanatory power. To investigate the relative role of intervention politics in the estimated variance,  $Var(p_t)/\sigma_t^2$  is derived from (3.9). In absence of censoring the relation is one, whereas it is close to zero in the presence of strong censoring effects. (Chavas and Kim, 2006) Figure 3.8 shows the results. Predicted standard deviation of butter price from (3.9) is shown in the lower graph. The smallest standard deviation is in the 1980s. This finding can be explained by the upper graph, which shows strong censoring effect in the 1980s. Also, largest standard deviation comes along with moderate censoring effect due to reducing support prices in 1990s until today.



### 3.5 Prediction of market development

		ADF	PP	KPSS
	lag	1	5	5
$N = 297$	cv	-2.58/ - 1.95	-3.45/ - 2.87	0.74/0.46
<hr/>				
Export refund				
$r_t$		-0.5977	-2.5885	0.7465
World market				
price $p_t^w$		-0.2416	-2.4572	1.7012

Table 3.13: Unit root tests July 1979 to March 2004. ADF: Augmented Dickey-Fuller test; PP: Phillips-Perron test (Z-tau statistic); KPSS: Kwiatkowski et al. (1992); cv: critical value at 1%, 5%, and 10% level. Source: Own calculation using the `urca`-package for the statistical software environment R.

## 3.5 Prediction of market development

Since mid 1990s we can observe that German market price is detached from purchase price (see Figure 3.6). Basing on Figure 3.8 we refine our first definition and presume period starting in Oct 1994 to be *market regime*. Chavas and Kim (2006) show that there is still influence of the intervention price although it is below market price and no intervention takes place, i.e. *market regime*. To be able to estimate the degree of influence, we performe a prediction of market development without reform. For this, we use the estimated Tobit model. Further, we have to forecast the development of export subsidies without the reform. This is necessary in order to be able to calculate the export price needed in the Tobit model. In accordance with Loy and Steinhagen (2009) for this an additional ECM is estimated using data until the reform (March 2004). Using historical data for butter stock and world market price we simulate the forward path of the expected price until May 2012. Results are compared to the actual development.

The ECM utilises cointegration within export subsidies and world market price. To confirm cointegration we check the time series properties. Table 3.13 shows that we cannot reject the hypothesis of non-stationarity in ADF and PP test. Additional we have to reject the null of stationarity in the KPSS test. Testing first differences results in stationarity. Hence, both variables have integration

*Tobit regression to estimate impact of EU market intervention*

Lütkepohl test			
rank	test statistic	5%	1%
$r \leq 1$	6.57	6.79	10.04
$r = 0$	26.12	15.83	19.85

Table 3.14: Johansen trace test procedure by Lütkepohl, Saikkonen, and Trenkler (2004). Estimated break point is July 1985. Source: Own calculation using the `urca`-package for the statistical software environment R.

order one. To examine cointegration Johansen trace test procedure by Lütkepohl, Saikkonen, and Trenkler (2004) is applied. It confirms cointegration (see Table 3.14).

Estimation of this auxiliary ECM is done using the Engle and Granger two-step procedure. At first, long run (equilibrium) relationship  $r_t = \beta_0 + \beta_1 \cdot p_t^w + \epsilon_t$  is estimated by OLS. Second, ECM

$$\Delta r_t = \eta_0 + \eta_1 \Delta p_t^w + \eta_2 \epsilon_{t-1} + \eta_3 TT + \eta_4 D^{BP} + \sum_{k=1}^K \eta_{t-k}^r \Delta r_{t-k} + \sum_{k=1}^K \eta_{t-k}^w \Delta p_{t-k}^w + \nu_t \quad (3.15)$$

is estimated by OLS. (Engle and Granger, 1987) Lag order  $K$  is selected by Schwarz criterion. With respect to results of the Johansen trace test an additional time trend  $TT$  is used as well as a dummy variable to indicate the estimated break point. Table 3.15 shows the result of estimation.

Comparing to Loy and Steinhagen (2009) we find long-term equilibrium  $\hat{\beta}_1 = -0.45$  smaller than  $-0.5$ . If world market price increases by one unit, export subsidies decrease by 0.45 units. It results in an export price increase by 0.55 EUR/kg. The adjustment towards the equilibrium is estimated to be only 6 percent per period. It is slower than 10 percent estimated by Loy and Steinhagen (2009).

Using the ECM we can forecast the development of export subsidies basing on world market price fluctuation. With predicted export subsidies we can calculate the expected export price after March 2004. Therefore, we only need to employ observed data for butter stock and world market price to predict market movement from the estimated Tobit model without the reform. Figure 3.9 shows the result.

### 3.5 Prediction of market development

Long run			
relationship	Value	Std. error	p-value
(Intercept)	2.3486	0.0393	0.000 ***
$p_t^w$	-0.4548	0.0254	0.000 ***
$R^2$	0.5217		
ECM			
	Value	Std. error	p-value
(Intercept)	-0.0050	0.0058	0.3857
$\Delta p_t^w$	-0.4402	0.0324	0.0000 ***
$\epsilon_{t-1}$	-0.0612	0.0179	0.0007 ***
$TT$	-0.0001	0.00005	0.0654
$D^{BP}$	0.0246	0.0104	0.0189 *
$\Delta r_{t-1}$	0.1381	0.0576	0.0172 *
$\Delta p_{t-1}^w$	0.1207	0.0410	0.0035 **
$R^2$	0.4194		

Table 3.15: Engle and Granger two-step procedure for cointegration analysis.  
Source: Own calculation.

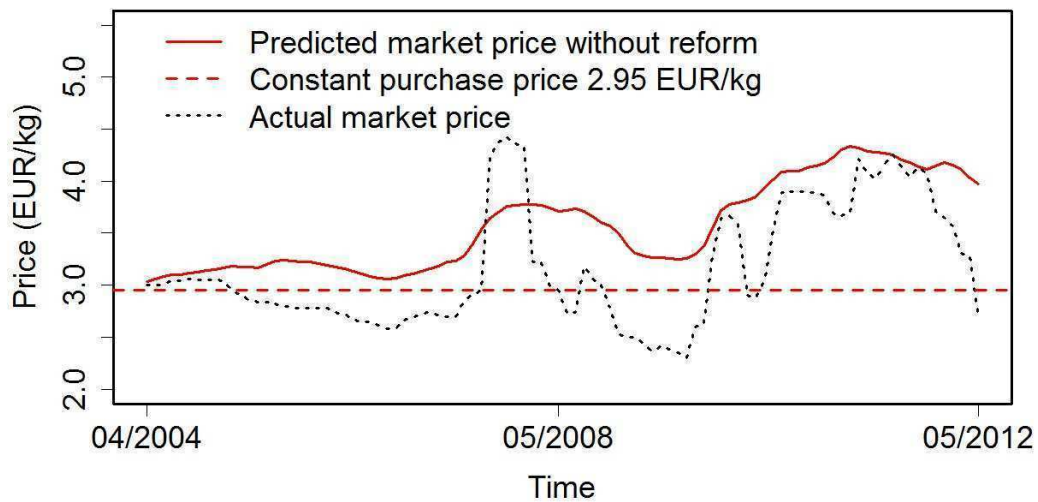


Figure 3.9: Comparing predicted market development to the actual market development.

### *Tobit regression to estimate impact of EU market intervention*

Similar to the findings in the study by Chavas and Kim (2006) the prediction indicates that the purchase price is a determining factor in price development. The expected price level is all the time above the observed price except Aug 2007 to Dec 2007 and Oct 2011. The expected price level is always higher than the assumed purchase price 2.95 EUR/kg. Consequently, no intervention takes place.

Figure 3.10 enables us to compare the impact of censoring between the situation with and without the reform. The upper graph compares the relative variance  $\sqrt{Var(p_t)}/\sigma_t$ . In case of censoring the relative variance is reduced. The reduction indicates how government intervention decreases volatility. (Chavas and Kim, 2006) Without the reform the relation is nearly one all the time. It denotes more or less an absence of censoring. This might be explained by the consideration that the intervention price can contribute to an increase in the expected market price. Through the increase in the expected market price and the constant purchase price, the probability of intervention decreases.

The lower graph in Figure 3.10 represents the estimated standard deviation with and without the reform. We can observe a reduction of standard deviation in the situation without the reform. This finding is already reflected in Figure 3.9. We observe strong peaks after the reform, but without the reform peaks are not predicted. It provides evidence that price support can reduce price instability in the German butter market.

## 3.6 Test for distributional assumption

Because the median of the observed variable  $p_t$  is not affected by censoring whereas the mean is, Powell (1984) proposed the least absolute deviations method for censored regression models. The CLAD estimator minimizes

$$\frac{1}{T} \sum_{i=1}^T |p_t - \max(s_t, \beta' \mathbf{X}_t)|. \quad (3.16)$$

The minimization estimates the median regression for  $p_t$  using the non-linear regression function  $\max(s_t, \beta' \mathbf{X}_t)$ . (Powell, 1986) CLAD method is robust to heteroscedasticity and nonnormality. (Amemiya, 1984; Powell, 1984; Powell, 1986)

### 3.6 Test for distributional assumption

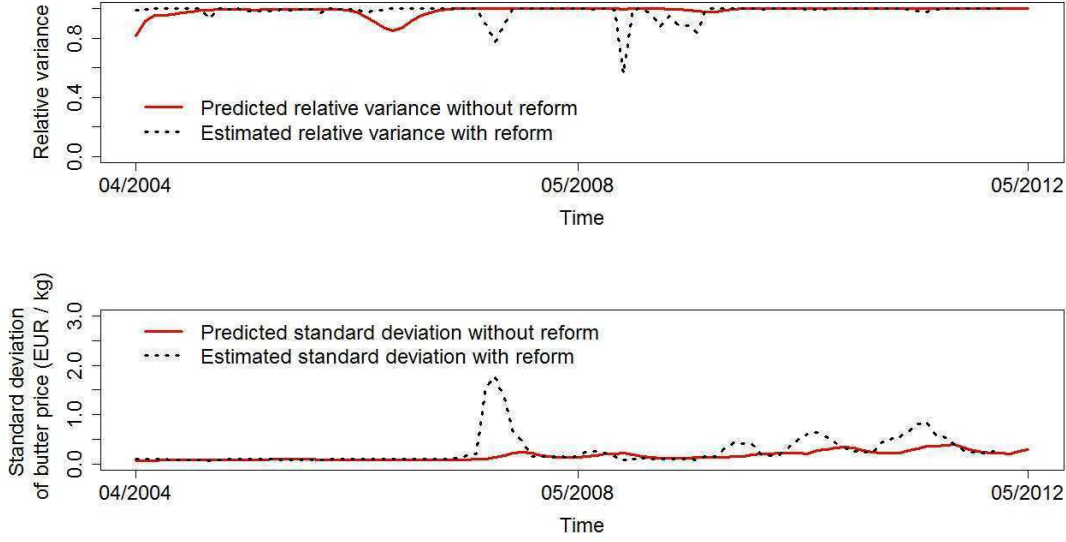


Figure 3.10: Comparing predicted values to the estimated ones: Relative variance  $Var(p_t)/\sigma_t^2$  (upper graph) and standard deviation (lower graph).

Since its robustness CLAD method is in some way the natural benchmark. (Chay and Powell, 2001) But due to distribution assumptions the ML method offers the possibility of numerical predictions. In the following the CLAD method is used to provide information on sources of misspecification in the ML estimation. Beside ML and CLAD estimation, Heckman's two-stage estimator is not considered because in large samples CLAD estimator performs better than Heckman's estimator. (Amemiya, 1984)

First estimation results in the model

$$\begin{aligned}
 p_t = & 0.0400 + 0.1994\Delta p_t^e + 0.9723p_{t-1} + 0.0185p_{t-1}^e \\
 & + 0.1168\Delta p_{t-1}^e + 0.0924\Delta p_{t-2}^e + 0.1994\Delta p_{t-1} - 0.0071\Delta p_{t-2}. \quad (3.17)
 \end{aligned}$$

Estimation is done using Powell's estimator implemented in the `quantreg`-package for the statistical software environment R. A Hausman test should be used to test, whether the differences between ML and CLAD estimates are statistically significant.

## 3.7 Conclusions

Our goal is to examine the effects of European intervention politics in the butter market. German butter market is taken to exemplify the effects. Basing on the analysis of the U.S. butter market done by Chavas and Kim (2006), we estimate a heteroscedastic Tobit model. The Tobit model accounts for censoring due to intervention politics on the market. Specification of the model originates from an ECM and considers world market price of butter to explain German market development. Monthly data for the period July 1979 through May 2012 are used. The estimated model gives some insight to the dynamics of butter prices and their changing volatility. We show how price volatility has changed. Since the 1990s the price volatility increases. It is associated with lowering support price.

Our analysis provides evidence on the price stabilization effects of intervention politics. Our prediction of market development without lowering the intervention price beginning in April 2004 illustrates that it tends to reduce price volatility. Also, a positive influence on price expectation is observable due to the constant intervention price. One interesting finding is that despite the non-decreasing intervention price the censoring effect declines. It might be explained by the decreasing probability of intervention due to the increasing expected market price.

This result is probably useable to stabilize butter markets. Keeping support price below market price, government intervention purchases are still possible in case of sudden low prices. Otherwise, as predicted, the high intervention price induces high expected market price above purchase level and no intervention purchases have to take place. From this we can conclude the same like Chavas and Kim (2006): with an increase in intervention price "*it is possible for government policy to affect long-term market behavior at a relatively low cost to taxpayers*". But also, as Chavas and Kim (2006) mentioned, the role of expectation formation in price dynamics need further research for efficient usage of such intervention purchases and intervention price.

Altogether, these results suggest that the intervention price may play a key role in stabilizing and supporting European butter market. The ML estimates of the Tobit model including distribution assumptions provide a useful framework for analysis. With regard to the stabilizing effect of price support results do

### *3.7 Conclusions*

not deviate significantly from the analysis of the American butter market. The validity of the results requires correct specifications.





# Bibliography

- Agriculture and Horticulture Development Board (2012). *DairyCo - Market Information*. Ed. by DairyCo. URL: <http://www.dairyco.org.uk/market-information/processing-trade/eu-market-support/eu-export-refunds/> (visited on 10/07/2012).
- Amemiya, Takeshi (1984). “Tobit Models: A Survey”. In: *Journal of Econometrics* 24, pp. 3–61.
- Arabmazar, Abbas and Peter Schmidt (1981). “Further evidence on the robustness of the Tobit estimator to heteroskedasticity”. In: *Journal of Econometrics* 17, pp. 253–258.
- (1982). “An investigation of the robustness of the Tobit estimator to non-normality”. In: *Econometrica* 50, pp. 1055–1063.
- Benjamin, Catherine, Alexandre Gohin, and Hervé Guyomard (1999). “The Future of European Union Dairy Policy”. In: *Canadian Journal of Agricultural Economics* 47, pp. 91–101.
- Bureau of Agricultural Economics (1985). *Agricultural policies in the European Community: Their origins, nature and effects on production and trade*. Canberra: Australian Government Publishing Service.
- Central Marketing and Price Reporting Unit for Products of Agriculture, Forestry and the Food Industry (ZMP) (1979-2008). *ZMP-Bilanz Milch*. Bonn: Zentrale Markt- und Preisberichtsstelle.
- Chavas, Jean-Paul and Kwansoo Kim (2006). “An econometric analysis of the effects of market liberalization on price dynamics and price volatility”. In: *Empirical Economics* 31, pp. 65–82.
- Chay, Kenneth Y. and James L. Powell (2001). “Semiparametrics Censored Regression Models”. In: *Journal of Economic Perspectives* 15, pp. 29–42.

## Bibliography

- Engle, Robert F. and Clive W. J. Granger (1987). “Co-integration and Error Correction: Representation, Estimation, and Testing”. In: *Econometrica* 55(2), pp. 251–276.
- European Court of Auditors (2009). *Have the management instruments applied to the market in milk and milk products achieved their main objectives? Special Report 14/2009*. Ed. by Publications Office of the European Union. URL: <http://europa.eu/rapid/pressReleasesAction.do?reference=ECA/09/63&format=HTML&aged=0&language=DE&guiLanguage=en> (visited on 09/22/2011).
- Eurostat (2012). *Statistics database*. Ed. by European Commission. URL: [ec.europa.eu/eurostat](http://ec.europa.eu/eurostat) (visited on 09/02/2012).
- Federal Office for Agriculture and Food (1979-2012). *Statistischer Monatsbericht*. Tech. rep. Federal Ministry of Food, Agriculture and Consumer Protection.
- Henningsen, Arne and Ott Toomet (2011). “maxLik: A package for maximum likelihood estimation in R”. In: *Computational Statistics* 26, pp. 443–458.
- Hockmann, Heinrich and Éva Vöneki (2009). “Market Integration and Market Power in the Internal EU Market for Butter”. In: *Traditional food production and rural sustainable development: a European Challenge*. Ashgate, Surrey: Teheresa de Noronha Vaz, Peter Nijkamp, and Jean-Louis Rastoin.
- Hurd, Michael (1979). “Estimation in truncated samples when there is heteroskedasticity”. In: *Journal of Econometrics* 11, pp. 247–258.
- Kwiatkowski, Denis et al. (1992). “Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?” In: *Journal of Econometrics* 54, pp. 159–178.
- Loy, Jens-Peter and Carsten Steinhagen (2009). “Preissenkung und Kompensation auf dem EU-Milchmarkt”. In: *Agrarwirtschaft - Zeitschrift für Betriebswirtschaft, Marktforschung und Agrarpolitik* 58 (5/6), pp. 259–268.
- Lütkepohl, Helmut, Pentti Saikkonen, and Carsten Trenkler (2004). “Testing for the Cointegrating Rank of a VAR Process with Level Shift at Unknown Time”. In: *Econometrica* 72(2), pp. 647–662.
- Maddala, Gangadharrao S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. 1st ed. Cambridge (UK): Cambridge University Press.
- Powell, James L. (1984). “Least absolute deviations estimation for the censored regression model”. In: *Journal of Econometrics* 25, pp. 303–325.

## Bibliography

- (1986). “Censored Regression Quantiles”. In: *Journal of Econometrics* 32, pp. 143–155.
- Tobin, James (1958). “Estimation of Relationships for Limited Dependent Variables”. In: *Econometrica* 26, pp. 24–36.
- United States Department of Agriculture (2012). *Dairy Market News - International Reports*. URL: <http://www.ams.usda.gov> (visited on 07/10/2012).



# Kapitel 4

## BEURTEILUNG DER KOMPENSATION AUF DEM EU-MILCHMARKT AUF BASIS EINES TOBIT-ANSATZES

Autor:

Meike Wocken

## 4.1 Einleitung

Nachdem der Milchmarkt von der McSharry-Reform der gemeinsamen Agrarpolitik im Jahr 1992 nur in geringem Ausmaß betroffen war (Salamon, 1998), begann die Liberalisierung des Milchsektors mit der Agenda 2000 und der Halbzeitbewertung 2003. Die wichtigsten Maßnahmen der Reform auf dem Milchmarkt sind (Isermeyer, 2003)

- Senkung der Interventionspreise (Butter um 25%, Magermilchpulver (MMP) um 15%),
- Kompensation der Interventionspreissenkung durch Direktzahlung (in Höhe von 3,55 ct pro kg Milchreferenzmenge),
- Erhöhung der Milchreferenzmengen (+1,5%).

Es stellt sich die Frage, ob die Zahlung der sogenannten Milchprämie ausreichend ist zur Kompensation der Interventionspreissenkungen. Eine erste Untersuchung dazu haben Loy und Steinhagen (2009) für den deutschen Markt durchgeführt. Zur Abschätzung des Kompensationsgrades wird ein Referenzpreis benötigt, der die Preisentwicklung ohne Reform prognostiziert. Loy und Steinhagen berechnen einen Referenzpreis, der aus der Summe von Weltmarktpreis und geschätzter Exporterstattung besteht, da dieser gemäß eines ersten von ihnen geschätzten Modells den Inlandspreis gut abbildet. Die Entwicklung der Exporterstattung wird dabei mittels Fehlerkorrekturmodell auf Grundlage des Zusammenhangs mit dem Weltmarktpreis vor der Reform ermittelt. Im Ergebnis wird eine deutliche Unterkompensation festgestellt. Bei der Untersuchung ist Loy und Steinhagen besonders die Dynamik des Weltmarktpreises bei der Schätzung des Referenzpreises als Einfluss wichtig. Dabei vernachlässigen sie aber den zusätzlichen Einfluss der Interventionspreise. Kim und Chavas (2002), Chavas und Kim (2004) und Chavas und Kim (2006) konnten in ihren Untersuchungen auf dem amerikanischen Milchmarkt feststellen, dass Interventionspreise eine stützende Wirkung auf die Preisentwicklung haben, auch wenn der Marktpreis oberhalb des Interventionspreises liegt und keine Intervention in Form von Aufkäufen stattfindet. Auch konnten sie zeigen, dass ein hohes Interventionsniveau zur Reduktion der Volatilität auf dem Markt führt. Für eine vollständige Bewertung der Kompensation ist die Beachtung des Interventionspreisniveaus bei der Abschätzung des Referenzpreises daher notwendig.

Die vorliegende Untersuchung greift die Frage der ausreichenden Kompensation der Interventionspreissenkung unter Berücksichtigung des Interventionspreisniveaus auf. Die Prognose des inländischen Marktpreises als Referenzpreises geschieht dabei mit einem Tobit Modell, in Anlehnung an die Untersuchungen von Kim und Chavas (2002) und

## 4.2 Entwicklung Butter- und Magermilchpulvermarkt

Chavas und Kim (2006). Als erklärende Variable des Modells wird zusätzlich der Exportpreis, die Summe aus Weltmarktpreis und Exporterstattung, in die Modellspezifikation aufgenommen. Aufgrund der varianzmindernden Wirkung des Interventionspreises ist zu vermuten, dass der hier gewählte Referenzpreis geringere Preisanstiege aufweist, als bei Loy und Steinhagen (2009) und daher die gemessene Unterkompensation geringer ausfällt.

Im Folgenden wird zunächst die Entwicklung des Butter- und Magermilchpulvermarktes in Deutschland und weltweit dargestellt. Die Zeitreiheneigenschaften der verwendeten Preisreihen werden überprüft. Anschließend wird die Methode der Berechnung des Rohstoffwertes und des Kompensationsgrades eingeführt. Abschnitt 4 erläutert das verwendete Tobit Modell und in Abschnitt 5 werden die Schätzergebnisse des Modells dargestellt. Da das Tobit Modell zur Erklärung auch die Höhe der Exporterstattung verwendet, wird diese zur späteren Prognose mit einem eigenen Modell in Abschnitt 6 abgeschätzt. Abschnitt 7 beinhaltet die Ergebnisse der Berechnung des Kompensationsgrades. Diese Ergebnisse werden im letzten Abschnitt des Beitrags diskutiert. Der Beitrag schließt mit einem Fazit.

## 4.2 Entwicklung Butter- und Magermilchpulvermarkt

Die Umsetzungen der Reformbeschlüsse der Halbzeitbewertung 2003 finden schrittweise statt. Die resultierenden Preise und die Milchprämie sind in Tabelle 4.16 dargestellt. Zusätzlich fällt die Auslöseschwelle für die Intervention bei Butter weg. Das bedeutet, dass Butteraufkäufe im Zeitraum 1. März bis 31. August eines jeden Jahres zu einem Preis von 90% des Interventionspreises getätigt werden. Daher wird analog zum Vorgehen von Loy und Steinhagen (2009) anstelle des Interventionspreises der Ankaufspreis zur weiteren Analyse verwendet. Die Interventionsmengen für Butter werden auf 70.000 t ab dem Jahr 2004 und 30.000 t ab 2008 begrenzt, sowie für MMP auf 109.000 t. (Europäische Union, 2005)

Zur Gewährleistung der Vergleichbarkeit werden die gleichen Daten und der gleiche Beobachtungszeitraum wie bei Loy und Steinhagen (2009) verwendet. Die monatlichen Marktpreise für Butter und MMP auf dem deutschen Inlandsmarkt und die Interventionspreise sind in EUR pro kg angegeben. (Bundesanstalt für Landwirtschaft und Ernährung, verschiedene Jahrgänge) Als Weltmarktpreise werden die Exportpreise (FOB) aus Ozea-

### Beurteilung der Kompensation auf dem EU-Milchmarkt

	2003/2004	2004/2005	2005/2006	2006/2007	Ab 1. Juli 2007
Butter (EUR je kg)	3,28	3,05	2,82	2,6	2,46
Magermilchpulver (EUR je kg)	2,06	1,95	1,85	1,75	1,75
Milchprämie (ct je kg Milchreferenzmenge)		1,181	2,365	2,365	3,55

Tabelle 4.16: EU Interventionspreise und Milchprämie. (Europäische Union, 2005; Loy und Steinhagen, 2009)

nien, umgerechnet in EUR per kg, verwendet. (Food and Agricultural Organisation, 2012) Die Bestände an Butter und MMP sind in 1000 t notiert. (Bundesanstalt für Landwirtschaft und Ernährung, verschiedene Jahrgänge) Einen ersten Überblick über die Entwicklung der Märkte für Butter und MMP in der Zeit von Januar 1995 bis Juli 2008 vermittelt Abbildung 4.11. Der Exportpreis ist dabei die Summe aus Weltmarktpreis und Exporterstattung (Zentrale Markt- und Preisberichtsstelle (ZMP), verschiedene Jahrgänge) in EUR pro kg.

In der Abbildung 4.11 haben die Marktpreise und die Exportpreise parallele Verläufe. Loy und Steinhagen (2009) zeigen die Kointegration der beiden Größen vor der Reform und schätzen den Zusammenhang als Fehlerkorrekturmodell. Bei der Ermittlung des Referenzpreises fokussieren sie schließlich auf den Exportpreis, da sie aufgrund ihrer Schätzergebnisse schließen, dass dieser „*längerfristig die Inlandspreise eins zu eins widerspiegelt.*“ (Loy und Steinhagen, 2009) Hingegen ist gerade auf dem Buttermarkt zu Beginn des Beobachtungszeitraums, als das Interventionspreisniveau noch nicht gesenkt war, der Marktpreis oberhalb des Exportpreises. Eine Verbesserung der Prognose der Marktentwicklung durch die Berücksichtigung des Interventionsniveaus ist daher zu erwarten.

Die Zeitreiheneigenschaften für den Zeitraum vor der Reform, die bereits von Loy und Steinhagen (2009) getestet worden sind, können durch eigene Tests bestätigt werden (Tabelle 4.17). Die ersten Differenzen ergeben für alle Zeitreihen von 1995 bis 2004 Stationarität. Es handelt sich um integrierte Preisreihen erster Ordnung. Mit einem Johansen-Trace Test wird zusätzlich die Kointegration zwischen dem deutschen Marktpreis und dem Exportpreis, sowie den Exporterstattungen und dem Weltmarktpreis jeweils für Butter und MMP in der Zeit vor der Reform bestätigt.



### 4.3 Berechnung Rohstoffwert und Kompensationsgrad

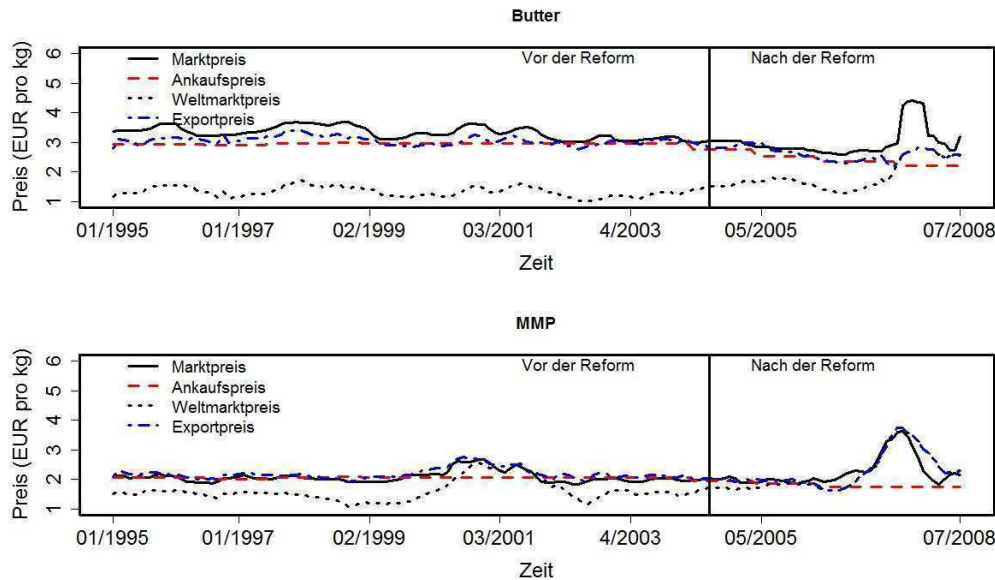


Abbildung 4.11: Entwicklung des Marktes für Butter und Magermilchpulver. (Bundesanstalt für Landwirtschaft und Ernährung, verschiedene Jahrgänge; Food and Agricultural Organisation, 2012)

### 4.3 Berechnung Rohstoffwert und Kompensationsgrad

Basierend auf den unterstellten Verarbeitungs- und Produktionskosten von 25,57 (24) ct pro kg Butter (MMP) und den technischen Koeffizienten von 22,649 (11) kg Milch je kg Butter (MMP) wird durch die Gleichung (4.18) der Schätzer für den Milchauszahlungspreis  $P$  auf Grundlage von Preisen für Butter ( $P_{Butter}$ , ct pro kg) und MMP ( $P_{MMP}$ , ct pro kg) ermittelt. (Loy und Steinhagen, 2009)

$$P = \frac{P_{Butter} - 25,57}{22,649} + \frac{P_{MMP} - 24}{11} \quad (4.18)$$

Analog zu Loy und Steinhagen (2009) wird entsprechend der reformbedingten Entwicklung der geschätzten Milchauszahlungspreise der Kompensationsgrad anhand der Größe

$$\frac{Z_D - (p^o - p^r)}{Z_D}.$$

## Beurteilung der Kompensation auf dem EU-Milchmarkt

$N = 114$		KPSS	ADF	ADF
	lag	2	1	5
Kritische Werte	1%/5%	0,74/0,46	-2,58/ - 1,95	-2,58/ - 1,95
Erstattung	Butter	0,16	-1,05	-0,44
	MMP	0,59	-1,31	-1,01
Exportpreis	Butter	0,80	-0,42	-0,06
	MMP	0,38	-0,57	-0,30
Marktpreis	Butter	1,46	-0,50	-0,54
	MMP	0,31	-0,40	-0,27
Weltmarktpreis	Butter	0,51	-0,08	-0,04
	MMP	0,53	-0,36	-0,22

Tabelle 4.17: KPSS Test (Kwiatkowski u. a., 1992) und Augmented Dickey Fuller Test (ADF) für den Zeitraum 1995 - 2004. Eigene Berechnung mit dem Statistik-Paket *urca* in der Software-Umgebung R.

ermittelt. Dabei ist  $Z_D$  die Milchprämie,  $p^o$  und  $p^r$  sind die Milchauszahlungspreise ohne und mit Reform. Aufgrund eines nur sehr geringen positiven Einkommenseffekts durch die Erhöhung der Milchreferenzmenge, wird dieser nicht weiter berücksichtigt. (Loy und Steinhagen, 2009)

## 4.4 Tobit Modellierung für Butter und Magermilchpulver

Um den Markt für Butter und den Markt für Magermilchpulver analysieren zu können, wird ein Modell benötigt, welches die Zensierung der abhängigen Variablen durch den politisch festgelegten Interventionspreis berücksichtigt. (Kim und Chavas, 2002; Chavas und Kim, 2006) Die Verwendung der Methode der kleinsten Quadrate würde die Schätzer verzerren, da die Annahme der Linearität aufgrund der Zensierung verletzt ist. (Amemiya, 1984) Das auf Tobin (1958) zurückgehende Tobit Modell (4.19) berücksichtigt die Linkszensierung durch den Interventions- bzw. Ankaufspreis  $s_t$ .

Die Variable  $p_t^*$  ist eine latente Variable in dem Modell.  $p_t^*$  repräsentiert dabei den Marktpreis, der ohne Intervention eintritt. Die abhängige Variable  $p_t$ , sowie auch die Regressoren  $\mathbf{X}_t$  sind hingegen zu jedem Zeitpunkt  $t$  beobachtbar. Tobin (1958) unterstellt, dass die Fehler  $u_t$  einer Normalverteilung  $N(0, \sigma_t)$  folgen. Es ergibt sich damit das folgende Modell

#### 4.4 Tobit Modellierung

$$\begin{aligned}
 p_t^* &= f(\mathbf{X}_t, \beta) + u_t & (4.19) \\
 p_t &= \begin{cases} p_t^*, & \text{falls } p_t^* > s_t, \\ s_t, & \text{falls } p_t^* \leq s_t, \end{cases} & t = 1, \dots, T.
 \end{aligned}$$

Die Indikatorvariable der Zensierung wird gemäß  $D_t = 1$  falls  $p_t^* > s_t$  definiert. Damit kann der Erwartungswert und die Varianz der zensierten Preisvariable  $p_t$  wie folgt berechnet werden:

$$\begin{aligned}
 E(p_t) &= Prob(D_t = 0) \cdot s_t \\
 &+ Prob(D_t = 1) \cdot [f(\mathbf{X}_t, \beta) + E(u_t | u_t > s_t - f(\mathbf{X}_t, \beta))] \\
 &= [1 - \Phi(h_t)] \cdot f(\mathbf{X}_t, \beta) + \sigma_t \cdot \phi(h_t) + \Phi(h_t) \cdot s_t & (4.20)
 \end{aligned}$$

$$Var(p_t) = \sigma_t^2 \cdot \left[ 1 - \Phi(h_t) + h_t \cdot \phi(h_t) + h_t^2 \cdot \Phi(h_t) - [h_t \cdot \Phi(h_t) + \phi(h_t)]^2 \right], \quad (4.21)$$

mit  $h_t = (s_t - f(\mathbf{X}_t, \beta))/\sigma_t$ .  $\phi$  und  $\Phi$  sind die Dichte und die kumulative Verteilungsfunktion der Standardnormalverteilung. (Chavas und Kim, 2006; Amemiya, 1984; Maddala, 1983; Tobin, 1958) Der erwartete Preis  $E(p_t)$  ist gemäß (4.20) ein gewichtetes Mittel von  $y_t^*$  und  $s_t$ . Die Gewichte resultieren aus den Wahrscheinlichkeiten eines Interventionseingriffes durch den Staat.

Wie in der Abbildung 4.11 zu sehen ist, sind auf den Märkten für Butter und MMP unterschiedliche Regime zu vermuten. Zum einen ein *Marktregime*, innerhalb dessen der beobachtbare Marktpreis sich vom Interventionspreis löst und der Interventionspreis zum Teil sinkt. Das ist vor allem auf dem Markt für Butter zu beobachten. Es gibt aber auch ein *Staatsregime*. Dort wird die Entwicklung des Marktpreises durch den staatlichen Interventionspreis dominiert, wie zum großen Teil auf dem Markt für MMP zu sehen ist. Das Tobit-Modell ermöglicht endogene Regimewechsel zwischen *Markt-* und *Staatsregimen*. (Kim und Chavas, 2002)

Aufgrund der Nicht-Stationarität der Daten wird die Idee des Fehlerkorrekturmodells zur Analyse der Beziehung der Inlandspreise zu den Exportpreisen von Loy und Steinhagen (2009) hier aufgegriffen. Ein Regressionsmodell ohne Fehlerkorrektur kann aufgrund der Nicht-Stationarität Scheinzusammenhänge der Variablen als Ergebnis aufweisen, die nicht den wahren Zusammenhängen entsprechen. Dieses Problem entsteht aufgrund der nicht-stationären Fehler des Modells. Ein Fehlerkorrekturmodell behebt dieses Problem.

## Beurteilung der Kompensation auf dem EU-Milchmarkt

Allerdings muss dafür Kointegration zwischen den Variablen gewährleistet sein. (Engle und Granger, 1987)

Ein Fehlerkorrekturmodell für die Untersuchung des Butter- bzw. MMP Marktes allein vernachlässigt ein mögliches *Staatsregime*. Innerhalb eines *Staatsregimes* ist kein Einfluss des Exportpreises auf die Marktentwicklung vorhanden. Die Preisentwicklung wird allein durch den Interventionspreis festgelegt. Da das Tobit Modell einen Regimewechsel erlaubt, wird versucht, eine Kombination aus beiden Ansätzen zu entwickeln. Ein allgemeines Fehlerkorrekturmodell nach Engle und Granger (1987) für den Zusammenhang des Inlandspreises  $p_t^*$  und des Exportpreises  $p_t^e$  lautet wie folgt

$$\Delta p_t^* = c + \gamma \Delta p_t^e + \alpha (p_{t-1} - \theta_0 - \theta_1 p_{t-1}^e) + \sum_{i=1}^L \beta_{t-i} \Delta p_{t-i}^e + \sum_{i=1}^L \xi_{t-i} \Delta p_{t-i} + u_t, \quad (4.22)$$

wobei  $\Delta$  den Differenzen Operator darstellt. Die Gleichung (4.22) ist äquivalent zu

$$p_t^* = (c - \alpha \cdot \theta_0) + \gamma \Delta p_t^e + (1 + \alpha) p_{t-1} - (\alpha \cdot \theta_1) p_{t-1}^e + \sum_{i=1}^L \beta_{t-i} \Delta p_{t-i}^e + \sum_{i=1}^L \xi_{t-i} \Delta p_{t-i} + u_t. \quad (4.23)$$

$p_t^*$  ist dabei weiterhin der latente Marktpreis, der nur beobachtbar ist, wenn  $p_t^* > s_t$  (*Marktregime*) gilt. Die Modellspezifikation  $f(\mathbf{X}_t, \beta)$  im Tobit Modell (4.19) wird nun gleichgesetzt der rechten Seite der Gleichung (4.23).  $\gamma$  gibt dabei den kurzfristigen Zusammenhang zwischen  $p_t^e$  und  $p_t^*$  an.  $\alpha$  gibt die Anpassungsrate an, mit der das Modell in sein langfristiges Gleichgewicht zurückkehrt.  $\theta_1$  beschreibt die langfristige Beziehung der beiden Variablen.

Gemäß Tobin (1958) wird angenommen, dass die Störterme  $u_t$  einer Normalverteilung  $N(0, \sigma_t)$  folgen. Kim und Chavas (2002) und Chavas und Kim (2006) entdecken bei ihren Untersuchungen der amerikanischen Märkte für Butter und fettfreier Milchtrockenmasse Heteroskedastizität in den Daten. Zur Berücksichtigung von Heteroskedastizität in den Daten, variiert die Standardabweichung  $\sigma_t$  über die Zeit. Variablen, die einen Einfluss auf  $\sigma_t$  haben können, werden untersucht. Dem Vorgehen von Kim und Chavas (2002) und Chavas und Kim (2006) folgend, ist die Spezifikation der Standardabweichung  $\sigma_t = \exp(\gamma' \mathbf{Z}_t)$ . Neben einem Zeittrend  $TT$  für langfristige Veränderungen in  $\sigma_t$  wird auch die Differenz des Marktpreises zum Interventionspreis ( $p_{t-1} - s_{t-1}$ ) in  $\mathbf{Z}_t$  aufgenommen. Eine geringe Differenz weist dabei auf ein *Staatsregime* hin. Eine große Differenz geht sehr wahrscheinlich mit einem *Marktregime* einher. (Chavas und Kim, 2006) Zusätzlich

## 4.5 Ergebnisse Tobit Modell

wird, wie bei Chavas und Kim (2006), der Einfluss der öffentlichen Lagermengen auf die Standardabweichung untersucht. Der Einfluss soll dabei variieren dürfen gemäß dem jeweiligen Regime. Daher wird der Interaktionsterm  $((p_{t-1} - s_{t-1}) \cdot ST_{t-1})$  in die Modellierung aufgenommen, um Volatilitätsveränderungen aufzuspüren. Schließlich ist der zeitlich verzögerte Lagerbestand  $ST_{t-1}$  auch noch Teil von  $\mathbf{Z}_t$ . In der Theorie führt ein hoher Lagerbestand in der Vorperiode zur Reduktion der Preisschwankungen. (Wright und Williams, 1982) Damit lautet die Spezifizierung der Standardabweichung

$$\sigma_t = \exp(\gamma_0 + \gamma_1 ST_{t-1} + \gamma_2(p_{t-1} - s_{t-1}) + \gamma_3((p_{t-1} - s_{t-1}) \cdot ST_{t-1}) + \gamma_4 TT). \quad (4.24)$$

## 4.5 Ergebnisse Tobit Modell

Um die Vergleichbarkeit der Ergebnisse zu gewährleisten, werden die gleichen Beobachtungszeiträume für die Schätzung verwendet wie bei Loy und Steinhagen (2009). Es werden die Beobachtungen Januar 1995 bis Juni 2004 vor der Reform ( $n_1 = 114$ ) für die Schätzung des Tobit Modells genutzt. Der beobachtete Weltmarktpreis wird nach der Reform bis Juni 2008 ( $n_2 = 48$ ) zur späteren Preisprognose verwendet. Die Modelle werden für Butter und MMP einzeln geschätzt. Eine multivariate Modellformulierung wird nicht angewendet, da die Annahme unterstellt wird, dass keine gegenseitige Beeinflussung der Märkte existiert aufgrund der nicht vorhandenen Rohstoffrivalität.<sup>2</sup> Für die Schätzung wird die Maximum Likelihood Methode genutzt. Zur Maximierung wird der Nelder Mead Algorithmus (implementiert im Software Paket `maxLik` für die Software Umgebung `R`) herangezogen. Zur Wahl von  $L$  in Gleichung (4.23) wird das Schwarz Kriterium verwendet. Auf dem MMP Markt wird demnach  $L = 3$  gewählt und auf dem Buttermarkt  $L = 1$ . In Tabelle 4.18 sind die Ergebnisse für den Buttermarkt dargestellt. Die Ergebnisse für den MMP Markt finden sich in Tabelle 4.19.

Mittels Likelihood-Ratio Tests wird die statistische Gesamtsignifikanz der beiden Modelle überprüft. Sowohl das Modell für den Butterpreis ( $371,837 \sim \chi^2(9)$ , p-Wert = 0) ist statistisch gesamtsignifikant, als auch das Modell für den MMP Preis ( $258,312 \sim \chi^2(13)$ ,

---

<sup>2</sup>Rohmilch besteht aus Wasser, Laktose, Milhfett und Eiweiß. Das Milhfett wird hauptsächlich zur Herstellung von Butter verwendet. Aus dem Milcheiweiß wird Magermilch erzeugt. Magermilch dient der Produktion von Magermilchpulver. (European Court of Auditors, 2009, S. 10)

Beurteilung der Kompensation auf dem EU-Milchmarkt

	Schätzer	Std. fehler	p-Wert
(Intercept)	0,125	0,094	0,184
$\Delta p_t^e$	0,115	0,049	0,018 *
$p_{t-1}$	0,909	0,032	0,000 ***
$p_{t-1}^e$	0,060	0,047	0,209
$\Delta p_{t-1}^e$	-0,002	0,049	0,964
$\Delta p_{t-1}$	0,651	0,077	0,000 ***
(Intercept)	-4,202	0,467	0,000 ***
$ST_{t-1}$	-0,005	0,013	0,729
$p_{t-1} - s_{t-1}$	1,734	0,822	0,035 *
$(p_{t-1} - s_{t-1}) \cdot ST_{t-1}$	-0,024	0,031	0,448
$TT$	0,010	0,003	0,001 **
Log likelihood		202,689	

Tabelle 4.18: Geschätzte Parameter des Tobit Modells für den Buttermarkt. Die Asteriske geben statistische Signifikanz an (\*  $p < .05$ , \*\*  $p < .01$  und \*\*\*  $p < .001$ ).

	Schätzer	Std. fehler	p-Wert
(Intercept)	-0,018	0,077	0,815
$\Delta p_t^e$	0,554	0,092	0,000 ***
$p_{t-1}$	0,544	0,165	0,001 ***
$p_{t-1}^e$	0,444	0,171	0,010 **
$\Delta p_{t-1}^e$	-0,312	0,164	0,057
$\Delta p_{t-2}^e$	0,021	0,151	0,892
$\Delta p_{t-3}^e$	-0,167	0,133	0,209
$\Delta p_{t-1}$	0,615	0,157	0,000 ***
$\Delta p_{t-2}$	-0,211	0,138	0,125
$\Delta p_{t-3}$	0,359	0,131	0,006 **
(Intercept)	-3,437	0,227	0,000 ***
$ST_{t-1}$	-0,025	0,009	0,009 **
$p_{t-1} - s_{t-1}$	-1,197	0,931	0,199
$(p_{t-1} - s_{t-1}) \cdot ST_{t-1}$	0,133	0,063	0,034 *
$TT$	0,013	0,004	0,003 **
Log likelihood		86,855	

Tabelle 4.19: Geschätzte Parameter des Tobit Modells für den MMP Markt. Die Asteriske geben statistische Signifikanz an (\*  $p < .05$ , \*\*  $p < .01$  und \*\*\*  $p < .001$ ).

p-Wert = 0). Die von Loy und Steinhagen (2009) genutzte langfristige 1:1 Beziehung zwischen Inlands- und Exportpreis kann auf den beiden Märkten gemäß Likelihood-Ratio Test der Hypothese  $\theta_1 = 1$  nicht abgelehnt werden. Auf dem MMP Markt liegt der Schätzer der langfristigen Beziehung bei  $\hat{\theta}_1 = 0,974$  ( $0,616 \sim \chi^2(1)$ , p-Wert = 0,433). Auf dem Buttermarkt ist der Schätzer für den langfristigen Zusammenhang deutlich kleiner Eins ( $\hat{\theta}_1 = 0,659$ ;  $0,174 \sim \chi^2(1)$ , p-Wert = 0,677).

## 4.6 Prognose Exporterstattung

Um Prognosen eines Referenzpreises mit dem geschätzten Tobit-Modell zu ermöglichen, wird ein Instrument zur Vorhersage der Exporterstattung benötigt, mit dem der vom Tobit-Modell verwendete Exportpreis berechnet werden kann. Anlog zu dem Vorgehen von Loy und Steinhagen (2009) wird ein Fehlerkorrekturmodell dafür verwendet. Dieses nutzt die Kointegrationsbeziehung zwischen der Exportunterstützung  $r_T$  und dem Weltmarktpreis  $p_t^w$ . Die Schätzung dieses Hilfsmodells geschieht mit dem einfach umzusetzenden Engle and Granger 2-Schritt Verfahren. (Engle und Granger, 1987) Dafür ist zuerst die langfristige Gleichgewichtsbeziehung  $r_t = \beta_0 + \beta_1 \cdot p_t^w + \epsilon_t$  der zwei Größen mit der Methode der Kleinsten Quadrate zu schätzen. In dem zweiten Schritt wird das Fehlerkorrekturmodell

$$\Delta r_t = \eta_0 + \eta_1 \Delta p_t^w + \eta_2 \epsilon_{t-1} + \sum_{k=1}^K \eta_{t-k}^r \Delta r_{t-k} + \sum_{k=1}^K \eta_{t-k}^w \Delta p_{t-k}^w + \nu_t \quad (4.25)$$

mit der Methode der Kleinsten Quadrate geschätzt. (Engle und Granger, 1987) Die Lag Ordnung  $K$  wird mittels Bayes Informations Kriterium gewählt. Für beide Märkte resultiert die Ordnung  $K = 1$ . In Tabelle 4.20 sind die Ergebnisse zu sehen.

Die Geschwindigkeit beim Abbau von Ungleichgewichten ist für Butter mit  $-0,086$  geringer als bei Loy und Steinhagen ( $-0,112$ ). Dieses Ergebnis ist mit  $-0,153$  auch für MMP zu finden ( $-0,366$ ). Die Unterschiede scheinen daraus zu resultieren, dass Loy und Steinhagen simultan die Zusammenhänge mittels Vektor-Fehlerkorrekturmodell geschätzt haben.

## Beurteilung der Kompensation auf dem EU-Milchmarkt

	Buttermarkt			MMP Markt		
Langfristige Beziehung	Schätzer	Std. fehler	p-Wert	Schätzer	Std. fehler	p-Wert
(Intercept)	2,195	0,067	0,000***	1,514	0,047	0,000***
$p_t^w$	-0,354	0,050	0,000***	-0,585	0,029	0,000***
$R^2$	0,3107			0,7897		
ECM	Schätzer	Std. fehler	p-Wert	Schätzer	Std. fehler	p-Wert
(Intercept)	-0,002	0,003	0,437	-0,002	0,003	0,585
$\Delta p_t^w$	-0,116	0,047	0,014*	-0,296	0,059	0,000 ***
$\epsilon_{t-1}$	-0,086	0,038	0,025*	-0,153	0,043	0,000 ***
$\Delta r_{t-1}$	0,352	0,078	0,000***	0,383	0,081	0,000 ***
$\Delta p_{t-1}^w$	-0,075	0,048	0,122	0,074	0,074	0,321
$R^2$	0,2883			0,5731		

Tabelle 4.20: Engle and Granger 2-Schritt Verfahren zur Schätzung des Zusammenhangs von Weltmarktpreis und Exporterstattung.

## 4.7 Bestimmung der Kompensation

Um den Kompensationsgrad durch die Milchprämie abschätzen zu können, wird im Folgenden auf Basis der beobachteten Weltmarktpreise ab Juli 2004 zuerst mithilfe der Fehlerkorrekturmodelle die Höhe der jeweiligen Exporterstattungen unter Annahme der Fortführung der Marktpolitik vor der Reform geschätzt. Diese Werte werden anschließend genutzt, um die jeweiligen Exportpreise zu ermitteln und mit diesen dann mittels geschätzter Tobit Modelle die Entwicklung der deutschen Marktpreise abschätzen zu können. Auch im Tobit Modell wird die Fortführung der Marktpolitik vor Juli 2004 simuliert, d.h. der Ankaufs- bzw. Interventionspreis wird auf dem Niveau vor der Reform gehalten.

Bei den geschätzten Exporterstattungen für MMP treten zum Teil negative Werte auf. Diese werden auf Null gesetzt, da Exportsteuern nicht Teil der Marktpolitik vor Juli 2004 waren. Auch werden die Exportpreise über die Exporterstattung, so weit möglich, angepasst, dass sie die Maximalwerte vor der Reform nicht übersteigen. Dies geschieht analog zu dem Vorgehen bei Loy und Steinhagen (2009).

Abbildung 4.12 zeigt die tatsächlichen und geschätzten Rohstoffwerte berechnet mit der Formel (4.18). Exemplarisch für die EU werden die Rohstoffwerte mit den beobachteten und geschätzten Preisen auf dem deutschen Markt ermittelt. Daher werden diese Größen auch „Rohstoffwert EU-Markt“ bzw. „Rohstoffwert EU-Markt geschätzt“ genannt. In



#### 4.7 Bestimmung der Kompensation

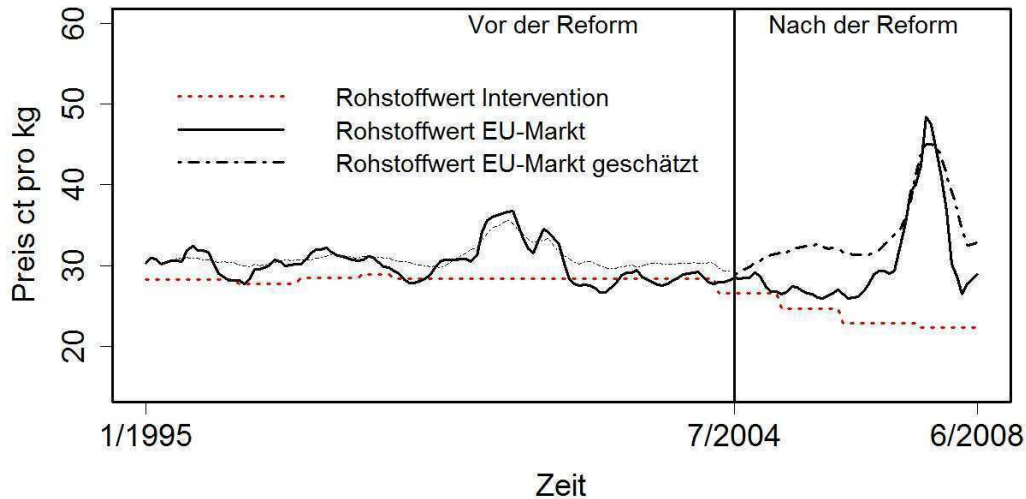


Abbildung 4.12: Beobachtete und geschätzte Rohstoffwerte auf dem deutschen Markt (exemplarisch für EU Markt).

		2004/05	2005/06	2006/07	2007/08	2004/08
Milchprämie ct/kg	$Z_D$	1,09	2,18	3,27	3,27	2,45
Rohstoffwert ct/kg	$p^o - p^r$	2,87	5,62	2,74	3,48	3,68
	$[Z_D - (p^o - p^r)] / Z_D$	-163%	-158%	16%	-6%	-50%

Tabelle 4.21: Ermittlung des Kompensationsgrads auf Basis von Rohstoffwerten EU Markt.

einem letzten Schritt werden die ermittelten Werte der gezahlten Milchprämie gegenübergestellt. Die Ergebnisse in Tabelle 4.21 zeigen, dass über den gesamten Zeitraum 2004 bis 2008 keine Überkompensation vorliegt. Die Berücksichtigung der Weltmarktentwicklungen bei der Abschätzung des Referenzpreises machen deutlich, dass zwar im Wirtschaftsjahr 2006/07 die Differenz der mittleren Rohstoffwerte am geringsten war und es daher auch zu einer Überkompensation kam, dennoch insgesamt im Durchschnitt eine Differenz zwischen den Rohstoffwerten von 3,68 ct pro kg vorliegt. Um die Senkung der Interventionspreise vollständig zu kompensieren, müsste die Milchprämie um 50% gesteigert werden. Der Grad der Unterkompensation fällt im Vergleich zu den Ergebnissen von Loy und Steinhagen (2009) nicht so stark aus. Die Tendenz ist aber ähnlich.

## 4.8 Diskussion der Ergebnisse und Fazit

Wie anfangs vermutet, wird eine geringere Unterkompensation festgestellt als bei Loy und Steinhagen (2009). Der als Referenzpreis prognostizierte Rohstoffwert für den inländischen Markt weist durch die Wirkung der Intervention gedämpfte Preisausschläge auf. Die Ergebnisse von Chavas und Kim (2004, 2006) lassen sich demnach auf den deutschen Markt und damit auch auf den europäischen Milchmarkt übertragen.

Die Wirkung der Interventionspolitik ist aber differenzierter zu betrachten. Neben der Reduktion der Volatilität haben Chavas und Kim gezeigt, dass der Interventionspreis das Preisniveau positiv beeinflusst, auch wenn der Marktpreis oberhalb des Interventionspreises liegt. Daher ist zu vermuten, dass im Fall von stark abfallenden Weltmarktpreisen, das deutsche Preisniveau auf dem Milchmarkt insgesamt nicht so stark abgefallen wäre, wie der auf dem Weltmarktpreis basierende Exportpreis. Die Interventionspolitik kann als „Sicherheitsnetz“, wie vom Rat der EU-Agrarminister nach dem Gesundheitscheck 2008 beschlossen (European Court of Auditors, 2009, S. 42), benutzt werden. Das bedeutet, dass zur Abmessung des Kompensationsgrades eine Betrachtung der Entwicklung beider Referenzpreise notwendig ist. Abhängig von der Entwicklung auf dem Weltmarkt ist die vorteilhaftere Absatzmöglichkeit zu wählen. Daher sind die Ergebnisse zusammenfassend dahingehend zu interpretieren, dass die Unterkompensation im Mittel zwischen  $-50\%$  und  $-103\%$  liegt.

Die Anwendbarkeit des hier verwendeten Tobit Modells ist beschränkt. Das Modell besitzt eine Schranke, unterhalb derer der modellierte Preis niemals fallen kann. Gemäß der Theorie wird diese Schranke durch die Interventionspreise gebildet. Es gibt allerdings Phasen auf dem deutschen Markt, in denen der Marktpreis unterhalb des Interventionspreises liegt. Gründe dafür können zum einen die Beschränkungen der Interventionsmenge und des Zeitraums des Aufkaufs, zum anderen auch die zeitliche Verzögerung zwischen Preisabfall und Aufkäufen sein. Diese Situationen sind mit dem Modell nicht abbildbar. Zusätzlich ist zur Verbesserung des Modells der Beobachtungszeitraum vor der Reform auszuweiten. Speziell auf dem Buttermarkt ist seit Anfang der 1990er Jahre der Marktpreis zum Teil deutlich höher als der Ankaufspreis der Intervention. Die Beobachtungen eines *Staatsregimes* gehen daher nicht in die Modellschätzung ein. Ein weitere Möglichkeit zur Modellverbesserung wäre die multivariate Modellierung zur Berücksichtigung möglicher gegenseitiger Einflüsse von Butter- und MMP Markt.

Schließlich hat diese Untersuchung gezeigt, dass auch die Wahl eines anderen Referenzpreises, als der von Loy und Steinhagen (2009) verwendete, das Ergebnis einer im

#### 4.8 *Diskussion der Ergebnisse und Fazit*

Zeitablauf schwankenden Kompensation hat. Es ist zu überlegen, die Kompensationszahlung variabel zu gestalten, um die Kompensation von Einkommensrückgängen der jeweiligen Marktsituation anzupassen.



# Literatur

- Amemiya, Takeshi (1984). »Tobit Models: A Survey«. In: *Journal of Econometrics* 24, S. 3–61.
- Bundesanstalt für Landwirtschaft und Ernährung (verschiedene Jahrgänge). *Statistischer Monatsbericht*. Techn. Ber. Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz.
- Chavas, Jean-Paul und Kwansoo Kim (2004). »A heteroskedastic multivariate Tobit analysis of price dynamics in the presence of price floors«. In: *American Journal of Agricultural Economics* 86, S. 576–593.
- (2006). »An econometric analysis of the effects of market liberalization on price dynamics and price volatility«. In: *Empirical Economics* 31, S. 65–82.
- Engle, Robert F. und Clive W. J. Granger (1987). »Co-integration and Error Correction: Representation, Estimation, and Testing«. In: *Econometrica* 55(2), S. 251–276.
- European Court of Auditors (2009). *Have the management instruments applied to the market in milk and milk products achieved their main objectives? Special Report 14/2009*. Hrsg. von Publications Office of the European Union. URL: <http://europa.eu/rapid/pressReleasesAction.do?reference=ECA/09/63&format=HTML&aged=0&language=DE&guiLanguage=en> (besucht am 22.09.2011).
- Europäische Union (2005). »Der Milchsektor«. In: *Die GAP-Reform 2003*. Luxemburg: Amt für Amtliche Veröffentlichungen der Europäischen Gemeinschaft.
- Food and Agricultural Organisation (2012). *International Commodity Prices*. Hrsg. von FAOSTAT. URL: <http://www.fao.org/economic/est/prices?lang=en> (besucht am 10.07.2012).
- Isermeyer, Folkhard (2003). *Umsetzung des Luxemburger Beschlusses zur EU Agrarreform in Deutschland : eine erste Einschätzung*. Braunschweig: Bundesforschungsanstalt für Landwirtschaft (FAL).

## Literatur

- Kim, Kwansoo und Jean-Paul Chavas (2002). »A dynamic analysis of the effects of a price support program on price dynamics and price volatility«. In: *Journal of Agricultural and Resource Economics* 27, S. 495–514.
- Kwiatkowski, Denis u. a. (1992). »Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?«. In: *Journal of Econometrics* 54, S. 159–178.
- Loy, Jens-Peter und Carsten Steinhagen (2009). »Preissenkung und Kompensation auf dem EU-Milchmarkt«. In: *Agrarwirtschaft - Zeitschrift für Betriebswirtschaft, Marktforschung und Agrarpolitik* 58 (5/6), S. 259–268.
- Maddala, Gangadharrao S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. 1. Aufl. Cambridge (UK): Cambridge University Press.
- Salamon, Petra (1998). »Überlegungen zur Neugestaltung der Milchmarktpolitik der EU«. In: *Agrarwirtschaft* 47, S. 313–320.
- Tobin, James (1958). »Estimation of Relationships for Limited Dependent Variables«. In: *Econometrica* 26, S. 24–36.
- Wright, Brian D. und Jeffrey C. Williams (1982). »The Economic Role of Commodity Storage«. In: *The Economic Journal* 92, S. 596–614.
- Zentrale Markt- und Preisberichtsstelle (ZMP) (verschiedene Jahrgänge). *ZMP-Bilanz Milch*. Bonn: Zentrale Markt- und Preisberichtsstelle.

# Kapitel 5

## EINSATZ EINES PROZESSMANAGEMENTZYKLUS ZUR OPTIMIERUNG VON ELEARNING: PRAXISBEISPIEL „ELEARNING QMM“

Autor:

Meike Wocken

Der Beitrag wird erscheinen in  
elead (e-learning and education) 9

Vorschau: <http://elead.campussource.de/archive/9>  
(besucht am 29.11.2012)

## Zusammenfassung

*Der Projektbericht stellt die Methode des Prozessmanagementzyklus zur Durchführung von eLearning in der universitären Lehre dar. Die Anwendung wird anhand des Praxisprojektes „eLearning QMM“ skizziert. Aufwand und Nutzen der Methode werden kurz diskutiert.*

### 5.1 Einleitung

Zur Schaffung einer zielgerichteten eLearning-Umgebung für eine Lehrveranstaltung, muss der eLearning-Einsatz geplant, entwickelt, umgesetzt und verbessert werden. Zur Unterstützung dieses Prozesses ist die Verwendung neuerer Managementmethoden sinnvoll (Fabac, 2006). Im Folgenden wird ein an den zyklischen eLearning-Prozess von Ćukušić u. a. (2010) angelehntes Prozessmanagement vorgeschlagen. Der praktische Einsatz dieser Managementmethode wird anhand eines Praxisprojektes skizziert. Es werden abschließend Aufwand und Nutzen gegenübergestellt, sowie Hinweise auf weitere Informationen zur Durchführung eines Prozessmanagements im eLearning gegeben.

### 5.2 Handlungsschritte Prozessmanagement im eLearning

Die Transformation von Inputs, u.a. die Bedürfnisse der Studenten und Dozenten, der zu vermittelnde Lehrstoff, Materialien, Hard- und Software, in den Output eLearning-Umgebung ist ein Prozess (vgl. Anupindi u. a. (2012, S. 3), welcher mit den Methoden des Prozessmanagements unterstützt werden sollte. Das Prozessmanagement regelt die Ausführung und Koordination der einzelnen Schritte des Prozesses (Becker und Kahn, 2010).

Abbildung 5.13 stellt den Prozesszyklus des eLearning-Einsatzes dar. Zu Beginn ist in der Planungsphase das Ziel des eLearning Einsatzes zu definieren. Aus der Zielsetzung heraus sind kritische Erfolgsfaktoren abzuleiten, die den Zielerreichungsgrad während der anschließenden Umsetzung und der Evaluation messen. Nach der Festsetzung kritischer Erfolgsfaktoren ist die Zielgruppe der Veranstal-



## 5.2 Handlungsschritte Prozessmanagement

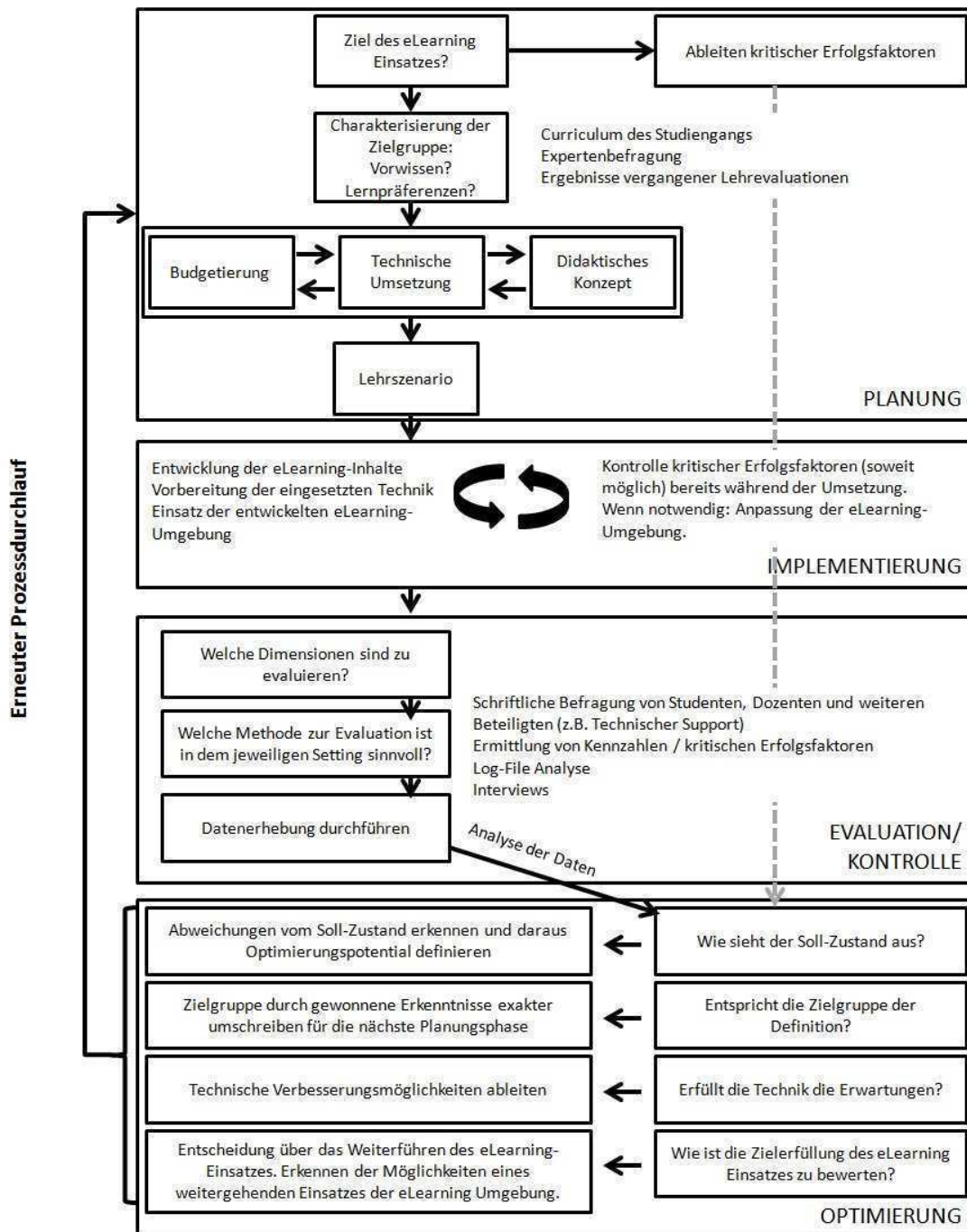


Abbildung 5.13: Schematische Darstellung Prozessmanagementzyklus. Eigene Ausarbeitung in Anlehnung an Ćukušić u. a. (2010).

tung bezüglich Vorwissens, Heterogenität, Lernpräferenzen und Motivation zu charakterisieren. Dies kann zum Teil durch das Curriculum des Studiengangs geschehen, zum Teil können die Ergebnisse vergangener Lehrevaluationen und Erfahrungen der Dozenten herangezogen werden. Mit Berücksichtigung der festgesetzten Ziele, der Zielgruppe und des zur Verfügung stehenden Budgets jetzt und in Zukunft (Gewährleistung der Fortführung der technischen Umgebung) sind der didaktische Rahmen sowie die technische Umsetzung auszuwählen. Eine Übersicht zu Evaluationshilfen für die Auswahl der passenden eLearning-Software und Systeme findet sich bei Hoppe und Breitner (2006).

Die anschließende Umsetzung und Durchführung der eLearning-Veranstaltung findet gemäß der vorhergehenden Planung statt. Die kritischen Erfolgsfaktoren sind im Blick zu behalten, um bei möglichen Fehlentwicklungen frühzeitig korrigierend eingreifen zu können. Eine umfassende Evaluation findet im Anschluss an die Umsetzung statt. Da die Kosten der Evaluation im Verhältnis zum Nutzen stehen müssen, sind nicht alle Dimensionen, die zum Erfolg von eLearning beitragen, zwingend zu evaluieren. Für einen Überblick dieser Dimensionen siehe Golzari u. a. (2011). Allerdings sind genügend Informationen und Daten zu erheben, um den Zielerreichungsgrad der in der Planung definierten Ziele zu messen. Die Analyse der Daten findet im letzten Schritt des Prozesses statt, der Optimierungsphase. Anhand der Abweichungen von dem definierten Soll-Zustand werden Optimierungspotentiale aufgedeckt. Basierend auf den gesamten Ergebnissen der Evaluation ist die Fortführung des eLearning-Einsatzes kritisch abzuwägen. Fällt die Evaluation positiv aus, so fließen Optimierungspotentiale, aber auch neue Erkenntnisse und Verbesserungsmöglichkeiten bezüglich der Zielgruppe und der technischen Umsetzung beim nächsten Einsatz der eLearning-Umgebung in die Planungsphase mit ein.

### **5.3 Praxisbeispiel „eLearning QMM“**

Der Einsatz des vorgestellten Projektmanagementzyklus wird anhand des Projektes „eLearning QMM“ verdeutlicht, welches im Rahmen der Förderung der nachhaltigen Integration von eLearning in der Präsenzlehre an der Christian-Albrechts-Universität zu Kiel (CAU) in den Sommersemestern 2010 und 2011

### 5.3 Praxisbeispiel „eLearning QMM“

durchgeführt wurde. Dabei wurde eLearning zur Unterstützung des Moduls „Quantitative Methoden der Marktanalyse“ (QMM) eingeführt. Inhalt des Moduls QMM ist die Theorie und Anwendung von statistischen und ökonometrischen Methoden. Zur praktischen Anwendung der Methoden wurde bis zum Sommersemester 2010 die kommerzielle Software STATA eingesetzt. Die Software STATA beinhaltet umfangreiche Funktionen zur Datenanalyse, wird aber aufgrund der Lizenzkosten von den Studenten in der Regel nur innerhalb des Computerlabors an der Universität genutzt. Zur Vermeidung der Kosten wird die frei verfügbare Open Source-Software GRETL (<http://GRETL.sourceforge.net/>) seit dem Sommersemester 2010 genutzt, welche ebenso alle in dem Modul benötigten Methoden für empirische Untersuchungen zur Verfügung stellt. Wie in STATA steht dabei sowohl die Eingabe über eine leicht verständliche grafische Benutzeroberfläche als auch über eine Kommandozeile zur Verfügung. Im Folgenden werden die Schritte des Projektmanagementzyklus kurz skizziert.

#### Planung

**Ziel:** Verbesserte „Fernbetreuung“ der Studenten, wenn diese auch zuhause mit der frei verfügbaren Software GRETL arbeiten. Der Aufwand für die Kursbetreuung soll durch den GRETL-Einsatz nicht steigen. Eine positive Lernerfahrung der Studenten durch das eLearning. **Kritischer Erfolgsfaktor:** Aufwand, approximativ gemessen in elektronischen Nachrichten, die der Kursbetreuer je Student schreiben muss, im Vergleich zum Vorjahr 2009 ohne eLearning und frei verfügbare Software. Diese Größe vernachlässigt zwar die auf das Verfassen verwendete Zeit, ist aber leicht zu erheben. **Zielgruppe:** Studenten des 4. Fachsemesters Agrarwissenschaften (B.Sc.) mit Schwerpunkt Agrarökonomie und Agribusiness. QMM ist ein Pflichtmodul. **Blended Learning Szenario:**

- Präsenzvorlesungen und -computerübungen zur Faktenvermittlung
- Online Lernplattform basierend auf dem Open-Source-System OLAT ([www.olat.org](http://www.olat.org)), welche von der CAU inklusive technischem Support zentral zur Verfügung gestellt wird (daher vermutlich auch in Zukunft verfügbar), mit folgenden Inhalten:
  - Aufnahmen von Ton und Bild vom Tablet PC aus der Vorlesung
  - Multiple Choice Tests, die interaktiv am Rechner gelöst und korrigiert werden

## *Quantitative Untersuchung zur Effizienz von eLearning*

- Kleine Datensätze und Aufgaben zum Üben der Anwendung der Software GRETL
  - Links zu externen Simulationsexperimenten (Java-Applets), z.B. zentraler Grenzwertsatz der Statistik
  - Beispiele der Anwendung von quantitativen Methoden außerhalb der Ökonomie
  - Forum für Fragen, Kommentare und Austausch
  - Materialsammlung (Vorlesungsfolien, Formelsammlung, statistische Tabellen, Fragenkatalog für das autonome Lernen, Literaturangaben und Skripte)
  - Betreuung einer in zweier Teams selbstständig durchzuführende Datenanalyse (Details dazu im Beitrag von Dickel und Loy (2010)).
- Prüfungsleistung des Moduls QMM:
    - 25% sind eine selbstständige Datenanalyse in Zweierteams
    - 75% sind eine 90 minütige schriftliche Klausur

**Budgetierung:** Studentische Hilfskräfte zur Erstellung der Inhalte (ohne Examen, 4 Monate à 30 Stunden) und die benötigten technischen Utensilien (Software, Mikrofon und Tablet PC für Vorlesungsaufnahmen) werden durch die zentrale eLearning-Förderung der CAU finanziert bzw. zur Verfügung gestellt; Betreuung der Lernplattform und der Inhalteerstellung durch eine viertel Stelle eines wissenschaftlichen Mitarbeiters.

### **Implementierung/Umsetzung**

Im Sommersemester 2010 sind auf der Lernplattform 71 Teilnehmer in der Veranstaltung registriert. Die Inhalte der Lernplattform sind durch studentische Hilfskräfte realisiert worden. Während des Semesters nutzen die Studenten die Lernplattform zur Bearbeitung der Datenanalyse nur zögerlich. Die Nutzung wird kurz vor dem Abgabetermin der Datenanalyse und dem Klausurtermin intensiviert. Der Aufwand der Betreuung (Anzahl elektronischer Nachrichten) während des Vorlesungszeitraums ist gering.

### **Evaluation/Kontrolle**

Zur Evaluierung des eLearning-Einsatzes wird neben der Erhebung der Anzahl

### 5.3 Praxisbeispiel „eLearning QMM“

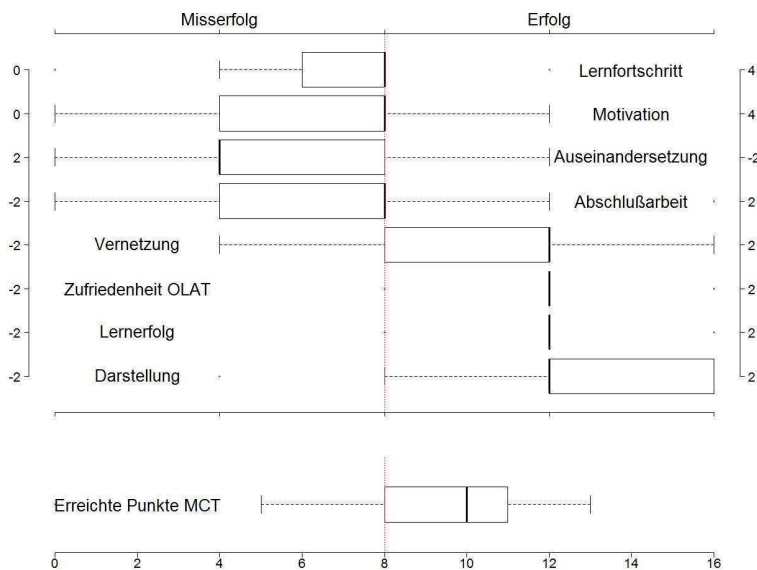


Abbildung 5.14: Ergebnis Untersuchung der Lernerfahrung der Studenten. Ska-lierung der Faktoren gemäß der Codierung der Antwortmöglich-keiten (siehe Fragen 1) - 8) des Fragebogensauszugs im Anhang), maximal zu erreichende Punktzahl des MCTs ist 16.

der Nachrichten eine schriftliche Befragung der Studenten in der letzten Präsenz-Lehrveranstaltung des Moduls durchgeführt. Der Fragebogen beinhaltet 32 Fragen, die zum größten Teil mit 5-stufigen äquidistanten Likertskalen zu beantworten sind. Neben Fragen zur Charakterisierung der Studenten (Vorwissen, Lernorte, Internet- und GRETL-Verfügbarkeit, Nutzungshäufigkeiten, Lerndauer, sozio-demographische Angaben) gibt es Fragen zu den eLearning-Inhalten (Bewertung, neu aufzunehmende und zu intensivierende Inhalte, negative Aspekte, technische Probleme, Verbesserungsvorschläge) und zu der Lernerfahrung der Studenten. Da die Lernerfahrung nicht nur an einem Prüfungsergebnis festzumachen ist (Schulmeister u. a., 2005; Preussler und Baumgartner, 2006), wird ein Konstrukt zur Messung entwickelt, welches die folgenden Faktoren beinhaltet:

1. Möglichkeit der Einschätzung des eigenen Lernfortschritts,
2. Motivation zur Auseinandersetzung mit dem Stoffgebiet,
3. Förderung der aktiven Auseinandersetzung mit dem Lehrstoff,
4. Erwägung einer selbständigen empirischen Untersuchung im Rahmen der Abschlussarbeit,

## Quantitative Untersuchung zur Effizienz von eLearning

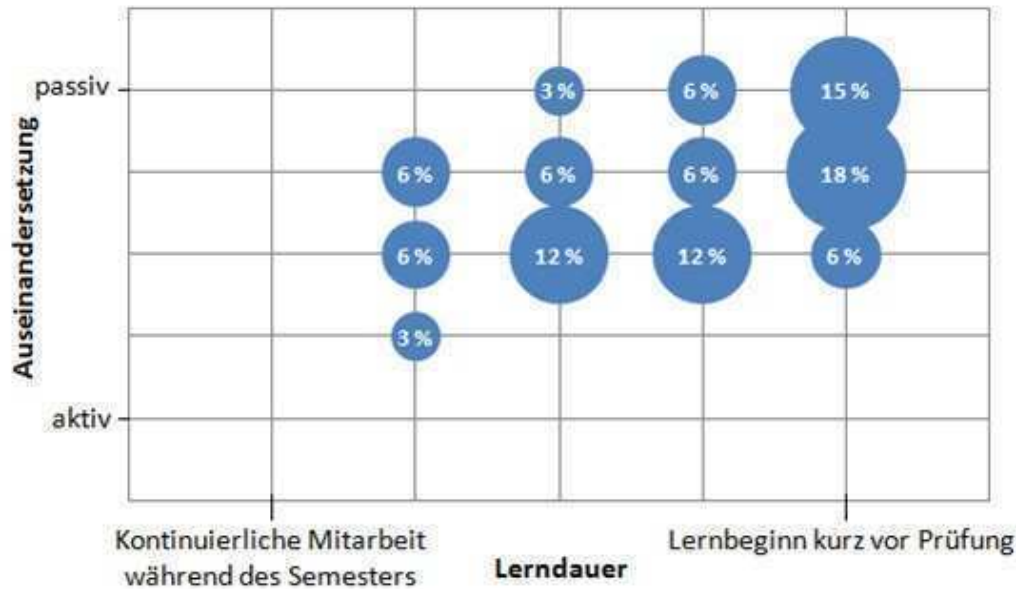


Abbildung 5.15: Auseinandersetzung mit den Lehrinhalten und Lerndauer der Studenten. Abbildung der Antworten zu Frage 6) (Ordinate) und 9) (Abzisse) des Fragebogensauszugs im Anhang.

5. Vernetzung der Inhalte von Vorlesung, Übung und Lernplattform,
6. Zufriedenheit bezüglich des Angebots der Lernplattform OLAT,
7. Beeinflussung des Lernerfolgs durch das Lernplattformangebot,
8. Unterstützung durch Darstellungsvarianten der Inhalte auf der Lernplattform.

Das Konstrukt wird mittels der Fragen 1) - 8) (siehe Auszug des Fragebogens im Anhang) quantifiziert. Zusätzlich beinhaltet der Fragebogen einen Multiple Choice Test (MCT), der das im Modul QMM vermittelte Wissen abfragt.

### Optimierung

Ein Wilcoxon-Rangsummentest weist keinen signifikanten Unterschied der Anzahl der elektronischen Nachrichten je Student zwischen der Veranstaltung 2009 (ohne eLearning und frei verfügbarer Software) und der Veranstaltung 2010 (mit eLearning und frei verfügbarer Software) auf ( $W=728,5$ ,  $p\text{-Wert}=0,2797$ ). Das Ziel eines nicht steigenden Betreuungsaufwands ist erreicht worden. 34 Studenten füllten den Fragebogen aus. Gemäß dem entwickelten Konstrukt bzgl. der Lernerfahrung ist

#### 5.4 Nutzen, Aufwand und weitere Informationen

der Soll-Zustand, dass sich in Abbildung 5.14 alle Boxplots auf der rechten Seite befinden. Das ist nicht der Fall. Es fand vermehrt eine passiv-konsumierende Nutzung der Inhalte statt und weniger eine aktive kognitive Auseinandersetzung. Auch sind die Motivation zur Auseinandersetzung mit dem Stoffgebiet, die verbesserte Einschätzung des eigenen Lernfortschritts und Organisation der Inhalte durch Vernetzung von Vorlesung, Lernplattform und Übung nicht wie gewünscht erfüllt worden. Die abgefragten Charakteristiken der Zielgruppe zeigen unter anderem, dass bei den Studenten zusätzlich mittels Anreizen das kontinuierliche Arbeiten und die aktive Auseinandersetzung mit dem Lehrstoff zu optimieren sind (siehe Abbildung 5.15).

Aufgrund der insgesamt positiven Evaluierung wird das eLearning Angebot unter Berücksichtigung der Optimierungspotentiale beibehalten. Neuerungen im zweiten Prozessdurchlauf im Sommersemester 2011 sind ein Vorlesungsskript als roter Faden der Veranstaltung, zur besseren Vernetzung der Angebote, und die gemeinsame, aktive Bearbeitung von Datensätzen mittels online Dateidiskussion. Insgesamt ist zu sagen, dass die Qualität des Moduls QMM durch das Prozessmanagement klar profitiert hat.

## 5.4 Nutzen, Aufwand und weitere Informationen

Ćukušić u. a. (2010) konnten empirisch zeigen, dass es eine positive Beziehung zwischen der systematischen Umsetzung gemäß ihrem entwickelten Prozess und dem Lernergebnis von eLearning gibt, unabhängig vom gelehrten Fachgebiet. Weitere Nutzen eines Prozessmanagementzyklus sind, dass der sich wiederholende Zyklus standardisiert und effizienter gestaltet wird. Dadurch werden (wiederkehrende) Fehler vermieden. Da das Endprodukt eLearning-Umgebung transparenter wird und leichter zu kommunizieren ist, werden Kooperationen von technischem Support und Lehrstühlen einfacher. (Page, 2010, S. 3). Der zyklische Aufbau unterstützt direkt die Wiederverwendung und zukünftige Anpassung der eLearning-Inhalte. Der Aufwand des Prozessmanagements ist insbesondere in der Planungsphase aufgrund der notwendigen Zielbestimmung, der Festlegung des methodologischen Rahmens und des Lehrszenarios unter Berücksichtigung der individuellen Gegebenheiten hoch. Unter Benutzung der gemachten Zielbestimmung ist hingegen

## *Quantitative Untersuchung zur Effizienz von eLearning*

eine Anpassung der häufig bereits im Rahmen des Qualitätsmanagements vieler Hochschulen vorhandenen Evaluationsprozesse mit geringem Aufwand verbunden, um hilfreiche Antworten zur Zielerreichung und dem spezifischen Lehr-Lernprozess zu erhalten. (Böss-Ostendorf und Senft, 2010, S. 280 f.) Eine Beschränkung der Durchführbarkeit des Prozessmanagements besteht darin, dass u.a. nicht immer alle Ziele quantifizierbar sind. Auch bestehen an Hochschulen Schwierigkeiten, die Budgetierung exakt zu analysieren, da beispielsweise Kostenzuweisungen im Rahmen zentral zur Verfügung gestellter Lernplattformen schwierig sind. Abschließend sei auf eine Reihe von Literaturquellen hingewiesen, die weitere Informationen zum Thema Projektmanagement und eLearning enthalten. Es gibt eine Vielzahl von Leitfäden zur Planungsphase eines eLearning-Prozesses (Schisler, 2009, S. 256), u.a. bietet Meier (2006) eine gute Orientierungshilfe an. Zur Evaluierung von eLearning im Allgemeinen gibt z.B. Glowalla u. a. (2009) einen guten Überblick. Die allgemeine Literatur zum Prozessmanagement (beispielsweise Anupindi u. a. (2012), Becker, Kugeler und Rosemann (), Page (2010) hat häufig betriebswirtschaftliche Aspekte als Schwerpunkt, bietet aber dennoch viele Informationen zu Instrumenten und Umsetzung. Speziell mit Prozessen im Rahmen von eLearning haben sich z.B. Hoppe und Breitner (2006) beschäftigt, die einen Prozess mit Schwerpunkt auf Evaluierung und Optimierung von eLearning vorstellen. Kandzia und Trahasch (2012) stellen die Vorteilhaftigkeit des Prozessmanagements für eLearning insgesamt an deutschen Hochschulen heraus.



# Literatur

- Anupindi, Ravi u. a. (2012). *Managing Business Process Flows: Principles of Operations Management*. 3. Aufl. New Jersey: Pearson.
- Becker, Jörg und Dieter Kahn (2010). »The Process in Focus«. In: *Process Management: A Guide for the Design of Business Processes*. Hrsg. von Jörg Becker, Martin Kugeler und Michael Rosemann. 2. Aufl. Heidelberg: Springer, S. 3–13.
- Becker, Jörg, Martin Kugeler und Michael Rosemann.
- Böss-Ostendorf, Andreas und Holger Senft (2010). *Einführung in die Hochschul-Lehre: Ein Didaktik-Coach*. Stuttgart: UTB.
- Ćukušić, Maja u. a. (2010). »E-Learning process management and the e-learning performance: Results of a European empirical study«. In: *Computers & Education* 55, S. 554–565.
- Dickel, Meike und Jens-Peter Loy (2010). »OLAT zur Unterstützung der individualisierten Betreuung studentischer Projektarbeiten«. In: *Hamburger eLMAGAZIN* 5, S. 27–28.
- Fabac, John N. (2006). »Project Management for Systematic Training«. In: *Advances in Developing Human Resources* 8(4), S. 540–547.
- Glowalla, Ulrich u. a. (2009). »Methoden und Ergebnisse der Evaluation elektronischer Lernangebote«. In: *Online-lernen: Handbuch für Wissenschaft und Praxis*. Hrsg. von Ludwig J. Issing und Paul Klimsa. München: Oldenbourg, S. 309–328.
- Golzari, Z. u. a. (2011). »Assessment model of E-learning for higher education«. In: *Journal of American Science* 7(7), S. 792–799.
- Hoppe, Gabriela und Michael H. Breitner (2006). »Evaluation and Optimization of E-Learning Scenarios«. In: *E-Learning Geschäftsmodelle und Einsatzkonzepte*. Hrsg. von Michael H. Breitner und Günter Fandel. Zeitschrift für Betriebswirtschaft Special Issue 2/2006. Wiesbaden: Gabler Verlag, S. 43–61.

- Kandzia, Paul und Stephan Trahasch (2012). »E-Learning an Hochschulen - Vom Projekt zum Prozessmanagement«. In: *i-com* 11(1), S. 34-37.
- Meier, Rolf (2006). *Praxis E-Learning: Grundlagen, Didaktik, Rahmenanalyse, Medienauswahl, Qualifizierungskonzept, Betreuungskonzept, Einführungsstrategie, Erfolgssicherung*. Offenbach: Gabal.
- Page, Susan (2010). *Business Process Improvement: 10 Simple Steps to Increase Effectiveness, Efficiency, and Adaptability*. New York: Amacom.
- Preussler, Annabell und Peter Baumgartner (2006). »Qualitätssicherung in mediengestützten Lernprozessen - zur Messproblematik von theoretischen Konstrukten«. In: *Qualitätssicherung im E-Learning*. Hrsg. von A. Sindler u. a. Waxmann, S. 73-85.
- Schisler, Peter (2009). »Planung und Entwicklung von Online-Lernangeboten in der Praxis«. In: *Online-lernen: Handbuch für Wissenschaft und Praxis*. Hrsg. von Ludwig J. Issing und Paul Klimsa. München: Oldenbourg, S. 255-262.
- Schulmeister, Rolf u. a. (2005). »Konzeption und Durchführung der Evaluation einer virtuellen Lernumgebung: Das Projekt Methodenlehre-Baukasten«. In: *Handbuch Medienpädagogik Bd. 5*. Hrsg. von Ben Bachmair, Peter Diepold und Cladua de Witt. Opladen: Verlag für Sozialwissenschaft, S. 37-52.

## Anhang

Die folgenden Fragen und Antwortmöglichkeiten sind ein Auszug des verwendeten Evaluations-Fragebogens. In Klammern ist die Codierung angegeben, die in den Abbildungen an den Achsenskalen verwendet wird.

### 1. Faktor Darstellung:

Neben Texten und Graphiken, vor allem im Forum, wurden auch Videos präsentiert. Hat diese Kombination von unterschiedlichen Darstellungsvarianten Sie im Lernprozess unterstützt oder verwirrt? Kreuzen Sie bitte an.

Antwortmöglichkeiten:

Unterstützende Darstellung (2) - Eher unterstützende Darstellung (1) -  
Zu gleichen Teilen unterstützende und verwirrende Darstellung (0) - Eher  
verwirrende Darstellung (-1) - Verwirrende Darstellung (-2)

### 2. Faktor Vernetzung:

Inwieweit stimmen Sie der Aussage zu, dass die Inhalte der verschiedenen Lehrangebote (Vorlesung, Lernplattform, Übung) gut vernetzt sind, d.h. die Überschneidungen der Inhalte sind sinnvoll und es gibt keine Lücken? Kreuzen Sie bitte an.

Antwortmöglichkeiten:

Stimme voll und ganz zu (2) - Stimme zu (1) - Teils/teils (0) - Lehne ab (-1) - Lehne voll und ganz ab (-2)

3. Faktor Abschlussarbeit:

Können Sie sich vorstellen, in Ihrer Abschlussarbeit selbständig eine empirische Analyse durchzuführen und dafür die erlernten quantitativen Methoden einzusetzen? Kreuzen Sie bitte an.

Antwortmöglichkeiten:

Kann ich mir sehr gut vorstellen. (2) - Kann ich mir eher gut vorstellen. (1) - Teils/teils (0) - Kann ich mir eher nicht vorstellen. (-1) - Kann ich mir gar nicht vorstellen. (-2)

4. Faktor Lernfortschritt:

Inwiefern hat Ihnen die Arbeit mit der Lernplattform eine Einschätzung Ihres Lernfortschrittes und Wissensstandes ermöglicht? Kreuzen Sie bitte an.

Antwortmöglichkeiten:

Einschätzung sehr gut möglich (4) - Einschätzung möglich (3) - Einschätzung mehr oder weniger möglich (2) - Einschätzung eher nicht möglich (1) - Keine Hilfe bei der Einschätzung. (0)

5. Faktor Motivation:

Hat die Lernplattform Sie motiviert, sich mit dem Stoffgebiet der Veranstaltung QMM auseinanderzusetzen? Kreuzen Sie bitte an.

Antwortmöglichkeiten:

Sehr stark motiviert (4) - Stark motiviert (3) - Mehr oder weniger motiviert (2) - Eher nicht motiviert (1) - Gar nicht motiviert (0)

6. Faktor Auseinandersetzung:

Die Auseinandersetzung mit den auf der Lernplattform angebotenen Materialien kann eher passiv oder aktiv sein. Lerntyp E (passiv) sieht sich die Vorlesung an, die er verpasst hat, überfliegt die Seminararbeiten und liest

angebotene Lösungsskizzen. Lerntyp F (aktiv) kontrolliert die Richtigkeit seiner erstellten Mitschriften, indem er die Vorlesungen noch einmal mithilfe der Aufnahmen durcharbeitet. Wenn beim Bearbeiten der GRETL-Aufgaben Probleme auftauchen, nutzt er das Forum zur Kommunikation und Problemlösung. Die Selbsttests dienen ihm zur Selbstkontrolle. Wie würden Sie Ihre eigene Auseinandersetzung mit den auf der Lernplattform angebotenen Materialien beschreiben. Kreuzen Sie bitte an.

Antwortmöglichkeiten:

Lerntyp E (passiv) (2) - Eher Lerntyp E (1) - Zu gleichen Teilen Lerntyp E und Lerntyp F (0) - Eher Lerntyp F (-1) - Lerntyp F (aktiv) (-2)

7. Faktor Zufriedenheit:

Wie würden Sie Ihre eigene Zufriedenheit bezüglich des Angebots der Lernplattform im Rahmen der Veranstaltung QMM bewerten? Bitte kreuzen Sie an und geben in kurzen Stichpunkten eine Begründung dazu.

Antwortmöglichkeiten: Ich bin mit der Lernplattform...

Sehr zufrieden (2) - Eher zufrieden (1) - Weder zufrieden, noch unzufrieden (0) - Eher unzufrieden (-1) - Sehr unzufrieden (-2)

8. Faktor Lernerfolg:

Wenn Sie alles in allem über die Lernplattform nachdenken: hat diese insgesamt Ihren Lernerfolg im Fach QMM positiv oder negativ beeinflusst? Bitte kreuzen Sie an.

Antwortmöglichkeiten: Die Lernplattform hat meinen Lernerfolg im Fach QMM...

Positiv beeinflusst (2) - eher positiv beeinflusst (1) - weder positiv noch negativ beeinflusst (0) - eher negativ beeinflusst (-1) - negativ beeinflusst (-2)

9. Es gibt unterschiedliche Lerndauern. Lerntyp C bevorzugt es, erst kurz vor den Prüfungen mit der Vorbereitung zu beginnen. Lerntyp D hingegen favorisiert es es, kontinuierlich von Anfang an während des Semesters zu lernen. Wie würden Sie Ihre eigene Lerndauer beschreiben? Bitte kreuzen Sie an.

Antwortmöglichkeiten:

Lerntyp C (Lernbeginn kurz vor Prüfung) - Eher Lerntyp C - Zu gleichen

Teilen Lerntyp C und D - Eher Lerntyp D - Lerntyp D (kontinuierliche  
Mitarbeit während des Semesters)



# Kapitel 6

## EVALUATION OF ELEARNING - A STUDY OF UNDERGRADUATE AGRICULTURAL ECONOMICS COURSE

Autoren:

Meike Wocken, Jens-Peter Loy

Posterbeitrag:

EAAE 2011 Congress, Change and Uncertainty: Challenges for Agriculture, Food and Natural Resources, 30. August-2. September 2011, Zürich, Schweiz.

Eine erste Version des Beitrags („Evaluierung von eLearning-Instrumenten im Bereich der universitären Lehre“) ist erschienen in Clasen, M., Schätzel, O., Theuvsen, B. (Hrsg.) (2011), Qualität und Effizienz durch informationsgestützte Landwirtschaft, Fokus: Moderne Weinwirtschaft, Referate der 31. GIL-Jahrestagung, 24.-25. Februar 2011, Oppenheim. Lecture Notes in Informatics. Bonn:Gesellschaft für Informatik. S. 181-184.

## Abstract

*In this study we evaluate the combination of traditional classroom education with additional content and support provided by an online Learn Content Management System (LCMS). Main focuses are individual learning characteristics and eLearning aspects. Findings are that not all dimensions of eLearning success meet the expected outcomes. Using achieved scores in examinations to measure successful application of eLearning is suitable only to a limited extent. eLearning objectives have to be defined with respect to eLearning's target group.*

## 6.1 Introduction

Over the last years the usage of electronic learning (eLearning) in university programs has increased rapidly. The term eLearning includes different types of learning, all supported by electronic media. Yaghoubi, 2009 find a positive students' perception of eLearning. From student's point of view eLearning can be used to improve education quality. But, Yaghoubi also indicates that universities should focus on students' attitudes regarding eLearning. This goes along with different publications pointing out that learning environment needs to be target-group specific (e.g. Ozkan and Koseler, 2009, Kuechler and Schmidt, 2007). Milliken and Barnes, 2002 tried to answer the question of the extent to which new technology aids improving educational quality. They found the need to develop the learning environment of students to ensure that they all had a positive experience.

In this study we analyze the combination of traditional classroom education with additional content and support provided by an online Learn Content Management System (LCMS), so called Blended Learning. Due to the positive link between planning and controlling of the eLearning process and its learning outcomes, an eLearning process management is conducted for implementation and application of the Blended Learning. The eLearning process management consists of four stages: planning, organizing/implementing, controlling and improving the eLearning (Ćukušić et al., 2010). After planning, organizing and implementing the LCMS an evaluation is conducted to measure user satisfaction and outcomes of specific eLearning aspects. In addition, the evaluation should provide detailed



target-group characteristics. To improve the eLearning in the fourth stage of process management, collected data analysis is used to customize and enhance the LCMS towards students' needs and usage. Particular emphasis is placed on exploring the different outcomes of specific eLearning aspects related to student's characteristics by means of principal components analysis. The exploration should give insight of how to deal with the, in the above listed publications, claimed consideration of target-group specifics in eLearning evaluation and improvement.

The study is organized as follows. First, the dataset is depicted including data collection and the instruments to measure outcomes of eLearning aspects. Next, the data analysis and used methods are described. Further, the empirical results are reported. Finally, conclusions of this study are summarized.

## 6.2 Data and Methods

This section presents course and data description. Data description comprises explanation of the theoretical construct of eLearning success. In the end of the section used analysis methods are shortly introduced.

### 6.2.1 Course description

This course aims to provide students with the opportunity to carry out statistical and econometrical research. All students studying agricultural sciences with a major in agricultural economics and agribusiness are required to undertake this course. By the end of the course students should have

- developed the ability to apply statistical and econometrical methods used in the agricultural market and food industry analysis;
- developed an understanding of using the theoretical foundations;
- developed substantive knowledge in application of statistical software;
- developed the capability to undertake independent research and writing.

## *Evaluation of eLearning*

Lecture is supported by multimedia applications. Via a tablet PC handwritten notes (Windows Journal), slides (PowerPoint), statistical examples (Open Source Statistic-Software GRETL) and simulations (Excel) are employed. A practical course taking place in the computer lab accompanies the lecture. Students have to submit their own data analyses project in addition to a written examination at the end of the course. Project's mark will be worth 25% of the overall mark.

Additional contents provided by the online LCSM are:

- **Record** of the in classroom used beamer presentation and corresponding voice of the lecturer;
- **Forum** to discuss and ask questions regarding the course;
- Material (Lecture slides, formulary, statistical tables, list of questions as guideline for autonomous learning);
- **Instructions** and **examples** how to realize the data analysis-project independently;
- Small **data sets and assignments** (including solutions) for own use of the statistical software GRETL;
- Multiple choice **self-tests** covering a wide range of course matters;
- **Examples of use** to motivate the subject beyond economics: Quantitative research publications from crop science, animal nutrition, animal physiology, and food science;
- **Link list**, bibliographical **references**.

### **6.2.2 Data description**

Student feedback is closest to student experience. Therefore, it is used in the controlling stage of the eLearning process management. To cope with the often emerging low response rate (Jara and Mellar, 2010), the sample was gathered in a practical course session in the computer lab at the end of the semester. The used survey-questionnaire consists of 34 different questions and a multiple-choice test (MCT) concerning course content. It has to be filled by hand. Survey and

	Mean	Median	Std. Dev.
Progress	1.844	2	0.7761
Encouragement	1.727	2	0.6261
Work	0.8182	1	0.8277
Success	1.031	1	0.5561
Display formats	1.194	1	0.7611
Thesis	-0.2121	0	1.0635
Points	9.647	10	2.0342
Frequency	2.441	3	0.8424
Network	0.5	1	0.7184
Duration of learning	0.9118	1	1.1288
Satisfaction LCMS	1.030	1	0.5040
Satisfaction Lecture	-0.09091	0	1.0080
Satisfaction practical course	0.9091	1	0.7399
Intrinsic	-0.5882	-1	0.8793
Extrinsic	1.0594	1	0.9070

Table 6.22: Objectives of eLearning implementation - Descriptive statistics

multiple-choice test are in students' native language German. 71 students do have access to online content provided by the LCMS. 34 students completed the questionnaire and MCT. In the sample are 20 males and 14 females. The average age is 23. Just over half (53%) of all respondents would like to earn a Master's degree. In contrast, extrinsic motivation overmatch intrinsic motivation to deal with course matter (Wilcoxon signed-rank test [Alternative: Median Extrinsic > Median Intrinsic]:  $V = 483 **$ ). We expected a more intrinsic motivation to be well prepared for master's program.

High score in an exam not necessarily accompanies satisfaction with the (e)learning environment. Therefore, success of eLearning is a construct of the attainment of didactic functions as well as impartial and subjective learning outcomes, and eLearning aspects, especially how eLearning supports the learning. (Preussler and Baumgartner, 2006) To specify this construct, we are using particular didactic objectives of eLearning implementation, which are theoretically pointed out at the beginning of the eLearning process management. They ought to be the additional

## *Evaluation of eLearning*

benefits from eLearning beside the advantages of overcoming restrictions related to physical location and time of usage. These didactic objectives are: To encourage learning activities, to represent knowledge (organizing and structuring knowledge, depict knowledge by different display formats), and to control and regulate the learning process (Kerres, 1999). Further, we consider that the implementation of eLearning gives impulses to work independently with quantitative methods and helps to achieve high scores in exams. To analyze eLearning success in terms of these different aspects, items listed in the appendix are asked in the survey-questionnaire. The scales are five-point Likert-scales (except points gained in the additional multiple-choice test). The scales are assumed to be equidistant, so interval scale is presumable. Table 6.22 contains corresponding descriptive statistics.

Impulses to work independently with quantitative methods are measured by *Thesis*. An objective learning success is quantified by the achieved score in the MCT (*Points*). Encouragement learning activities is measured by *Encouragement* and *Work*; *Display formats* quantifies the support of representation of knowledge through different display formats; *Network* is a parameter to check the structure of the different course offers and how interlocked they are; *Progress* is a variable to detect students' problems in controlling their learning process. Box plots of the mentioned variables are represented in Figure 6.16. On the vertical axes, the corresponding values of the Likert-Scales are written. Figure 6.16 provides an easy to interpret graphical representation of eLearning success basing on our construct: eLearning success means to have all box plots situated on the right hand side of the scale. We cannot observe such situation. We have to perceive, the implementation of eLearning is not as successful as expected. Further analyzes is conducted. In the next section employed methods are described.

### **6.2.3 Methods**

To analyze the collected data different methods are adopted. At first a frequency analysis is made graphically: It displays relative usage frequency of an online-content versus the appraisal in terms of how helpful the content is. Further, the parameters of a linear regression model are estimated using linear least squares

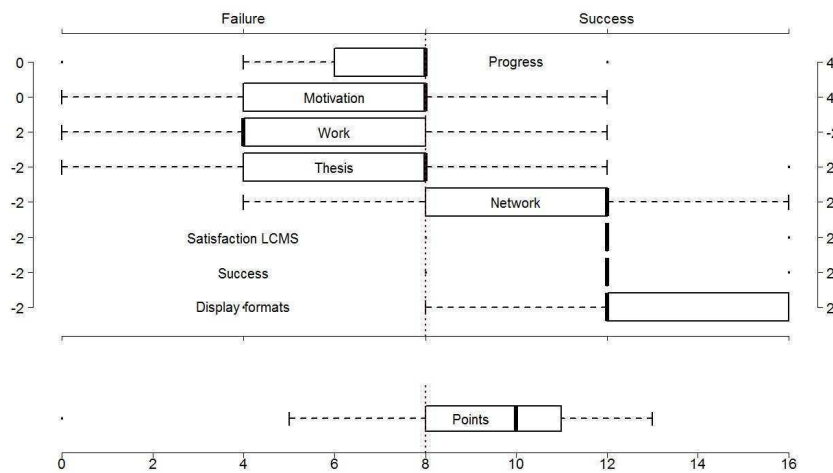


Figure 6.16: Dimensions of eLearning success

techniques. Finally, to understand response behavior related to student's characteristics principal components analysis (PCA) is applied. To extract variables for PCA the Anti-Image matrix is consulted. Diagonal elements on anti-image matrix for each variable are a measure of sampling adequacy (MSA). If MSA-value is bigger than 0.5, these variables should stay in the PCA. Before using PCA, Kaiser-Meyer-Olkin Test (KMO) and Bartlett's test of sphericity are executed to test if the chosen variables are suitable for PCA. KMO compares observed correlation coefficients to the partial correlation coefficients. KMO value is in  $[0, 1]$ . If it is bigger than 0.5, PCA is suitable. (Ahmad and Härdle, 2008) Analysis is done in the software environment R.

## 6.3 Results

Figure 6.17 shows the frequency analysis. Self-test, link list, references, and examples of use are relative often not used. Instructions for the project work, examples of such project works, but also small datasets including practical assignments and solutions are more frequent used. Students appraise records, forum, and lecture material to be most helpful, in relation to the other offered content, and use it frequently. We assume that this usage pattern results from the fact to be an obligatory course. Presumably, most students only want to pass the exam. Self-test, link list, bibliographical references, and practical examples of use

## Evaluation of eLearning

Points	Estimator	Std. Dev.	p-value
Intercept	5.97	1.24	$3.68e - 05 * **$
Frequency	1.27	0.49	0.014*
Examination	0.83	0.89	0.35
F-Statistic (2, 31) = 4.925		$R^2$ : 0.24	
p-Value = 0.01389		adj. $R^2$ : 0.19	

Table 6.23: Linear Model explaining points gained in Multiple Choice Test. Examination is dummy-variable (=1 if student is registered for the 1st examination date and therefore probably already preparing and learning for examination).

exceed subject matter, which is asked in the examination or needed in project work.

Results for the estimated linear model are summarized in Table 6.23. Our findings highlight significant positive relation between frequency of usage LCMS and gained points in MCT. Beside the poor Goodness-of-fit, the effect of Frequency might be overestimated. We have to consider the positive link between frequency using the LCMS and general interest in quantitative methods. Frequency is a measure of relative time spent on the LCMS. The group of students who is particular interested in quantitative methods is likely to use the LCMS more often. Also, this group is likely to achieve in general better test results independent of the LCMS. Altogether, we receive first impression of eLearning's impact: Either the objective learning success is positively concerned by usage-frequency, or at least the group of interested students receives an additional support by LCMS content.

To be able to identify relationships between variables in the data-set, e.g. patterns in students' characteristics and reported survey specifications, PCA is used in the last part of data analysis. The values on the diagonal of the anti-image correlation matrix are a measure of sampling adequacy (MSA). Based on these MSA-values, variables are excluded from PCA. The remaining variables are: *Duration of learning, points, extrinsic, satisfaction LCMS, work, success, frequency, and thesis*. The KMO value of these variables is 0.64 (Table 6.24). It is an acceptable value to continue PCA. The accomplishment of PCA is also supported by Bartlett's test of

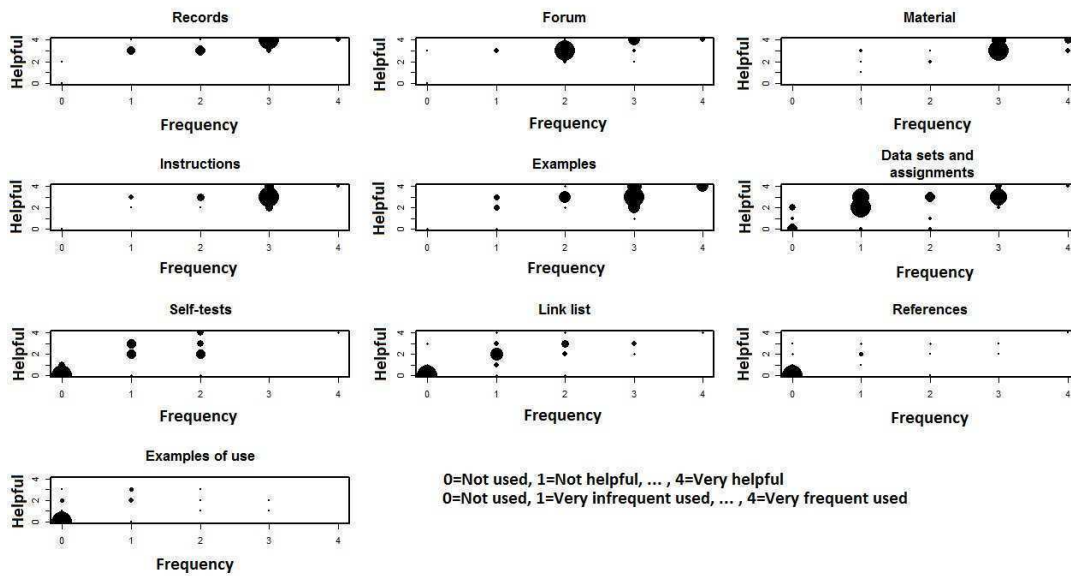


Figure 6.17: Relative usage frequency of an online-content versus the appraisal in terms of how helpful the content is.

	Value
KMO Kaiser-Meyer-Olkin Measure of Sampling Adequacy	0.64
Bartlett's Test of Sphericity Approx. Chi-Square	110.46
df	28
Significance	0.00

Table 6.24: Kaiser-MeyerOlkin Measur of Sampling Adequacy and Bartlett's Test of Sphericity

sphericity. The hypothesis that the variables are uncorrelated in the population has to be rejected (Table 6.24: Significance p-value= 0.00).

For PCA the correlation matrix is used. Four principal components are extracted. They explain nearly 80% of variation in the total data. To present the components in a more meaningful configuration, orthogonal varimax rotation is applied. Table 6.25 shows rotated loadings. The loadings express the correlation of the respective variable with the principal component. The variability in the original variables is almost retained, shown by the communalities in Table 6.26.

## Evaluation of eLearning

Variables	Component 1	Component 2	Component 3	Component 4
Duration of learning	-0.76			
Points	0.52			
Extrinsic	0.83			
Satisfaction LCMS		0.54		0.55
Work		0.61		
Success		0.88		
Frequency			-0.94	
Thesis				0.94
Rotation Sum of Squared Loadings	1.91	1.73	1.35	1.32
% of variance	23.9	21.6	16.9	16.5
Cumulative %	23.9	45.5	62.4	78.8

Table 6.25: Rotated component matrix. Rotation Method: Varimax with Kaiser Normalization. Values lower than 0.50 are not presented.

Considering the loadings of the rotated components, the first principal component is highly correlated with *Duration of learning* and *Extrinsic*, and it is moderately correlated with *Points*. This indicate, that high score in an exam not necessarily accompanies satisfaction with the eLearning environment. Instead, high score in an exam is attended by extrinsic motivation to achieve a good mark in the final exam and effort to learn continuously. Keyword for interpretation component 1 is **Effort**. This finding confirms our approach to treat success of eLearning as a construct of different aspects, not only objective test-results. The second component is strongly correlated with *Success* and it is correlated with *Work* and *Satisfaction LCMS*. This result can be associated to **Encouragement**, with regard to encourage the passive working student through the LCMS. The overall success of a passive working student is positively influenced by the LCMS. The variable *Frequency* determines the third component. It indicates that **Frequency** is nearly uncorrelated to other components as well as variables. Component 4 is strongly correlated with *Thesis* and moderately correlated with *Satisfaction LCMS*. Here, satisfaction of the LCMS is attended by conceiving a thesis, which is likely to be written using the learned quantitative methods. This might concatenates with **Econometrical Interest**.



	Communalities
Duration of learning	0.68
Points	0.89
Extrinsic	0.76
Satisfaction LCMS	0.73
Work	0.73
Success	0.90
Frequency	0.84
Thesis	0.80

Table 6.26: Communalities: proportion of variable's variance explained by the four components.

## 6.4 Conclusion

This study investigates success of eLearning. Success of eLearning is constructed by different aspects. A survey including a multiple-choice test was conducted to quantify this construct. First result is that the implementation of eLearning is not as successful as expected. Not all dimensions of eLearning success meet the expected outcomes. The insufficient aspects of the course must be corrected. According to the analyses of relative usage frequency of an online-content versus the appraisal in terms of how helpful the content is, students appraise organizing knowledge functions to be most helpful, in relation to the other offered content, and use it more frequent. Content, which exceed subject matter asked in the examination or needed in project work, is not used by students. Our assumption is that this usage pattern results from the fact being an obligatory course.

As an empirical finding, there is a significant positive relationship between LCMS usage frequency and achieved points in the multiple-choice test. The fitted linear model has limitation due to lack of exact knowing latent influence on the variable frequency. Usage frequency might be influenced by general interest in quantitative methods. Students with this interest tend to gain higher score in tests than students without this interest. Therefore, to explore variables correlation, PCA is employed. It shows that points achieved in the MCT come along with effort and extrinsic motivation. Hence, using Points to measure successful application

### *Evaluation of eLearning*

of eLearning is suitable only to a limited extent. Further result of PCA is joint appearance of satisfaction with LCMS with either econometrical interest or passive learning-attitude towards offered content.

Back to the initial question how to deal with the consideration of target-group specifics in eLearning evaluation and improvement: Students with econometrical interest are satisfied with LCMS as well as passive working students. Therefore, we have to reassess aims of implementing eLearning. Success is not only encouraging active learning attitude. Being an obligatory course, success is also to spend satisfaction with the learning environment in terms of LCMS, although there is only passive learning attitude. The different objectives within the construct of eLearning success have to be weighted by target-group specifics. Our conclusion is that we want to improve LCMS for a more active learning-behavior. But, to support passive learning we decide to improve the knowledge organization, still providing the existing content of LCMS. At least, this might help students without econometrical interest to pass the course satisfied within learning environment. With this, we try to fulfill Milliken and Barnes, 2002 conclusions, to meet individual learner's need.

Jara and Mellar, 2010 mention student representation amongst others to gain student feedback. But, they also constrain the approach because of problems finding proper student representation in eLearning courses with only few class attendances. Our findings from PCA suggest arbitrary choice for student representation needs to be avoided. Otherwise, the impact of student characteristics on eLearning success distorts the results. Solution of this problem might be weighting student representation according to their characteristics to be representative. For this, future research is needed. Moreover, an identical research where the course is elective should be conducted to see if there is any difference in the result. Also, further research should cover more courses in the sample. The group of student which clearly benefits from eLearning needs to be identified.

# Bibliography

- Ahmad, Taleb and Wolfgang Härdle (2008). *Statistics E-learning Platforms Evaluation: Case Study*. SFB 649 Discussion Paper 2008-058. visited on 13/01/2011. Berlin: Humboldt-Universität Berlin. URL: <http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2008-058.pdf>.
- Ćukušić, Maja et al. (2010). “E-Learning process management and the e-learning performance: Results of a European empirical study”. In: *Computers&Education* 55, pp. 554–565.
- Jara, Magdalena and Harvey Mellor (2010). “Quality enhancement for e-Learning courses: The role of student feedback”. In: *Computers&Education* 54, pp. 709–714.
- Kerres, Michael (1999). “Didaktische Konzeption multimedialer und telemedialer Lernumgebungen”. In: *HMD - Praxis der Wirtschaftsinformatik* 205, pp. 9–21.
- Kuechler, Tilman and Karen Schmidt (2007). “From e-Learning to integrated learning architectures: a novel approach to learning management in corporate and higher education contexts”. In: *Management & Marketing* 2(4), pp. 27–34.
- Milliken, John and L.Philip Barnes (2002). “Teaching and technology in higher education: student perceptions and personal reflections”. In: *Computers&Education* 39, pp. 223–235.
- Ozkan, Sevgi and Refika Koseler (2009). “Multi-dimensional students’ evaluation of e-learning systems in the higher education context: An empirical investigation”. In: *Computers&Education* 53, pp. 1285–1296.
- Preussler, Annabell and Peter Baumgartner (2006). “Qualitätssicherung in medien-gestützten Lernprozessen - zur Messproblematik von theoretischen Konstrukten”. In: *Qualitätssicherung im E-Learning*. Ed. by A. Sindler et al. Waxmann, pp. 73–85.

## *Bibliography*

Yaghoubi, Jafar (2009). “Assessment of agricultural extension and education graduate students’ perceptions of e-Learning in Iran”. In: *Procedia Social and Behavioral Sciences* 1, pp. 1914–1918.

# Appendix

Main items measuring success of eLearning implementation:

Items	Scale
<p>Progress</p> <p>To what extent does work with the LCMS help to appraise the own progress in learning and level of knowledge?</p>	<p>4 ... 0</p>
	<p>Appraising the learning progress is very good practicable. ... No help for appraising the learning progress.</p>
<p>Encouragement</p> <p>Does the LCMS encourage you to deal with the content of the course?</p>	<p>4 ... 0</p>
	<p>Very strong encouraging. ... Not encouraging at all.</p>
<p>Work</p> <p>How do you work with the content provided by the LCMS?</p>	<p>2 ... -2</p>
	<p>Passive work (watching missed lectures, reading offered solutions, quick scanning examples of projects). ... Active work (controlling correctness of notes from lecture by watching the records, using self-tests for self-control, communicating in the forum for solving any problems).</p>

Success	Taken all together, does the LCMS influence your success in the course positively or negatively?	2	...	-2
		Positive influence.	...	Negative influence.
Display formats	Does the combination of different display formats in the LCMS support you or bewilder you in your learning process?	2	...	-2
		Supporting display formats.	...	Bewildering display formats.
Thesis	Can you imagine writing your bachelor or master thesis using the learned quantitative methods?	2	...	-2
		Very well imaginable.	...	Not at all imaginable.
Frequency	In relation to the total time, how frequent do you use the LCMS?	4	...	0
		Very frequent.	...	Do not use it.
Network	Do you agree with the statement, the different offers (Lecture, LCMS, practical course) are well networked?	2	...	-2
		Strong agreement.	...	Strong disagreement.

Duration of learning	How do you describe your own duration of learning?	2	...	-2
		Learning on the verge of the examination.	...	Steady learning over the whole semester.
Satisfaction LCMS	How do you appraise your satisfaction concerning the LCMS?	2	...	-2
		Very satisfied.	...	Very unsatisfied.
Satisfaction Lecture	How do you appraise your satisfaction concerning the lecture?	2	...	-2
		Very satisfied.	...	Very unsatisfied.
Satisfaction practical course	How do you appraise your satisfaction concerning the practical course?	2	...	-2
		Very satisfied.	...	Very unsatisfied.
Intrinsic	Do you agree with the statement, the matter of lecture encourages me to learn for the subject?	2	...	-2
		Strong agreement.	...	Strong disagreement.
Extrinsic	Do you agree with the statement, a good mark in the examination encourages me to learn for the subject?	2	...	-2
		Strong agreement.	...	Strong disagreement.
Points	Points gained in the Multiple Choice Test	Maximum score: 16		





# Kapitel 7

## QUANTITATIVE UNTERSUCHUNG ZUR EFFIZIENZ VON ELEARNING

Autor:

Meike Wocken

Der Beitrag ist erschienen in  
Hambach, S., Martens, A., Urban, B. (Hrsg.) (2011), eLearning Baltics 2011:  
Proceedings of the 4th International eLBa Conference. Fraunhofer Verlag.  
Stuttgart. S. 233 -242.

## Zusammenfassung

*Der Einsatz von eLearning in der Hochschullehre wird häufig per se als effizientes und effektives Instrument angesehen. In einigen Veröffentlichungen wird allerdings die Vermutung geäußert, dass vermehrt Lernende einen Nutzen durch den Einsatz von eLearning haben, aber nicht Lehrende. Um dies zu untersuchen, quantifiziert die vorliegende Studie approximativ Kosten und Nutzen des Einsatzes von eLearning-Instrumenten.*

## 7.1 Einleitung

Die Effizienz und Effektivität des Einsatzes von eLearning in der Hochschullehre wird im Allgemeinen vorausgesetzt (z.B. in Ćukušić u. a. (2010), Figlio, Rush und Yin (2010), Milliken und Barnes (2002), Selim (2007)). Dabei gibt es Kosten-Nutzen-Analysen, die zeigen, dass eLearning-Szenarien sowohl Kosten einsparen, aber auch vermehrt Kosten verursachen können, die den Nutzen weit übersteigen (Schwarz, 2001; Twigg, 2003). Werden Aufwandsschätzung des Einsatzes von eLearning durchgeführt, so berücksichtigen diese oft nur die Erstellungskosten der Inhalte, aber nicht die Kosten der Durchführung und Betreuung einer Veranstaltung mit eLearning-Instrumenten (z.B. in Brosowski und Cleve (2009), Ćukušić u. a. (2010)). Einer der Gründe ist die Schwierigkeit der exakten Kostenzuordnung im Hochschulbereich, sowie die Quantifizierung des Nutzens. Häufig wird die technische Infrastruktur zentral zur Verfügung gestellt. Die Kosten der Bereitstellung werden daher auf der Ebene der einzelnen Lehrveranstaltung meistens nicht berücksichtigt. Es besteht auch die Schwierigkeit, Kosten, die explizit durch die mediengestützte Realisierung einer Veranstaltung zustande kommen, zu erheben. Insgesamt ist die Wirtschaftlichkeitsbetrachtung von eLearning im Verhältnis zur gesamten eLearning-Literatur bisher nur vereinzelt, zumeist im Rahmen privatwirtschaftlicher Unternehmungen (z.B. Bensberg, Grob und Brocke (2007), Brocke u. a. (2007)), untersucht worden (Bremer, Krömker und Voß, 2010; Lehner, Schäfer und Proksch, 2003). Es besteht Forschungsbedarf zur Entwicklung von praxisrelevanten, präzisen Instrumenten, um die Wirtschaftlichkeit von eLearning zu untersuchen (Saghafian, 2011). Der vorliegende Beitrag stellt einen praxisorientierten Ansatz zur Abschätzung des Aufwands der Durchführung eines

eLearning-Szenarios am Beispiel der Betreuung von Projektarbeiten vor. Dafür wird der Aufwand approximativ durch die Anzahl geschriebener Nachrichten des Kurs-Betreuers an die Studenten quantifiziert und mit einem Szenario ohne eLearning-Elemente verglichen. Diese Untersuchung ersetzt keine vollständige Kosten-Nutzen-Analyse. Sie soll vielmehr einen ersten Einblick geben, ob eLearning wirklich per se ein effizientes Instrument ist. Ziel ist es dabei, eine qualitative Befragung von Steffens (Steffens und Reiss, 2010) zu quantifizieren. Dort geben die meisten Befragten (2/3 der Stichprobe sind Mitarbeiter an deutschen Universitäten) an, dass die Lernenden den größten Nutzen hinsichtlich Effizienz durch eLearning erzielen, nicht aber Universitäten und/oder Lehrende. Im Folgenden werden zunächst die für die Untersuchung gewählten Variablen und Daten erläutert, sowie die zu untersuchenden Hypothesen vorgestellt. Anschließend werden die zur Analyse verwendeten Methoden beschrieben. Es werden die empirischen Ergebnisse zusammengefasst. Der Beitrag schließt mit einem Fazit.

## 7.2 Daten und Hypothesen

Die Betreuung von Projektarbeiten bietet Herausforderungen. In dem hier vorgestellten Szenario handelt es sich bei der Projektarbeit um eine eigenständige ökonomische Analyse anhand selbstgewählter Daten. Die schriftlichen Ausarbeitungen sind verknüpft mit den jeweiligen Datensätzen und Befehlsdokumentationen der durchgeführten Datenanalyse für jede einzelne Projektgruppe zu betrachten. Die Projektgruppen bestehen dabei aus jeweils zwei Studenten. Grundsätzlich sollte die Kommunikation offen und transparent für alle Beteiligten einer Projektarbeit stattfinden. Eine Betreuung per eMail hat die Nachteile, dass keine thematische und hierarchische Gliederung, sowie keine Funktion für ein Dateimanagement existieren. Die Kommunikation der Mitglieder einer Projektgruppe untereinander wird nicht unterstützt. Die Betreuung der Studenten mittels Gruppenfunktion in einem online Learn Content Management System (LCMS) schafft hier Abhilfe. Die Gruppenfunktion enthält einen Dateiordner zum hochladen der jeweiligen Daten und Ergebnisse und ist zusätzlich mit einem Forum ausgestattet. Dieses Forum soll soziale Interaktionen zwischen den Studenten der jeweiligen Projektgruppe auslösen, so dass sich der Betreuungsaufwand für den Lehrenden verringert. Ori-

## *Quantitative Untersuchung zur Effizienz von eLearning*

entierend an Laurillard (2007) zur Erstellung eines Kosten-Nutzen-Modells wurde versucht, den Haupttreiber des Aufwands bzw. der Kosten der Veranstaltung zu identifizieren. Das ist in dem hier betrachteten Fall die individuelle Betreuung der Studenten. Dafür werden zum größten Teil elektronische Nachrichten verwendet. Um den Aufwand abzuschätzen, wurde daher zur Approximation die Anzahl der vom Betreuer geschriebenen Nachrichten je Student herangezogen. Die Anzahl elektronischer Nachrichten vernachlässigt zwar den Umfang der jeweiligen Nachricht und die darauf verwendete Zeit zum Verfassen, ist aber eine für Studienzwecke leicht zu erhebende Größe. Elektronische Nachrichten umfassen eMails und Nachrichten via LCMS-Foren innerhalb der Gruppenfunktion. Durch den Einsatz des LCMS wird die Betreuung übersichtlicher gestaltet, als die zuvor verwendete reine eMail-Betreuung. Es wird angenommen, dass sich die Betreuung vereinfacht und sich die Anzahl der notwendigen Nachrichten aufgrund der verbesserten Übersichtlichkeit verringert. Die Qualität der Betreuung wird dadurch gesteigert, da durch die permanente Möglichkeit der Kontrolle ein frühzeitiges Eingreifen durch den Lehrenden bei ersichtlichen Verständnisfehlern der Studenten durchführbar ist. Das erhöht den Nutzen der Studenten, was näherungsweise in besseren Noten der Projektarbeit gemessen werden soll. Daraus werden die folgenden Hypothesen abgeleitet:

**Hypothese 1:** Durch den Einsatz des LCMS ist der Arbeitsaufwand zur studentischen Betreuung gesunken.

**Hypothese 2:** Durch den Einsatz des LCMS steigt der Nutzen der studentischen Betreuung. Dadurch steigen die Leistungen der Studenten in den Projektarbeiten.

Für die Analyse werden Daten von 75 Studierenden genutzt, die in dem obigen Szenario eine Projektarbeit angefertigt haben. 40% der Studenten haben die Veranstaltung 2009 (ohne Einsatz des LCMS) besucht, die übrigen 60% haben 2010 an der Veranstaltung teilgenommen und wurden mittels LCMS betreut. Jede Beobachtung beinhaltet die Anzahl der vom Betreuer an den jeweiligen Studenten gesendeten elektronischen Nachrichten und die Note der Projektarbeit. Einen ersten Eindruck der Anzahl der Nachrichten verschafft Abbildung 7.18.

## 7.3 Methoden

Aufgrund der Verteilungsfreiheit werden neben graphischen Analysen Wilcoxon-Rangsummentests durchgeführt, um allgemeine Gruppenunterschiede festzustellen (Duller, 2008, S.165 ff.). Dies soll zur Untersuchung der Hypothese 1 dienen. Die Variable mit den Noten der Projektarbeiten ist ordinal skaliert. Daher wird zur Überprüfung der Hypothese 2 ein Logit Modell geschätzt. Die Noten werden als dichotome Variable codiert (gute / schlechte Leistung) (Fahrmeir, Kneib und Lang, 2009, S. 235 ff.). Die Aufteilung geschieht anhand des Medians der Noten. Dieser beträgt sowohl in der Gruppe mit LCMS-Betreuung als auch in der Gruppe ohne LCMS-Betreuung jeweils 2,3. Dementsprechend werden Noten besser-gleich 2,3 mit 1 codiert. Das Modell soll den Effekt des Einsatzes des LMCS, aber auch den Effekt der Anzahl der Nachrichten auf die Leistungen im Rahmen der Projektarbeiten erklären. Zur Erklärung werden daher die Nachrichten verwendet, auf die der jeweilige Student zugreifen kann, d.h. die Anzahl der Nachrichten pro Projektarbeitsgruppe plus die Anzahl der eMail-Nachrichten. Die Analyse wird mit der Statistiksoftware R durchgeführt.

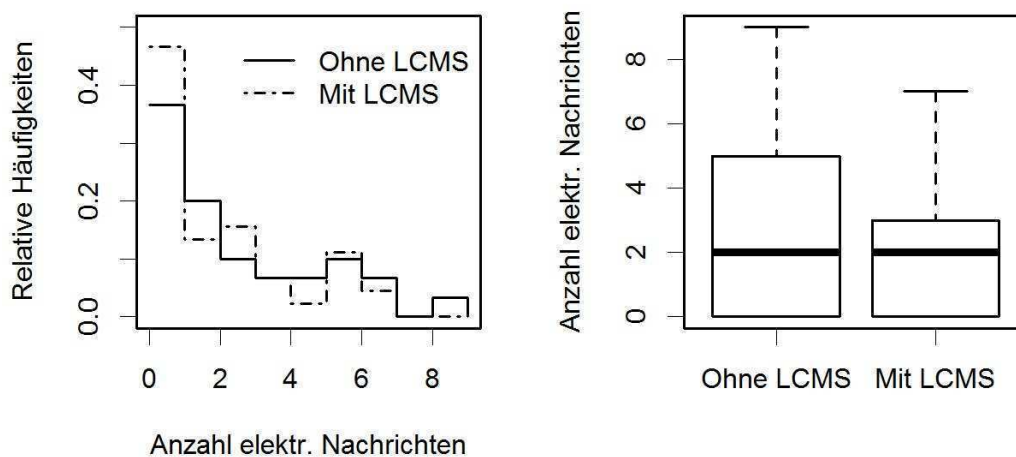


Abbildung 7.18: Histogramm der Häufigkeiten in Relation zur Anzahl der Studenten und Boxplots der Anzahl der elektronischen Nachrichten.

## Quantitative Untersuchung zur Effizienz von eLearning

Nullhypothese	Alternative	Emp. Prüfgröße	p-Wert
Die zentrale Tendenz der erzielten Noten ist in den Gruppen mit und ohne LCMS gleich.	Die mittels LCMS betreute Gruppe hat im Mittel höhere Noten erzielt.	W = 688	0,4456
Die zentrale Tendenz der Anzahl der geschriebenen Nachrichten ist in den Gruppen mit und ohne LCMS gleich.	In der Gruppe ohne LCMS-Betreuung wurden im Mittel mehr Nachrichten je Student versendet.	W=728,5	0,2797

Tabelle 7.28: Ergebnisse Wilcoxon-Rangsummentest auf Gruppen-Unterschiede.

## 7.4 Ergebnisse

Die Boxplots in Abbildung 7.18 zeigen bereits, dass kein signifikanter Gruppenunterschied für die Anzahl der geschriebenen Nachrichten existiert. Das wird vom Wilcoxon-Rangsummentest bestätigt. Auch kann kein statistisch signifikanter Unterschied der Gruppen in der Tendenz der erreichten Noten in den Projektarbeiten nachgewiesen werden. Die Prüfgrößen und zugehörigen p-Werte sind in Tabelle 7.28 aufgeführt. Tabelle 7.29 enthält die Ergebnisse einer Logit-Regression zur Erklärung der erreichten Noten. Die Variable LCMS hat ein positives Vorzeichen, d.h. die Betreuung durch ein LCMS vergrößert die Wahrscheinlichkeit, eine gute Leistung zu erzielen. Dieses Ergebnis ist statistisch nicht signifikant. Tabelle 7.29 zeigt einen negativen Einfluss der Anzahl der geschriebenen Nachrichten auf die Wahrscheinlichkeit, eine gute Note zu erzielen. Dieser Einfluss ist wiederum statistisch nicht signifikant. Ursache dafür ist, dass die bedingte Wahrscheinlichkeit einer guten Note nicht linear mit der Anzahl der Nachrichten variiert, sondern einem zyklischen Verlauf folgt (wie in Abbildung 7.19 zu sehen ist). Es ist zu vermuten, dass der zyklische Verlauf der bedingten Wahrscheinlichkeit aufgrund von latenten Gruppen innerhalb der Studierenden zustande kommt. Diese Gruppen zu identifizieren könnte eine gezielte Förderung ermöglichen.

	Schätzer	Std. Fehler	z Wert	p-Wert
Konstante	-0,44629	0,45049	-0,991	0,322
Nachrichten insgesamt	-0,03714	0,09122	-0,407	0,684
LCMS (=1 bei LCMS-Betreuung)	0,04749	0,48846	0,097	0,923
AIC		104,93		

Tabelle 7.29: Ergebnisse Logit-Modell zur Erklärung der dichotom codierten Noten.  
Abhängige Variable =1, wenn beobachtete Note besser-gleich 2,3 ist.

## 7.5 Fazit

Die empirischen Ergebnisse zeigen, dass das hier verwendete eLearning-Szenario keinen signifikanten Unterschied bezüglich der erzielten Noten und der Anzahl der geschriebenen Nachrichten (Betreuungsaufwand) zu herkömmlichen Lehrformen liefert. Das geht einher mit dem bereits aus früheren Untersuchungen bekannten „no significant difference“ Phänomen (Twigg, 2003). Dass der Betreuungsaufwand bei online Betreuung im Allgemeinen nicht sinkt, ist bereits mehrfach ohne empirische Belege in der Literatur erwähnt. Aus Sicht der Lehrenden ist die persönliche Betreuung der Studierenden mittels Kommunikation über Lernplattformen ein hoher Aufwand, welcher die in anderen Bereichen der Lehre durch eLearning erreichten Skaleneffekte zu Nichte machen kann (Kuhn, 2009; Laurillard, 2007; Lehner, Schäfer und Proksch, 2003).

Obwohl keine Anhebung des Notenschnitts durch Einführung des LCMS erreicht wurde, ist davon auszugehen, dass die Studierenden einen Nutzengewinn erfahren. Es muss bedacht werden, dass eine gute Note keine Aussage über die Lernzufriedenheit der Studenten macht und die Note allein zu wenig Information über den Gesamtnutzen liefert (Preussler und Baumgartner, 2006). Da eine Befragung aller Studenten nicht möglich war, wurden nur Studenten, die mit LCMS betreut worden sind, zum Abschluss des Kurses zum eLearning-Einsatz befragt. Die Befragung hat ergeben, dass die Studierenden das LCMS beibehalten möchten und den eigenen Lernerfolg als positiv beeinflusst einschätzen<sup>3</sup>. Das unterstützt die

<sup>3</sup>Alle befragten Studenten halten die Fortführung des Angebots der Lernplattform für *eher empfehlenswert* bis *sehr empfehlenswert*. Alle befragten Studenten bewerten die Beeinflussung des eigenen Lernerfolgs durch das LCMS als *eher positiv* bis *positiv*. Die Antworten konnten auf

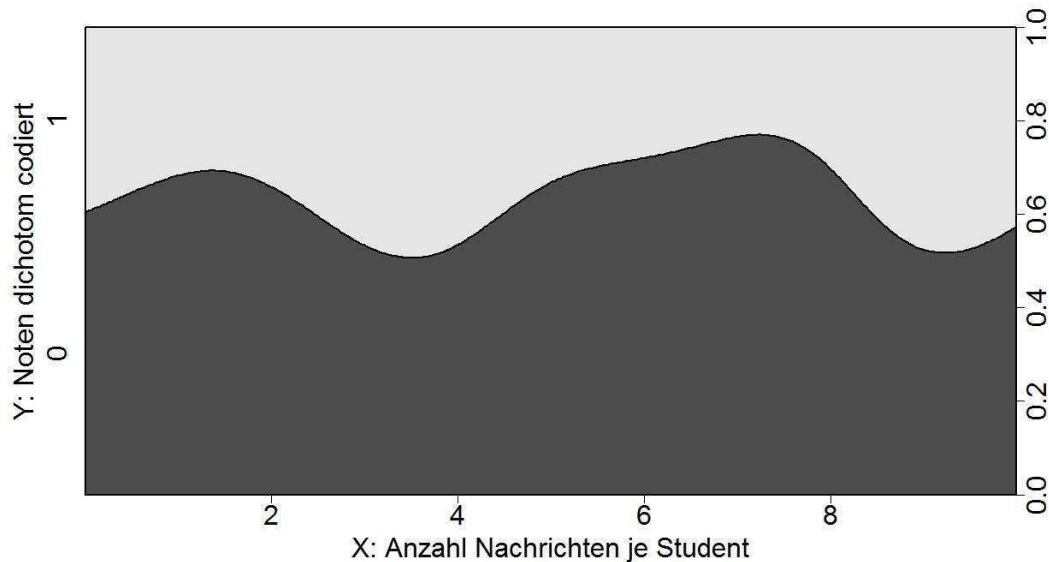


Abbildung 7.19: Bedingte Wahrscheinlichkeit  $P(Y|X)$ . Die dunkle Fläche gibt die bedingte Wahrscheinlichkeit an, eine Note schlechter als 2,3 zu erzielen.

Hypothese 2 zum Teil. Allerdings wird die studentische Einschätzung des eigenen Lernerfolgs eben nicht durch tatsächlich bessere Noten unterlegt. Die Befragung musste aus datenschutzrechtlichen Gründen anonym durchgeführt werden, so dass keine Verknüpfung zum Betreuungsaufwand durch die Anzahl der geschriebenen Nachrichten ermittelt werden konnte.

Trotz der Mängel der Modellierung können wir empirisch zeigen, dass eLearning-Einsatz nicht per se zu einer Vereinfachung des Betreuungsaufwand führen muss. Die Hypothese 1 wird nicht durch die Daten unterstützt. Dieses Ergebnis ist nicht verallgemeinerbar, da der Betreuungsaufwand sehr stark vom jeweiligen eLearning-Szenario abhängt (Bremer, Krömker und Voß, 2010). Da der Einsatz von eLearning neben der Betreuung auch die Erstellung und Aktualisierung der LCMS-Inhalte bedeutet, tritt insgesamt ein Mehraufwand gegenüber herkömmlicher Lehre auf. Die vorliegenden Daten untermauern damit die Studie von Steffens (Steffens und Reiss, 2010). Basierend auf diesen Ergebnissen ist für die Zukunft die Beantwortung von zwei Fragen wichtig: 1) Wie können Anreizmechanismen für

---

5-stufigen Likert-Skalen gegeben werden.



## 7.5 Fazit

Lehrende aussetzen, damit weiterhin eLearning (auch nachhaltig) eingesetzt wird, trotz der für sie ausbleibenden Nutzenzugewinne. Das könnte zum Beispiel in Form der Anrechnung aufs Lehrdeputat sein. 2) Wie können eLearning-Projekte gestaltet werden, damit Kosten für Lehrende eingespart werden bei gleichem oder höherem Nutzen für die Studierenden. Zur Beantwortung dieser Frage, wäre ein praxisorientierter Ausbau der einfachen Messung von Kosten und Nutzen sinnvoll. Die oben erwähnten Mängel des hier vorgestellten Vorgehens müssen behoben werden.



# Literatur

- Bensberg, Frank, Heinz Lothar Grob und Jan vom Brocke (2007). »Kostenorientiertes Controlling von E-Learning-Plattformen mit dem TCO-Konzept - Methodische Grundlagen, Softwareunterstützung und Entwicklungsperspektiven«. In: *eLearning and education* 3. Besucht am 03.02.2011. URL: <http://elearn.campussource.de/archive/3/1080/>.
- Bremer, Claudia, Detlef Krömker und Sarah Voß (2010). »Wirtschaftlichkeits- und Wirksamkeitsanalysen sowie Vorgehensmodelle zur Einführung und Umsetzung von eLearning an Hochschulen«. In: *E-Learning in der Hochschule und Weiterbildung: Einsatzchancen und Erfahrungen*. Hrsg. von R. Holten und D. Nittel. Bertelsmann, S. 61–80.
- Brocke, Jan vom u. a. (2007). »Gestaltung und Bewertung von E-Learning Geschäftsmodellen - Ein Vorgehensmodell am Fallbeispiel e-teaching.org«. In: *Zeitschrift für e-learning* 2(3), S. 7–18.
- Brosowski, Thomas und Jürgen Cleve (2009). »Qualitatives E-Learning versus Machbarkeit - Strategien der Fakultät für Wirtschaftswissenschaften«. In: *E-Learning-Inhalte: effizient entwickeln und effektiv einsetzen. Tagungsband zur AKWI-Fachtagung am 13. bis 15.09.2009*. Hrsg. von T. Barton u. a. Besucht am 03.02.2011. Verlag News & Media, S. 19–26. URL: <http://www.fh-worms.de/fileadmin/medien/wsinf/Professoren/barton/AKWI/AKWI-Tagungsband2009.pdf>.
- Ćukušić, Maja u. a. (2010). »E-Learning process management and the e-learning performance: Results of a European empirical study«. In: *Computers & Education* 55, S. 554–565.
- Duller, Christine (2008). *Einführung in die nichtparametrische Statistik mit SAS und R: Ein anwendungsorientiertes Lehr- und Arbeitsbuch*. Heidelberg: Physica.
- Fahrmeir, Ludwig, Thomas Kneib und Stefan Lang (2009). *Regression: Modelle, Methoden und Anwendungen*. Berlin: Springer.

## Literatur

- Figlio, David N., Mark Rush und Lu Yin (2010). *Is it live or is it internet? Experimental estimates of the effects of online instruction on student learning*. NBER Working Paper 16089. Besucht am 21.06.2010. NBER Working Paper Series. URL: <http://www.nber.org/papers/w16089>.
- Kuhn, Elvira (2009). »E-Learning - Innovativer Beitrag zur Steigerung der Qualität der Lehre?« In: *E-Learning-Inhalte: effizient entwickeln und effektiv einsetzen. Tagungsband zur AKWI-Fachtagung am 13. bis 15.09.2009*. Hrsg. von T. Barton u. a. Besucht am 03.02.2011. Verlag News & Media, S. 27–44. URL: <http://www.fh-worms.de/fileadmin/medien/wsinf/Professoren/barton/AKWI/AKWI-Tagungsband2009.pdf>.
- Laurillard, Diana (2007). »Modelling benefits-oriented costs for technology enhanced learning«. In: *Higher Education* 54, S. 21–39.
- Lehner, Franz, Klaus J. Schäfer und Marianne Proksch (2003). »Was kostet E-Learning?« In: *DeLFI 2003, Tagungsband der 1. E-Learning Fachtagung Informatik, 16.-18.9.2003 in Garching bei München, Lecture Notes in Informatics vol. 37*. Hrsg. von A. Bode u. a. Gesellschaft für Informatik, S. 240–249.
- Milliken, John und L.Philip Barnes (2002). »Teaching and technology in higher education: student perceptions and personal reflections«. In: *Computers & Education* 39, S. 223–235.
- Preussler, Annabell und Peter Baumgartner (2006). »Qualitätssicherung in medienegestützten Lernprozessen - zur Messproblematik von theoretischen Konstrukten«. In: *Qualitätssicherung im E-Learning*. Hrsg. von A. Sindler u. a. Waxmann, S. 73–85.
- Saghafian, Marzieh (2011). »A critical review of research on technology-based training in business organizations«. In: *Research in Higher Education Journal* 10. Besucht am 29.03.2011. URL: <http://www.aabri.com/manuscripts/10632.pdf>.
- Schwarz, Christine (2001). »eLearning und Bildungspolitik: Von der Nachhaltigkeit hoher Erwartungen«. In: *Virtueller Campus. Szenarien - Strategien - Studium*. Hrsg. von E. Wagner und M. Kindt. Waxmann, S. 374–384.
- Selim, Hassan M. (2007). »Critical success factors for e-learning acceptance: Confirmatory factor models«. In: *Computers & Education* 49, S. 396–413.
- Steffens, Dirk und Michael Reiss (2010). »Performance of Blended Learning in University Teaching: Determinants and Challenges«. In: *eLearning and education*

## *Literatur*

6. Besucht am 03.02.2011. URL: <http://eeced.campussource.de/archive/6/2627/>.

Twigg, Carol A. (2003). »Improving learning and reducing costs: new models for online learning«. In: *EDUCAUSE Review Magazine* 38, S. 28–38.



# Kapitel 8

## Tobit Modell

Das Tobit Modell ist ein Regressionsmodell für beschränkte Variablen. Die abhängige Variable ist aufgrund einer Zensierung nur innerhalb eines bestimmten Bereichs zu beobachten. Das zensierte Regressionsmodell und die Schätzung der zugehörigen Parameter wurden von Tobin (1958) zum ersten Mal diskutiert. (Maddala, 1983, S. 151) Tobin entwickelte das Modell, um Ausgaben der Haushalte für Gebrauchsgüter, wie zum Beispiel ein Auto, zu erklären. Dabei sollte das Problem berücksichtigt werden, dass viele Haushalte keine Ausgaben in einem bestimmten Jahr diesbezüglich haben, andere Haushalte hingegen eine große Variation in den Ausgaben vorweisen. Die Ausgaben sind dementsprechend durch die Null nach unten beschränkt. (Tobin, 1958, S. 25). In den Kapiteln 3 *Tobit regression to estimate impact of EU market intervention in the German butter market* und 4 *Beurteilung der Kompensation auf dem EU-Milchmarkt auf Basis eines Tobit-Ansatzes* in dieser Arbeit wird das Tobit Modell verwendet zur Erklärung der Wirkung der EU Interventionspolitik auf den Märkten für Butter und Magermilchpulver. Die Verwendung des Tobit Modells ist in diesen Fällen notwendig, da die jeweiligen Marktpreise für Butter und Magermilchpulver durch die politisch festgelegten Interventionspreise nach unten beschränkt sind.

Im folgenden Kapitel wird das Tobit Modell umfassend dargestellt. Zuerst wird die Modellformulierung von Tobin (1958) erläutert. Im zweiten Abschnitt wird die Interpretation des Modells anhand der Erwartungswerte, marginalen Effekte und der Varianz besprochen. Der darauf folgende Abschnitt zeigt die Probleme der Methode der Kleinsten Quadrate in einem Umfeld beschränkter Daten. Zum einen werden die Probleme anhand eines simulierten Beispiels verdeutlicht, zum anderen aber auch analytisch aufgezeigt. Es zeigt sich dabei, dass die Verwendung

## *Tobit Modell*

der Methode der kleinsten Quadrate nicht passend für diesen Zusammenhang ist. (Amemiya, 1984, S. 5) Daher werden in Abschnitt 4 drei Schätzverfahren vorgestellt: Maximum Likelihood Methode, Heckmann 2-Schritt Verfahren und der Censored Least Absolute Deviations Schätzer. Abschnitt 5 behandelt Fehlspezifikationen im Tobit Modell. Schließlich wird in Abschnitt 6 der Likelihood Ratio Test kurz vorgestellt. Mithilfe dieses Tests können Hypothesen überprüft werden.

### **8.1 Modellformulierung**

Sei  $y$  eine zensierte Variable. In der Anwendung von Tobin (1958) misst die Variable  $y$  die Haushaltsausgaben für Gebrauchsgüter, wie zum Beispiel ein Auto. Bei einer Befragung von Haushalten zu dem Haushaltseinkommen  $x$  und den Ausgaben  $y$  bezüglich eines Autokaufs innerhalb eines bestimmten Zeitraums wird eine große Anzahl der Haushalte Ausgaben in Höhe von Null angeben. Eine Möglichkeit der Untersuchung dieser Ausgaben ist die Schätzung eines Probit Modells, welches die Wahrscheinlichkeiten der Kauf- und Nicht Kauf-Entscheidung erklärt. Dieses Vorgehen ist aber ineffizient, denn es verursacht einen Informationsverlust über die Höhe der möglicherweise gemachten Ausgaben. (Tobin, 1958, S. 25) Eine weitere Möglichkeit ist die Methode der kleinsten Quadrate (OLS) zur Schätzung eines linearen Modells für den Zusammenhang zwischen Haushaltsausgaben  $y$  und dem Haushaltseinkommen  $x$ . Diese Methode ist allerdings unpassend in dieser Situation. Unter anderem unterstellt die OLS Methode einen konstanten marginalen Effekt einer exogenen Variablen  $x$  auf  $y$ . (Wooldridge, 2009, S. 588) Die Abbildung 8.20 zeigt simulierte Daten für ein Beispiel. Wichtig ist dabei, dass es einige Beobachtungen mit Ausgaben in Höhe von Null gibt. Linearität und konstante marginale Effekte zwischen Haushaltsausgaben und Einkommen sind daraus nicht ersichtlich. Auch ist deutlich, dass eine stetige Dichtefunktion nicht geeignet ist, die bedingte Verteilung von Ausgaben bei gegebenem Einkommen aufgrund der Häufung von Beobachtungen der Null zu erklären. (Amemiya, 1984, S. 5) Im dritten Abschnitt dieses Kapitels wird erklärt, dass das OLS Vorgehen zu verzerrten und inkonsistenten Schätzern führt. (Greene, 1981, S. 505)

Das Tobit Modell zur Erklärung von  $y$  geht von einer latenten Variablen  $y^*$  aus. Eine Annahme ist, dass die Variable  $y^*$  einen linearen Zusammenhang mit den



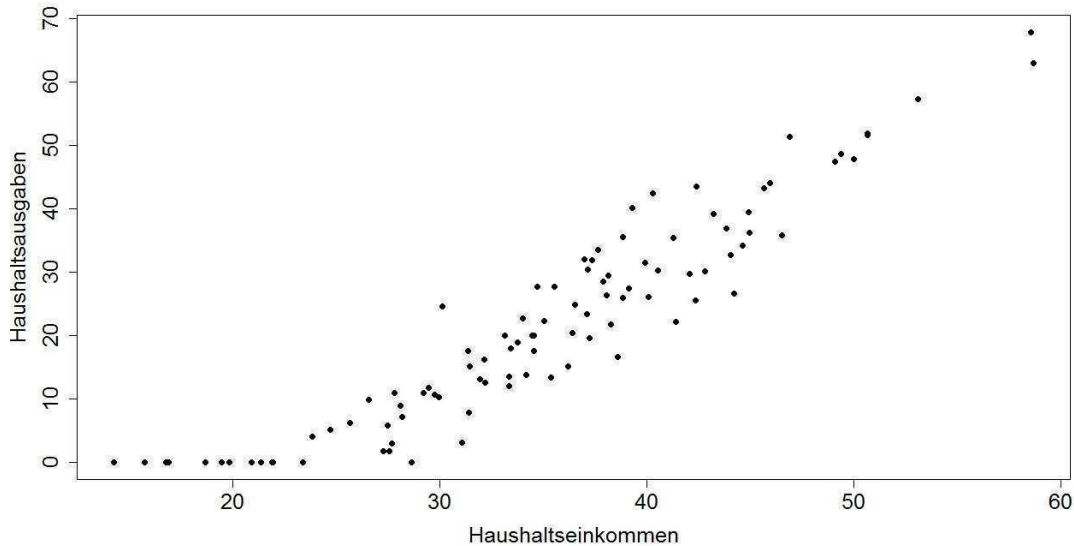


Abbildung 8.20: Beispielhafte Simulation von Haushaltsausgaben und Haushaltseinkommen. Eigene Darstellung in Anlehnung an Amemiya (1984, S. 5).

unabhängigen Variablen  $x_1, \dots, x_m$  besitzt (Tobin, 1958, S. 25f.; Wooldridge, 2009, S. 588):

$$y_i^* = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_m x_{mi} + u_i = \mathbf{x}_i \boldsymbol{\beta} + u_i \quad i = 1, \dots, n. \quad (8.26)$$

Es ist  $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_m)'$  und  $\mathbf{x}$  ist eine Matrix, deren erste Spalte nur 1en enthält. Die weiteren  $m$  Spalten bestehen aus den  $n$  Beobachtungen der  $m$  Variablen, d.h.  $\mathbf{x} = (\mathbf{1}, x_1, \dots, x_m)$ .  $\mathbf{x}_i$  ist dabei die  $i$ -te Zeile der Matrix  $\mathbf{x}$ . Es gilt dabei, dass die  $m$  Variablen  $x_1, \dots, x_m$  für alle  $n$  Haushalte beobachtbar sind. Hingegen gilt für die Variable  $y$

$$y = \begin{cases} y^*, & \text{wenn } y^* > y_0 \\ y_0, & \text{wenn } y^* \leq y_0. \end{cases}$$

Das bedeutet, dass die Variable  $y^*$  nur beobachtbar ist, wenn sie größer als der linksseitige Schwellenwert  $y_0$  ist. In dem Beispiel hier gilt also  $y_0 = 0$ . Eine weitere

## Tobit Modell

Annahme ist, dass die weiteren Unterschiede der Haushalte in ihren Ausgaben  $y$ , die nicht durch exogene Variablen  $x_1, \dots, x_m$  und den Grenzwert  $y_0$  erklärt werden, als zufällig angenommen und durch die normalverteilte Störgröße  $u$  mit Erwartungswert Null und Standardabweichung  $\sigma$  reflektiert werden. (Tobin, 1958, S. 26)

Die Wahrscheinlichkeitsdichte von  $y_i$ ,  $\mathbf{x}_i$  sei gegeben, bestehend aus zwei Teilen. Ein Teil beschreibt die Dichte der zensierten Beobachtungen und der zweite Teil gibt die Dichtefunktion der unzensierten Beobachtungen an. Die Wahrscheinlichkeitsdichte unzensierter Beobachtungen von  $y$  ist identisch zu der Dichte von  $y^*$ . Aufgrund der angenommenen Normalverteilung der Störgröße  $u$  gilt daher

$$P(y = y_i | \mathbf{x}_i) = \frac{1}{\sigma} \phi \left( \frac{y_i - \mathbf{x}_i \boldsymbol{\beta}}{\sigma} \right). \quad (8.27)$$

Die bedingte Wahrscheinlichkeitsdichte für  $y = y_0$  ist außerdem wie folgt gegeben (Wooldridge, 2009, S. 588)

$$\begin{aligned} P(y = y_0 | \mathbf{x}) &= P(y^* < y_0 | \mathbf{x}) \\ &= P(\mathbf{x} \boldsymbol{\beta} + u < y_0 | \mathbf{x}) \\ &= P(u/\sigma < (y_0 - \mathbf{x} \boldsymbol{\beta})/\sigma | \mathbf{x}) \\ &= \Phi((y_0 - \mathbf{x} \boldsymbol{\beta})/\sigma). \end{aligned} \quad (8.28)$$

Dabei sind  $\Phi$  und  $\phi$  jeweils die Verteilungsfunktion und Dichtefunktion der Standardnormalverteilung.

## 8.2 Interpretation des Modells

Zur Interpretation des Tobit Modells sind die drei bedingten Erwartungswerte  $E(y^* | \mathbf{x})$ ,  $E(y | \mathbf{x})$  und  $E(y | \mathbf{x}, y > y_0)$  zu berücksichtigen. Es handelt sich dabei um die Erwartungen bezüglich der latenten Variablen, der Variablen  $y$  (unabhängig ob zensiert oder unzensiert) und der unzensierten Beobachtungen von  $y$ . Die Berechnung der Erwartungswerte ergibt (Greene, 2008, S. 870ff.):

$$E(y^*|\mathbf{x}) = \mathbf{x}\boldsymbol{\beta} \quad (8.29)$$

$$\begin{aligned} E(y|\mathbf{x}, y > y_0) &= E(\mathbf{x}\boldsymbol{\beta} + u|\mathbf{x}, \mathbf{x}\boldsymbol{\beta} + u > y_0) \\ &= \mathbf{x}\boldsymbol{\beta} + E(u|\mathbf{x}, u > y_0 - \mathbf{x}\boldsymbol{\beta}) \\ &= \mathbf{x}\boldsymbol{\beta} + \sigma \left[ \frac{\phi((y_0 - \mathbf{x}\boldsymbol{\beta})/\sigma)}{1 - \Phi((y_0 - \mathbf{x}\boldsymbol{\beta})/\sigma)} \right] \\ &= \mathbf{x}\boldsymbol{\beta} + \sigma \left[ \frac{\phi((y_0 - \mathbf{x}\boldsymbol{\beta})/\sigma)}{\Phi((\mathbf{x}\boldsymbol{\beta} - y_0)/\sigma)} \right] = \mathbf{x}\boldsymbol{\beta} + \sigma \cdot \lambda \end{aligned} \quad (8.30)$$

$$\begin{aligned} E(y|\mathbf{x}) &= P(y = y_0|\mathbf{x}) \cdot y_0 + P(y > y_0|\mathbf{x}) \cdot E(y|\mathbf{x}, y > y_0) \\ &= P(y = y_0|\mathbf{x}) \cdot y_0 + P(y > y_0|\mathbf{x}) \cdot E(y|\mathbf{x}, y^* > y_0) \\ &= \Phi((y_0 - \mathbf{x}\boldsymbol{\beta})/\sigma) \cdot y_0 + (1 - \Phi((y_0 - \mathbf{x}\boldsymbol{\beta})/\sigma)) \cdot (\mathbf{x}\boldsymbol{\beta} + \sigma \cdot \lambda) \\ &= (1 - \Phi((\mathbf{x}\boldsymbol{\beta} - y_0)/\sigma)) \cdot y_0 \\ &\quad + \Phi((\mathbf{x}\boldsymbol{\beta} - y_0)/\sigma) \cdot \mathbf{x}\boldsymbol{\beta} + \sigma \cdot \phi((\mathbf{x}\boldsymbol{\beta} - y_0)/\sigma) \end{aligned} \quad (8.31)$$

$$\text{mit } \lambda = \frac{\phi((y_0 - \mathbf{x}\boldsymbol{\beta})/\sigma)}{\Phi((\mathbf{x}\boldsymbol{\beta} - y_0)/\sigma)}. \quad (8.32)$$

Für Vorhersagen von  $y_i$  bei gegebenem  $\mathbf{x}_i$  ist eine der drei Funktionen des Erwartungswerts zu benutzen. Intuitiv sollte dies (8.31) sein, doch die Wahl des Erwartungswertes hängt auch von der Fragestellung der Untersuchung ab. Wird ein Tobit Modell zur Erklärung von Kartenverkäufen in einem Stadion genutzt, so ist die Variable *verkaufte Eintrittskarten*  $y$  nach oben durch die maximale Anzahl an Plätzen in dem Stadion begrenzt. Für die Fragestellung, wie viele Karten bei einer zukünftigen Veranstaltung verkauft werden, wird dann (8.31), unter Anpassung des Grenzwertes als obere Schranke, verwendet. Um die Frage der Notwendigkeit einer Stadionvergrößerung zu beantworten ist hingegen der Mittelwert der latenten Variablen von Interesse. (Greene, 2008, S. 872) Es sollte auf jeden Fall klar sein, dass der Mittelwert aller **beobachteten**  $y$  Werte durch (8.31) und nicht durch (8.29) (Mittelwert der **potentiellen**  $y$  Werte) gegeben ist. (Maddala, 1983, S. 160)

Gemäß den drei Erwartungswerten gibt es auch die dazugehörigen Ableitungen zur Abschätzung der Auswirkungen von Änderungen in den unabhängigen Variablen. Es gilt für die marginalen Effekte (Greene, 2008, S. 872f.; Wooldridge, 2009, S. 590f.):

$$\frac{\partial E(y^*|\mathbf{x})}{\partial x_j} = \beta_j \quad (8.33)$$

$$\begin{aligned} \frac{\partial E(y|\mathbf{x}, y > y_0)}{\partial x_j} &= \beta_j + \sigma \cdot \frac{\partial \lambda}{\partial x_j} \\ &= \beta_j \cdot \left( 1 - \lambda \cdot \left( \frac{\mathbf{x}\boldsymbol{\beta} - y_0}{\sigma} + \lambda \right) \right) \end{aligned} \quad (8.34)$$

$$\begin{aligned} \frac{\partial E(y|\mathbf{x})}{\partial x_j} &= \frac{\partial \Phi((\mathbf{x}\boldsymbol{\beta} - y_0)/\sigma)}{\partial x_j} \cdot (\mathbf{x}\boldsymbol{\beta} - y_0) \\ &+ \Phi\left(\frac{\mathbf{x}\boldsymbol{\beta} - y_0}{\sigma}\right) \cdot \beta_j + \sigma \cdot \frac{\partial \phi((\mathbf{x}\boldsymbol{\beta} - y_0)/\sigma)}{\partial x_j} \\ &= \Phi\left(\frac{\mathbf{x}\boldsymbol{\beta} - y_0}{\sigma}\right) \cdot \beta_j. \end{aligned} \quad (8.35)$$

Der marginale Effekt einer exogenen Variablen  $x_j$  besitzt immer das Vorzeichen von  $\beta_j$ , doch hängt die Größenordnung der jeweiligen Auswirkung bei (8.34) und (8.35) von allen erklärenden Variablen und der Standardabweichung des Störterms ab. Damit sind die marginalen Effekte (8.34) und (8.35) zwar schwieriger als im Fall eines einfachen linearen Modells, dafür ist der Erwartungswert (8.31) aber auf jeden Fall größer-gleich  $y_0$  und auch die Nichtlinearität von  $y$  wird berücksichtigt. (Wooldridge, 2009, S. 590) Das entspricht der Berücksichtigung der Eigenschaften von  $y$ , wie sie bereits im Rahmen des Beispiels der Erklärung der Haushaltsausgaben in Abbildung 8.20 zu sehen waren.

Nicht nur auf den Erwartungswert hat die Zensierung eine Auswirkung. Auch die Varianz von  $y$  wird durch die Zensierung beeinflusst. Die Varianz ist dabei wiederum davon abhängig, ob nur unzensierte  $y$  Werte im Fokus sind, oder alle beobachtbaren Werte von  $y$ . Die Varianzen lauten (Greene, 2008, S. 866; Kim und Chavas, 2002, S. 498; Chavas und Kim, 2006, S. 69)

### 8.3 Probleme der OLS Schätzung bei zensierten Daten

$$Var(y|\mathbf{x}, y > y_0) = \sigma^2 \cdot \left[ 1 - \lambda \cdot \left( \lambda - \frac{y_0 - \mathbf{x}\boldsymbol{\beta}}{\sigma} \right) \right] \quad (8.36)$$

$$\begin{aligned} Var(y|\mathbf{x}) = \sigma^2 \cdot & \left[ 1 - \Phi \left( \frac{y_0 - \mathbf{x}\boldsymbol{\beta}}{\sigma_t} \right) + \frac{y_0 - \mathbf{x}\boldsymbol{\beta}}{\sigma_t} \cdot \phi \left( \frac{y_0 - \mathbf{x}\boldsymbol{\beta}}{\sigma_t} \right) \right. \\ & + \left( \frac{y_0 - \mathbf{x}\boldsymbol{\beta}}{\sigma_t} \right)^2 \cdot \Phi \left( \frac{y_0 - \mathbf{x}\boldsymbol{\beta}}{\sigma_t} \right) \\ & \left. - \left[ \frac{y_0 - \mathbf{x}\boldsymbol{\beta}}{\sigma_t} \cdot \Phi \left( \frac{y_0 - \mathbf{x}\boldsymbol{\beta}}{\sigma_t} \right) + \phi \left( \frac{y_0 - \mathbf{x}\boldsymbol{\beta}}{\sigma_t} \right) \right]^2 \right]. \quad (8.37) \end{aligned}$$

Chavas und Kim (2002, 2006) leiten aus (8.37) die relative Varianz  $V(y)/\sigma^2$  her. Sie gibt die Varianz von  $y$  in Relation zur Modellvarianz der latenten linearen Beziehung von  $y^*$  zu den exogenen Variablen an. Damit ist  $V(y)/\sigma^2$  ein Maß für den Zensierungseffekt. Liegt keine Zensierung vor, so ist die Relation 1. Durch Zensierung nimmt die relative Varianz gemäß der Reduktion der Volatilität in  $y$  ab. (Kim und Chavas, 2002, S. 499; Chavas und Kim, 2006, S. 69)

## 8.3 Probleme der OLS Schätzung bei zensierten Daten

Das Beispiel der Haushaltsausgaben und der Zusammenhang mit dem Haushaltseinkommen kann auch vor dem Hintergrund einer Nutzenmaximierung durch den Haushalt beschrieben werden. Dafür seien

- $y$  Haushaltsausgaben für ein Gebrauchsgut,
- $y_0$  der Preis des günstigsten Gebrauchsgutes,
- $z$  weitere Haushaltsausgaben,
- $x$  Haushaltseinkommen.

Der Haushalt versucht, den Nutzen  $U(y, z)$  unter der Budgetbedingung  $y + z \leq x$  und den Grenzen  $y \geq y_0$  oder  $y = 0$  zu maximieren.  $y^*$  löst dann das Maximie-

## Tobit Modell

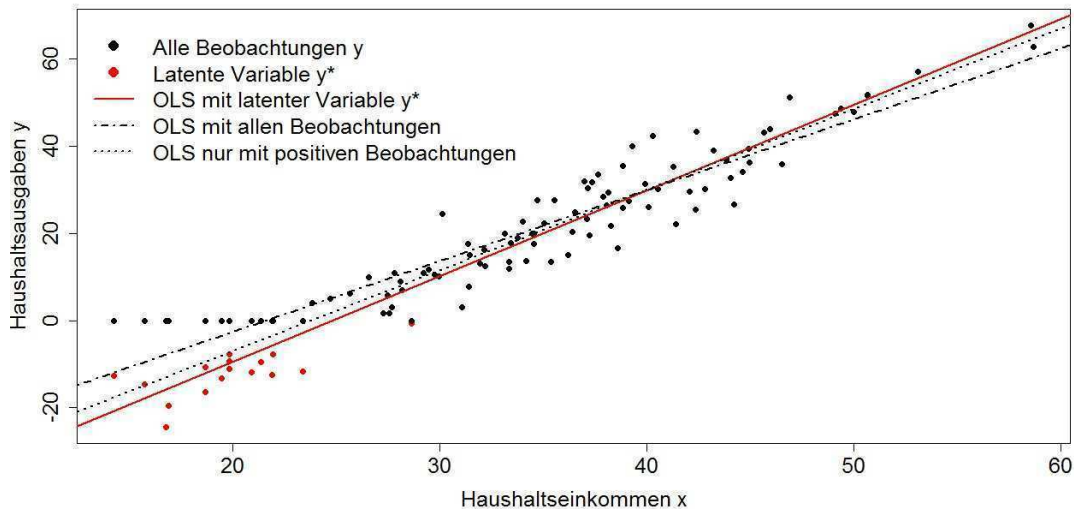


Abbildung 8.21: Erweiterung der beispielhaften Simulation von Haushaltsausgaben und Haushaltseinkommen aus Abbildung 8.20. Es werden die zugehörigen (latenten) Beobachtungen  $y^*$ , sowie OLS Schätzungen des linearen Zusammenhangs hinzugefügt. Die OLS Schätzung des Zusammenhangs der latenten Variablen  $y^*$  und  $x$  ist als wahrer Zusammenhang anzunehmen.

rungsproblem unter Berücksichtigung des Budgets, aber unter Vernachlässigung der Beschränkung von  $y$  durch die unteren Grenzen. (Amemiya, 1984, S. 5f.)

Die grafische Darstellung des Beispiels in Abbildung 8.20 zeigt nur die beobachteten (zensierten) Haushaltsausgaben. In Abbildung 8.21 werden zusätzlich die latenten Ausprägungen von  $y^*$  gezeigt. Gemäß unserer Annahme folgen die unzensierten  $y^*$  einem positiven linearen Zusammenhang mit dem Haushaltseinkommen  $x$ . Wie kann dieser lineare Zusammenhang ermittelt werden, wenn nur die beobachteten  $y$  Werte zur Verfügung stehen?

Unter der Verwendung des OLS Schätzers gibt es zwei mögliche Herangehensweisen. Zum einen kann die Schätzung auf Grundlage der unzensierten Beobachtungen ( $y > y_0$ ) durchgeführt werden, zum anderen basierend auf allen Beobachtungen  $y$ . Tabelle 8.30 zeigt die Schätzergebnisse, die zusätzlich in Abbildung 8.21 eingezeichnet sind.

### 8.3 Probleme der OLS Schätzung bei zensierten Daten

$y^*$	wahrer		OLS	
	Zusammenhang	Tobit	( $y > y_0$ )	(alle Beobachtungen)
(Intercept)	-48,809 (1,99)	-46,526 (2,65)	-44,073 (2,91)	-35,099 (2,11)
Haushaltseinkommen	1,967 (0,06)	1,911 (0,07)	1,853 (0,08)	1,628 (0,06)

Tabelle 8.30: Schätzer des linearen Zusammenhangs der simulierten Daten aus Abbildung 8.21. Zusätzlich sind die Schätzer des Tobit Modells angegeben. Alle Schätzer sind statistisch signifikant zum Signifikanzniveau 0.

Es ist zu sehen, dass die OLS Schätzer starke Abweichungen aufweisen. Das kann auch analytisch gezeigt werden. Für den Fall, dass nur die unzensierten Beobachtungen ( $y > y_0$ ) in der OLS Schätzung verwendet werden, so ist der zugehörige Erwartungswert durch Gleichung (8.30) gegeben. Gleichung (8.30) zeigt, dass der Erwartungswert für die unzensierten  $y$  gleich  $\mathbf{x}\boldsymbol{\beta}$  ist, plus der positive Ausdruck  $\sigma \cdot \lambda$ . Die OLS Schätzung berücksichtigt aber nur  $\mathbf{x}\boldsymbol{\beta}$ . Das Weglassen von  $\lambda$  führt zu einem *Omitted Variable Bias*. (Wooldridge, 2009, S. 590)  $\sigma \cdot \lambda$  wird bei der Schätzung in den Störterm der linearen Gleichung aufgenommen. Eine Standardannahme der OLS Schätzung ist die Unkorreliertheit der Störgröße mit exogenen Variablen. Das ist hier aber nicht gewährleistet, da  $\lambda$  von  $\mathbf{x}$  abhängt. Damit wird die OLS Schätzung keinen Parameter konsistent schätzen. (Wooldridge, 2010, S. 66)

Für den zweiten Fall, bei dem alle  $y$  Beobachtungen, zensiert und unzensiert, für die OLS Schätzung verwendet werden, konnte Greene (1981) zeigen, dass für die OLS Schätzer der Steigungsparameter  $\hat{\beta}_{OLS}$  gilt (Greene, 1981, S. 507)<sup>4</sup>

$$plim \hat{\beta}_{OLS} = \beta \cdot \Phi(\mu^*/\sigma^*). \quad (8.38)$$

Gleichung (8.31) zeigt, dass  $E(y|\mathbf{x})$  nicht linear in  $\mathbf{x}$  und  $\boldsymbol{\beta}$  ist. Gemäß (8.38) verzerrt die dennoch angewendete OLS Methode alle Schätzer Richtung Null im gleichen Verhältnis. (Greene, 1981, S. 513)  $\Phi(\mu^*/\sigma^*)$  ist die Wahrscheinlichkeit, bei einer zufällig gezogenen Beobachtung  $y_i^*$  aus der Normalverteilung  $N(\mu^*, \sigma^{*2})$  eine unzensierte Beobachtung zu erhalten. Die Wahrscheinlichkeit kann durch das Verhältnis  $n_1/n$ , mit  $n_1$  als die Anzahl aller unzensierten Beobachtung, angegeben

<sup>4</sup>Unter der Annahme  $y^* \sim N(\mu^*, \sigma^{*2})$  und ebenso normalverteilter exogener Regressoren.

## Tobit Modell

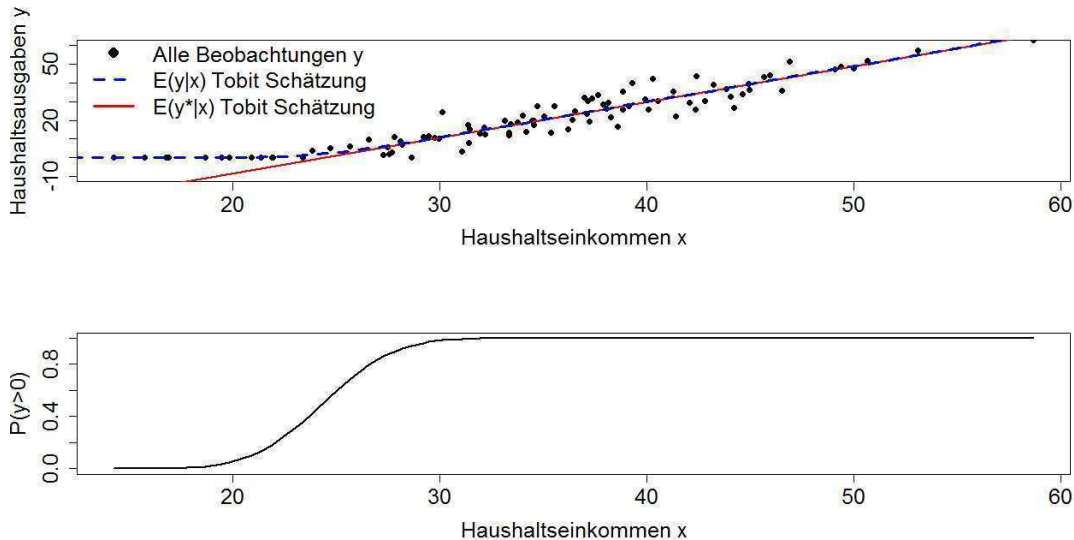


Abbildung 8.22: Beispielhafte Simulation von Haushaltsausgaben und -einkommen mit zugehörigen Erwartungswerten  $E(y|\mathbf{x})$  und  $E(y^*|\mathbf{x})$ . Die Tobit Parameter sind mit der Maximum Likelihood Methode geschätzt worden.

werden. Die OLS Schätzer einfach durch das Verhältnis  $n_1/n$  zu teilen, führt aber nur zu guten Schätzern, wenn auch die exogenen Regressoren  $\mathbf{x}$  auf jeden Fall einer Normalverteilung folgen. (Amemiya, 1984, S. 11)

OLS als Schätzmethode im Kontext eines Tobit Modells ist gemäß (8.38) nur konsistent, wenn  $\Phi(\mu^*/\sigma^*) \rightarrow 1$ , d.h. zensierte Beobachtungen nicht auftreten. Dann liegt ein klassisches lineares Regressionsmodell vor. (Greene, 1981, S. 507) Die Verwendung einer Tobit Formulierung ist dann allerdings nicht mehr zu rechtfertigen. (Wooldridge, 2009, S. 591) Das Ergebnis (8.38) konnte von Chung und Goldberger (1984) auf den Fall nicht notwendigerweise normalverteilter Variablen verallgemeinert werden. (Amemiya, 1985, S. 472)

In Tabelle 8.30 sind zum Vergleich die mittels Maximum Likelihood Methode geschätzten Parameter des Tobit Modells dargestellt. Abbildung 8.22 zeigt den dazugehörigen berechneten Erwartungswert  $E(y|\mathbf{x})$  gemäß (8.31) für die simulierten Beobachtungen. Der untere Teil der Abbildung gibt die Wahrscheinlichkeit einer unzensierten Beobachtung  $P(y > 0) = \Phi((\mathbf{x}\boldsymbol{\beta} - 0)/\sigma)$  an. Es ist zu sehen, dass der marginale Effekt (8.35) sich für  $P(y > 0) \rightarrow 1$  dem konstanten marginalen



Effekt  $\hat{\beta}_1 = 1,911$  annähert. Der marginale Effekt kann mittels Produktregel aus (8.31) als sogenannte McDonald-Moffit Zerlegung dargestellt werden (Hill, Griffiths und Lim, 2008, S. 446; Greene, 2008, S. 873; Wooldridge, 2009, S. 591; Wooldridge, 2010, S. 674):

$$\frac{\partial E(y|\mathbf{x})}{\partial x_j} = P(y > 0) \cdot \frac{\partial E(y|\mathbf{x}, y > 0)}{\partial x_j} + E(y|\mathbf{x}, y > 0) \cdot \frac{\partial P(y > 0)}{\partial x_j}.$$

Eine Änderung einer exogenen Variablen hat demnach nicht nur einen Effekt auf den Erwartungswert von  $y$  im unzensierten Teil der Verteilung, sondern auch Auswirkungen auf die Wahrscheinlichkeit, dass die zugehörige Beobachtung zu dem unzensierten Teil der Beobachtung gehören wird. (Greene, 2008, S. 873) Zusammenfassend ist zu sagen, dass weder  $E(y|\mathbf{x}, y > y_0)$  noch  $E(y|\mathbf{x})$  lineare Funktionen sind. Sie entsprechen nicht  $\mathbf{x}\boldsymbol{\beta}$  und haben keine konstanten marginalen Effekte. Die OLS Methode ist für den Fall zensierter Beobachtungen inkonsistent.

## 8.4 Schätzverfahren

Der vorhergehende Abschnitt hat gezeigt, dass OLS nicht geeignet ist, die Parameter des Tobit Modells fehlerfrei zu schätzen. In diesem Abschnitt werden konsistente Schätzverfahren vorgestellt. Zuerst wird auf die Maximum Likelihood (ML) Methode, die von Tobin (1958) verwendete wurde, eingegangen. Als weitere Möglichkeit wird dann das zweistufige Schätzverfahren von Heckmann erklärt. In der ersten Stufe des Schätzers wird ein Probit Modell mittels ML geschätzt. Bezüglich der Rechenkomplexität ist dieses Verfahren daher im Standard Tobit Modell nicht vorteilhafter als der effiziente ML Schätzer. Doch in komplexeren Tobit Modellformulierungen spielt der 2-Schritt Schätzer zur Vereinfachung der Berechnung eine Rolle. (Amemiya, 1984, S. 12) Als dritte Methode wird das robuste Censored Least Absolute Deviations (CLAD) Verfahren erläutert.

### 8.4.1 Maximum Likelihood

Das Vorgehen der Maximum Likelihood Schätzung basiert auf der Idee, dass diejenigen Parameter zu finden sind, für die die Wahrscheinlichkeit des Auftretens

## Tobit Modell

der Beobachtungen maximal ist. (Komlos und Süßmuth, 2010, S. 122) Da den Beobachtungen Unabhängigkeit unterstellt wird, ist die gemeinsame Verteilung der  $n$  Beobachtungen das Produkt ihrer Dichtefunktionen. (Wooldridge, 2009, S. 761) Gemäß (8.28) und (8.27) ist die Likelihood Funktion (Amemiya, 1984, S. 7)

$$L = \prod_{y_i^* > y_0} \frac{1}{\sigma} \phi\left(\frac{y_i - \mathbf{x}_i \boldsymbol{\beta}}{\sigma}\right) \cdot \prod_{y_i^* \leq y_0} \Phi\left(\frac{y_0 - \mathbf{x}_i \boldsymbol{\beta}}{\sigma}\right). \quad (8.39)$$

(8.39) gibt die Wahrscheinlichkeit an, genau die Beobachtungen  $y_1, \dots, y_n$  zu erhalten. (Greene, 2008, S. 485) Die Likelihood ist eine Funktion der unbekanntem Modellparameter. Der Maximum Likelihood Schätzer ist für die Parameter so zu wählen, dass er die Likelihood Funktion maximiert. In der Regel wird die Log-Likelihood Funktion maximiert, da sie technisch häufig einfacher zu handhaben ist. Die Log-Likelihood Funktion entsteht durch logarithmieren der Likelihood Funktion. (Fahrmeir, Kneib und Lang, 2009, S. 467f.) Die notwendige Bedingung für ein Maximum ist, dass die erste Ableitung der Log-Likelihood Funktion gleich Null ist. (Greene, 2008, S. 486)

Unter Verwendung der Indikator Variable  $D_i = 1$ , wenn  $y_i > y_0$ , ergeben sich die Log-Likelihood Funktion und die notwendigen Bedingungen für ein Maximum wie folgt (Wooldridge, 2009, S. 588; Wooldridge, 2010, S. 676; Maddala, 1983, S. 153)

$$\log L = \sum_{i=1}^n \left[ D_i \cdot \log\left(\frac{1}{\sigma} \phi\left(\frac{y_i - \mathbf{x}_i \boldsymbol{\beta}}{\sigma}\right)\right) + (1 - D_i) \log\left(\Phi\left(\frac{y_0 - \mathbf{x}_i \boldsymbol{\beta}}{\sigma}\right)\right) \right] \quad (8.40)$$

$$\frac{\partial \log L}{\partial \boldsymbol{\beta}} = \sum_{i=1}^n \left[ D_i \cdot \frac{1}{\sigma^2} \cdot (y_i - \mathbf{x}_i \boldsymbol{\beta}) \mathbf{x}_i' + (1 - D_i) \cdot \left(-\frac{1}{\sigma}\right) \cdot \frac{\phi\left(\frac{y_0 - \mathbf{x}_i \boldsymbol{\beta}}{\sigma}\right)}{\Phi\left(\frac{y_0 - \mathbf{x}_i \boldsymbol{\beta}}{\sigma}\right)} \mathbf{x}_i' \right] \stackrel{!}{=} 0 \quad (8.41)$$

$$\begin{aligned} \frac{\partial \log L}{\partial \sigma^2} = \sum_{i=1}^n \left[ D_i \cdot (-1) \cdot \frac{1}{2 \cdot \sigma^2} + D_i \cdot \frac{(y_i - \mathbf{x}_i \boldsymbol{\beta})^2}{2 \cdot \sigma^4} \right. \\ \left. + (1 - D_i) \cdot (\mathbf{x}_i \boldsymbol{\beta}) \cdot \frac{1}{2 \cdot \sigma^3} \cdot \frac{\phi\left(\frac{y_0 - \mathbf{x}_i \boldsymbol{\beta}}{\sigma}\right)}{\Phi\left(\frac{y_0 - \mathbf{x}_i \boldsymbol{\beta}}{\sigma}\right)} \right] \stackrel{!}{=} 0. \end{aligned} \quad (8.42)$$

Zur Bestimmung der Nullstellen sind numerische Verfahren notwendig. (Wooldridge, 2009, S. 588) Diese stehen in Statistik Softwarepaketen zur Verfügung. In

dem für die Schätzungen in Kapitel 3 und 4 genutzten Softwarepaket `maxLik` in der Statistik Umgebung `R` sind die Optimierungs-Algorithmen Newton-Raphson, Berndt-Hall-Hall-Hausmann, Broyden-Fletcher-Goldfarb-Shanno, Nealder-Mead und SANN implementiert. (Henningsen und Toomet, 2011, S. 444f.) Da die Likelihood Funktion des Tobit Modells nur ein einziges Maximum besitzt, wird der Algorithmus auf jeden Fall konvergieren und das globale Maximum der Funktion finden. (Amemiya, 1984, S. 157f.) Die Maximum Likelihood Schätzer sind unter den gemachten Annahmen konsistent und asymptotisch normalverteilt. Die Kovarianzmatrix der Schätzer kann bestimmt werden (siehe Wooldridge (2010, S. 677) für Details). (Hill, Griffiths und Lim, 2008, S. 445)

Aufgrund der Eigenschaften des Maximum Likelihood Schätzers im Kontext des Tobit Modells, können einfache Restriktionen eines einzelnen Parameters mittels asymptotischer  $t$  Statistik geprüft werden. Multiple Hypothesen werden mittels Likelihood Ratio Test getestet. (Wooldridge, 2010, S. 677) Dieser wird im letzten Abschnitt dieses Kapitels vorgestellt.

### 8.4.2 Heckmann 2-Schritt Verfahren

Zur Vereinfachung wird für das 2-Schritt Verfahren unterstellt, dass  $y_0 = 0$  ist. Falls  $y_0 \neq 0$  ist, kann die Null als Schranke durch abziehen von  $y_0$  auf beiden Seiten von (8.26) erreicht werden. (Maddala, 1983, S. 159) Für den ersten Schritt des 2-Schritt Verfahrens wird die bereits im vorherigen Abschnitt definierte Indikator Variable  $D_i = 1$ , wenn  $y_i > 0$ , benutzt. Mittels Probit Likelihood werden die Parameter  $\gamma$  der Wahrscheinlichkeit  $\Phi_i(\mathbf{x}_i\gamma)$ , eine unzensierte Beobachtung ( $D_i = 1$ ) zu erhalten, geschätzt. Damit ist es möglich, einen Wert für (8.32) zu schätzen mit

$$\hat{\lambda} = \frac{\phi(\mathbf{x}\hat{\gamma})}{\Phi(\mathbf{x}\hat{\gamma})}.$$

$\hat{\lambda}$  wird zur Schätzung der Regressionsgleichung (8.30) als erklärende Variable verwendet. Damit ergibt sich für die unzensierten Beobachtungen die Schätzgleichung

$$y_i = \mathbf{x}_i\boldsymbol{\beta} + \beta_\lambda \cdot \hat{\lambda}_i + v_i \quad i = 1, \dots, n.$$

Die OLS Methode liefert konsistente Schätzer für  $\boldsymbol{\beta}$ . (Hill, Griffiths und Lim, 2008,

## Tobit Modell

S. 448f.)

Anstatt nur die unzensierten Beobachtungen zu verwenden, ist es auch möglich, Gleichung (8.31) mit allen Beobachtungen zu schätzen. Aufgrund der Vereinfachung, dass  $y_0 = 0$  gesetzt wird, ergibt sich

$$\begin{aligned} E(y_i|\mathbf{x}_i) &= P(y_i > 0) \cdot E(y_i|y_i > 0) + P(y_i = 0) \cdot E(y_i|y_i = 0) \\ &= \Phi\left(\frac{\mathbf{x}_i\boldsymbol{\beta}}{\sigma}\right) \cdot \left(\mathbf{x}_i\boldsymbol{\beta} + \sigma \cdot \frac{\phi\left(\frac{\mathbf{x}_i\boldsymbol{\beta}}{\sigma}\right)}{\Phi\left(\frac{\mathbf{x}_i\boldsymbol{\beta}}{\sigma}\right)}\right) + \left(1 - \Phi\left(\frac{\mathbf{x}_i\boldsymbol{\beta}}{\sigma}\right)\right) \cdot 0 \\ &= \Phi\left(\frac{\mathbf{x}_i\boldsymbol{\beta}}{\sigma}\right) \cdot \mathbf{x}_i\boldsymbol{\beta} + \sigma\phi\left(\frac{\mathbf{x}_i\boldsymbol{\beta}}{\sigma}\right) \end{aligned} \quad (8.43)$$

Die Gleichung (8.43) wird unter Verwendung der geschätzten Wahrscheinlichkeiten aus dem ersten Schritt für  $\phi_i$  und  $\Phi_i$  mittels OLS geschätzt. (Maddala, 1983, S. 159) Auch das 2-stufige Verfahren liefert unter den gemachten Annahmen konsistente, asymptotisch normalverteilte Schätzer, deren Varianz-Kovarianzmatrix bestimmt werden kann. (Amemiya, 1984, S. 13ff.)

### 8.4.3 Censored Least Absolute Deviations

Die beiden vorhergehenden Schätzer basieren auf der Annahme der Normalverteilung der Störgröße. Im Gegensatz dazu benötigt der Censored Least Absolute Deviations (CLAD) Schätzer keine Verteilungsannahme. (Wooldridge, 2010, S. 789) Das CLAD Verfahren ist damit das einzige Verfahren, das auch bei Verletzung der Normalverteilungsannahme konsistente Schätzer liefert. (Amemiya, 1984, S. 28)

Das Modell (8.26) wird beibehalten. Anstatt der Annahme, dass die Störgröße einer Normalverteilung mit Erwartungswert gleich Null folgt, unterstellt das CLAD Verfahren, dass der bedingte Median der Störgröße Null ist:

$$\text{Med}(u|\mathbf{x}) = 0.$$

Es gilt damit

$$\text{Med}(y|\mathbf{x}) = \max(y_0, \text{Med}(\mathbf{x}\boldsymbol{\beta} + u|\mathbf{x})) = \max(y_0, \mathbf{x}\boldsymbol{\beta}).$$

Das bedeutet, dass der bedingte Median einer linearen Funktion entspricht, für den Fall  $\mathbf{x}\boldsymbol{\beta} > y_0$ . (Wooldridge, 2010, S. 671 und S. 687f.) Auf Grundlage der Formulierung des Medians löst der Least Absolute Deviations Schätzer  $\boldsymbol{\beta}$  im zensierten Regressionsmodell (CLAD) (Powell, 1984, S. 305)

$$\min_{\boldsymbol{\beta}} \sum_{i=1}^n |y_i - \max(y_0, \mathbf{x}_i\boldsymbol{\beta})|. \quad (8.44)$$

Die zu minimierende Funktion (8.44) ist nicht stetig differenzierbar. (Powell, 1984, S. 304) Auch ist die Funktion (8.44) nicht konvex, so dass die Möglichkeit lokaler anstatt globaler Minima besteht. (Koenker, 2008, S. 4) Daher ist eine hohe Rechenleistung zur Optimierung notwendig. Das ist ein Nachteil des CLAD Verfahrens. (Greene, 2008, S. 880)

Der CLAD Schätzer ist ein konsistenter Schätzer. Er ist asymptotisch normalverteilt. (Powell, 1984, S. 311f.) Die konsistente Schätzung einer asymptotischen Varianz-Kovarianzmatrix ist möglich (siehe Powell (1984, S. 312ff.) für Details). Der CLAD Schätzer benötigt für diese Eigenschaften keine strikten Annahmen, der bedingte Median muss gleich Null sein. Aufgrund seiner Robustheit gegenüber Verteilungen ist er daher eine natürliche Referenzgröße (Chay und Powell, 2001, S. 36), auf deren Grundlage Spezifikationstests, z.B. ein Hausman Test, durchführbar sind. (Greene, 2008, S. 880)

## 8.5 Fehlspezifikationen des Modells

Die Probleme der seriellen Korrelation, Heteroskedastizität und der Abweichung von der Normalverteilung der Störgröße sind, ebenso wie in der linearen Regression, auch ein Problem im Tobit Modell. (Maddala, 1983, S. 185) Es kann gezeigt werden, dass ML und Heckmann 2-Schritt Schätzer weiterhin konsistent im Fall von serieller Korrelation sind. Hingegen geht die Konsistenz durch Heteroskedastizität und einer nicht-normalverteilten Störgröße im Heckmann 2-Schritt Verfahren verloren. (Amemiya, 1984, S. 23)

## Tobit Modell

Das Ignorieren der Heteroskedastizität führt auch beim ML Verfahren zu einem inkonsistenten Schätzer. (Maddala, 1983, S. 179) Zur Behebung des Problems innerhalb der ML Schätzung ist  $\sigma$  durch ein spezifiziertes  $\sigma_i$  in der Likelihood (8.39) zu ersetzen. (Greene, 2008, S. 876) Für die Spezifikation sind vernünftige Annahmen bezüglich der Form der Heteroskedastizität zu treffen. (Maddala, 1983, S. 179)

Eine Möglichkeit der Spezifikation ist (Chavas und Kim, 2006, S. 70)

$$\sigma_i = \exp(\mathbf{z}_i \boldsymbol{\gamma}). \quad (8.45)$$

In  $\mathbf{z}_i$  befinden sich exogene Variablen zur Erklärung der Standardabweichung gemäß Formel (8.45). Um Homoskedastizität zu testen, ist ein Likelihood Ratio Test (siehe Abschnitt 8.6) mit der Nullhypothese  $H_0 : \boldsymbol{\gamma} = 0$  durchzuführen. (Greene, 2008, S. 876) Zur Schätzung des unrestringierten Modells ergeben sich die notwendigen Bedingungen zur Maximierung der Likelihood wie folgt

$$\frac{\partial \log L}{\partial \boldsymbol{\beta}} = \sum_{i=1}^n \left[ D_i \cdot \frac{1}{\sigma_i^2} \cdot (y_i - \mathbf{x}_i \boldsymbol{\beta}) \mathbf{x}_i' + (1 - D_i) \cdot \left( -\frac{1}{\sigma_i} \right) \cdot \frac{\phi \left( \frac{y_0 - \mathbf{x}_i \boldsymbol{\beta}}{\sigma_i} \right)}{\Phi \left( \frac{y_0 - \mathbf{x}_i \boldsymbol{\beta}}{\sigma_i} \right)} \mathbf{x}_i' \right] \stackrel{!}{=} 0 \quad (8.46)$$

$$\begin{aligned} \frac{\partial \log L}{\partial \boldsymbol{\gamma}} = \sum_{i=1}^n \left( D_i \frac{((y_i - \mathbf{x}_i \boldsymbol{\beta})^2 - \sigma_i^2)}{\sigma_i^2} \cdot \mathbf{z}_i' \right. \\ \left. + (1 - D_i) \frac{\phi \left( \frac{y_0 - \mathbf{x}_i \boldsymbol{\beta}}{\sigma_i} \right) (\mathbf{x}_i \boldsymbol{\beta} - y_0)}{\Phi \left( \frac{y_0 - \mathbf{x}_i \boldsymbol{\beta}}{\sigma_i} \right) \sigma_i} \cdot \mathbf{z}_i' \right) \stackrel{!}{=} 0. \end{aligned} \quad (8.47)$$

Ebenso wie für den Heckmann 2-Schritt Schätzer, ist die Verletzung der Annahme der Normalverteilung auch für den ML Schätzer ein Problem. Das Ergebnis der Maximierung der Likelihood (8.39) führt zu einem inkonsistenten ML Schätzer. (Amemiya, 1984, S. 25) Zur Behebung des Problems sind entweder andere Schätzverfahren zu verwenden, oder es ist eine Transformation durchzuführen, um Normalität herzustellen. (Maddala, 1983, S. 187) Eine Box-Cox Transformation wird in Maddala (1983, S. 190ff.) erläutert. Aufgrund seiner Robustheit gegenüber der Verteilung der Störgröße empfiehlt sich der CLAD Schätzer in dieser Situation.

Es ist aber auch weiterhin die ML Schätzung unter Benutzung einer sinnvollen Alternativverteilung möglich. (Maddala, 1983, S. 190) Jedoch muss dies nicht zwingend zur Lösung des Problems führen, sondern kann die Abweichung von der wahren Verteilung der Störgröße noch vergrößern. (Greene, 2008, S. 880)

Testverfahren zur Überprüfung der Verteilungsannahme sind sinnvoll, damit inkonsistente Schätzer vermieden werden. Eine Möglichkeit zur Überprüfung der Verteilung ist ein Hausman Test. Dabei werden die ML Schätzer  $\hat{\theta}_{ML}$  unter der Nullhypothese der Normalverteilung geschätzt. Der Schätzer  $\hat{\theta}_{ML}$  wird mit einem alternativen Schätzer  $\tilde{\theta}$  verglichen.  $\tilde{\theta}$  muss dafür unter der Alternativhypothese konsistent sein.

$$(\hat{\theta}_{ML} - \tilde{\theta})' V^{-1} (\hat{\theta}_{ML} - \tilde{\theta}) \quad (8.48)$$

mit  $V = V(\tilde{\theta}) - V(\hat{\theta}_{ML})$

ist dann unter der Nullhypothese asymptotisch  $\chi^2$ -verteilt mit  $K$  Freiheitsgraden.  $K$  gibt die Anzahl der Element in  $\theta$  an. (Amemiya, 1984, S. 25) Die Herleitung eines Schätzers für  $V$  für den Fall, dass für  $\tilde{\theta}$  der CLAD Schätzer verwendet wird, ist in Wilhelm (2008, S. 563f.) zu finden. Anstelle des CLAD Schätzers kann ebenso der Symmetrically Censored Least Squares Schätzer für  $\tilde{\theta}$  verwendet werden. Eine weitere Möglichkeit zur Überprüfung der Verteilungsannahme ist ein Test auf Basis der bedingten Momente. (Wilhelm, 2008, S. 563)

## 8.6 Likelihood Ratio Test

Zum Testen multipler Hypothesen im Rahmen einer ML Schätzung ist der Likelihood Ratio Test geeignet. Der Test vergleicht dabei die Funktionswerte der maximierten Log-Likelihood des unrestringierten und des restringierten Modells. (Wooldridge, 2010, S. 481) Werden erklärende Variablen aus dem Modell herausgenommen, so verringert das die Wahrscheinlichkeit und damit auch den Funktionswert der Likelihood. Der Funktionswert der Likelihood ist dabei vergleichbar mit der Residuenquadratsumme im linearen Regressionsmodell. Durch das Weglassen von erklärenden Variablen im linearen Modell erhöht sich die

### *Tobit Modell*

Residuenquadratsumme. Damit muss das  $R^2$  eines restringierten Modells immer mindestens kleiner-gleich zu dem  $R^2$  des unrestringierten Modells sein. Dieses Prinzip ist auch auf den Vergleich der Likelihood-Werte übertragbar. Die Likelihood Ratio Statistik ist zweimal die Differenz der Log-Likelihoods (Wooldridge, 2009, S. 580):

$$LR = 2 \cdot (\log L_{ur} - \log L_r). \quad (8.49)$$

$\log L_{ur}$  und  $\log L_r$  sind jeweils die Log-Likelihoods des unrestringierten und restringierten Modells.  $LR$  ist unter der Nullhypothese asymptotisch  $\chi^2$ -verteilt. Die Freiheitsgrade der Verteilung entsprechen der Anzahl der Restriktionen in der Hypothese. (Greene, 2008, S. 534) Der Likelihood Ratio Test fand bereits bei Tobin (1958) Anwendung. Tobin (1958) hat einen Test auf Gesamtsignifikanz des formulierten Tobit Modells mittels Likelihood Ratio Statistik beschrieben. (Tobin, 1958, S. 28f.)



# Literatur

- Amemiya, Takeshi (1984). »Tobit Models: A Survey«. In: *Journal of Econometrics* 24, S. 3–61.
- (1985). *Advanced econometrics*. Cambridge, Mass.: Harvard Univ. Pr.
- Chavas, Jean-Paul und Kwansoo Kim (2006). »An econometric analysis of the effects of market liberalization on price dynamics and price volatility«. In: *Empirical Economics* 31, S. 65–82.
- Chay, Kenneth Y. und James L. Powell (2001). »Semiparametrics Censored Regression Models«. In: *Journal of Economic Perspectives* 15, S. 29–42.
- Chung, Ching-Fan und Arthur S. Goldberger (1984). »Proportional projections in limited dependent variable models«. In: *Econometrica* 52, S. 531–534.
- Fahrmeir, Ludwig, Thomas Kneib und Stefan Lang (2009). *Regression: Modelle, Methoden und Anwendungen*. Berlin: Springer.
- Greene, William H. (1981). »On the asymptotic Bias of the Ordinary Least Squares Estimator of the Tobit Model«. In: *Econometrica* 49, S. 505–513.
- (2008). *Introductory econometrics: a modern approach*. 6. Aufl. Upper Saddle River, New Jersey: Pearson Prentice Hall.
- Henningsen, Arne und Ott Toomet (2011). »maxLik: A package for maximum likelihood estimation in R«. In: *Computational Statistics* 26, S. 443–458.
- Hill, R.Carter, William E. Griffiths und Guay C. Lim (2008). *Principles of Econometrics*. 3. Aufl. Hoboken, New Jersey: John Wiley & Sons.
- Kim, Kwansoo und Jean-Paul Chavas (2002). »A dynamic analysis of the effects of a price support program on price dynamics and price volatility«. In: *Journal of Agricultural and Resource Economics* 27, S. 495–514.
- Koenker, Roger (2008). »Censored Quantile Regression Redux«. In: *Journal of Statistical Software* 27, S. 1–25.
- Komlos, John und Bernd Süßmuth (2010). *Empirische Ökonomie: Eine Einführung in Methoden und Anwendungen*. Heidelberg: Springer.

## Literatur

- Maddala, Gangadharrao S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. 1. Aufl. Cambridge (UK): Cambridge University Press.
- Powell, James L. (1984). »Least absolute deviations estimation for the censored regression model«. In: *Journal of Econometrics* 25, S. 303–325.
- Tobin, James (1958). »Estimation of Relationships for Limited Dependent Variables«. In: *Econometrica* 26, S. 24–36.
- Wilhelm, Mark Ottoni (2008). »Practical Considerations for Choosing Between Tobit and SCLS and CLAD Estimators for Censored Regression Models with an Application to Charitable Giving«. In: *Oxford Bulletin of Economics and Statistics* 70, S. 559–582.
- Wooldridge, Jeffrey M. (2009). *Introductory econometrics: a modern approach*. 4. ed., internat. student ed. Mason, Ohio: South-Western, Cengage Learning.
- (2010). *Econometric Analysis of cross section and panel data*. 2. Aufl. Cambridge, Massachusetts: The MIT Press.

# Kapitel 9

## Schlussbetrachtung

Inhalt der vorliegenden Arbeit sind sechs verschiedene Untersuchungen. Alle Untersuchungen eint, dass sie den Fokus auf quantitativen Analysen der Agrar- und Ernährungsmärkte haben oder aber die Vermittlung der quantitativen Methoden in der universitären Lehre unter Einsatz von eLearning näher betrachten. In der folgenden Schlussbetrachtung werden die Ergebnisse der einzelnen Untersuchungen kritisch gewürdigt. Zusätzlich werden Ansätze für Erweiterungen und zukünftige Forschungsfragen diskutiert.

### **Validation study of consumer scanner and retail scanner data**

Der Beitrag untersucht den Messfehler in einem Haushaltspanel. Es kann gezeigt werden, dass die Preisangaben der Panelteilnehmer Messfehler beinhalten. Der gefundene Messfehler erfüllt dabei aber nicht die Annahmen eines klassischen Messfehlers, da er mit dem wahren Wert der Preisangabe korreliert. Zur Validierung werden Scannerdaten aus dem Lebensmitteleinzelhandel (LEH) verwendet. Es wird angenommen, dass die Daten aus dem Lebensmitteleinzelhandel aufgrund der automatischen Erfassung an den Kassen während des Bezahlvorgangs fehlerfrei sind.

Ein direkter Vergleich der beiden Datensätze gestaltet sich schwierig aufgrund fehlender Angabe der Einkaufsstätte im Haushaltspanel. Dort wird nur die Handelskette, zu der das aufgesuchte Geschäft gehört, angegeben. Um die Angabe der Einkaufsstätte vernachlässigen zu können, werden in der Untersuchung nur Artikel beachtet, für die einheitliche Preise innerhalb einer Kette je Kalenderwoche für die einzelnen Einkaufsstätten im LEH Datensatz vorliegen. Dieser einheitliche

## *Schlussbetrachtung*

Preis wird als wahrer Preis unterstellt.

Dieses restriktive Vorgehen dezimiert die untersuchten Beobachtungen erheblich. Die Repräsentativität des Datensamples wird eingeschränkt. Innerhalb der Kette Plus, dem einzigen Discounter in der Untersuchung, wird der größte Anteil an Artikeln untersucht. Da Discounter im Vergleich zu anderen Geschäftsformaten vermehrt konstant niedrige Preise anbieten, sind die untersuchten Artikel eher konstant im Preis. Daher ist der gefundene Fehler insgesamt als untere Grenze des Messfehlers anzusehen.

Die Ergebnisse beruhen auf der Annahme, dass für die untersuchten Artikel eine einheitliche Preispolitik innerhalb einer Kette angewendet wird. Die strikte Definition zur Auffindung einheitlicher Preispolitik soll dies gewährleisten. Eine nähere Betrachtung der Sonderangebots-Indikatorvariable zeigt aber, dass die einzelnen Einkaufsstätten einer Kette unabhängig voneinander Sonderangebote setzen. Es ist fraglich, ob die Definition der einheitlichen Preispolitik eher geeignet ist, Artikel mit nur einer geringen Anzahl an Sonderangeboten, d.h. Preisänderungen, aufzufinden.

Zur Steigerung der Repräsentativität der untersuchten Daten sollte in zukünftigen Untersuchungen eine nicht so stark einschränkende Methode zur Validierung der Daten verwendet werden. Falls dazu weiterhin Scannerdaten aus dem LEH verwendet werden, so ist deren Datenqualität kritisch zu überprüfen. Die automatische Erfassung der Beobachtung an den Kassen kann auch Fehler erzeugen. Neben der automatischen Erfassung der Daten an den Kassen sind auch kritisch die zusätzlichen Artikelbeschreibungen, zu der auch die Angabe der Stückzahl je EAN gehört, zu untersuchen. Unterschiede in den Artikelbeschreibungen deuten darauf hin, dass diese nachträglich von Hand dem Datensatz hinzugefügt worden sind. Aufgrund der Verwendung der Stückzahl in der Untersuchung zur Ermittlung eines Stückpreises, ist dies eine weitere Quelle für mögliche Ungenauigkeiten in der Untersuchung.

Allgemeines Ziel für zukünftige Untersuchungen sollte die Möglichkeit der Korrektur von Messfehlern in Haushaltspaneln sein. Ein Korrekturverfahren kann in der hier vorliegenden Untersuchung nicht angewendet werden, aufgrund fehlender Repräsentativität des Beobachtungssamples. Die vollständige Vermeidung von Messfehlern in Haushaltspanels wäre hingegen der zu erreichende Optimalzustand.

## **Tobit regression to estimate impact of EU market intervention in the German butter market**

Ziel der Untersuchung ist die Bestimmung des Effekts der EU Interventionspolitik auf dem Buttermarkt am Beispiel Deutschlands. Dies geschieht vor dem Hintergrund der Senkung des Interventionspreises im Rahmen der Milchmarktreform. Ergebnis ist, dass die Preisstützung eine varianzmindernde Wirkung auf die Preisentwicklung hat. Zusätzlich kann ein positiver Einfluss auf das Marktpreisniveau durch die Interventionspolitik beobachtet werden, auch wenn der Interventionspreis unterhalb des Marktpreises liegt und keine Interventionsaufkäufe stattfinden.

Für die Untersuchung wird ein Tobit Modell verwendet, welches die Linkszensierung des Marktpreises durch den Interventionspreis berücksichtigt. Für den Störterm des Modells wird eine heteroskedastische Normalverteilung angenommen. Auf Grundlage der Verteilungsannahme ist die Herleitung von Varianz und Erwartung des modellierten Marktpreises möglich. Auch das Schätzverfahren der Maximum Likelihood (ML) Methode benötigt eine Verteilungsannahme für die Likelihood Funktion.

Die zunehmende Volatilität auf dem Milchmarkt führt zu extremen Preisbeobachtungen. Die Annahme der Normalverteilung ist kritisch zu untersuchen. Es besteht die Möglichkeit, dass Extrema unterschätzt werden. Zur Untersuchung der möglichen Fehlspezifikation wird eine Schätzung mittels Censored Least Absolute Deviations (CLAD) Methode angewendet. Aufgrund des Verzichts auf Verteilungsannahmen ist diese Methode aus der Quantilregression ein sehr robustes Schätzverfahren. Die Schätzergebnisse der CLAD Schätzung weichen in der Größe nicht stark von den Ergebnissen der ML Schätzung ab. In einer zukünftigen Untersuchung ist ein Hausmann Test durchzuführen, um zu überprüfen, ob die Abweichungen statistisch signifikant sind. Ist dies der Fall, so ist die Untersuchung auf der Annahme einer anderen Verteilung durchzuführen.

Eine weitere Aufgabe für zukünftige Untersuchungen ist die Abklärung der Erwartungsbildung. In der Untersuchung wird der Erwartungswert des Marktpreises als gewichtetes Mittel zwischen dem Interventionspreis und dem Marktpreis, der sich auf einem freien Markt bilden würde, berechnet. Die Gewichte resultieren aus den Wahrscheinlichkeiten eines Interventionseingriffes durch den Staat. Es ist

## *Schlussbetrachtung*

fraglich, ob dies eine realistische Erwartungsbildung liefert.

### **Beurteilung der Kompensation auf dem EU-Milchmarkt auf Basis eines Tobit-Ansatzes**

Die Auswirkungen der Liberalisierung des EU-Milchmarkts sind Inhalt dieser Untersuchung. Es werden Tobit Modelle genutzt, um den Magermilchpuvermarkt und den Buttermarkt unter Berücksichtigung der Linkszensierung der Marktpreise durch die Interventionspreise zu modellieren. Die Modelle werden mit monatlichen Daten von Januar 1995 bis Juli 2008 geschätzt. Aus den Modellen werden Prognosen der Marktentwicklung ohne Reduktion der Interventionspreise ab dem Jahr 2004 abgeleitet und ein approximativer Milchpreis, der Referenzpreis, ermittelt. Daraus wird der Preisrückgang aufgrund der Reform berechnet und ins Verhältnis zu den staatlichen Zahlungen an die Landwirte zur Kompensation des Politikwandels gesetzt.

Die Zahlungen zur Kompensation des Einkommensrückgangs sind über den gesamten Zeitraum betrachtet zu gering. Die Unterkompensation ist in ihrer Höhe aber geringer als bei einer anderen Untersuchung, die die Entwicklung des Referenzpreises direkt gekoppelt an die Entwicklung des Exportpreises (gleich Weltmarktpreis plus Exportsubvention) abgeschätzt hat. Aufgrund der stabilisierenden Wirkung der Interventionspolitik sind hier in der Untersuchung die Preisausschläge auf dem inländischen Markt, verursacht durch die Weltmarktentwicklung, nicht so stark ausgeprägt. Doch zeigt der Einfluss des Weltmarktes auch hier, dass der Grad der Kompensation in den einzelnen Jahren schwankend ist. Kompensationszahlungen sollten in ihrer Höhe daher variabel an die jeweilige Marktsituation angepasst werden, um ihr Ziel erfüllen zu können.

Kritisch in dieser Untersuchung ist die Anwendung des Tobit Modells. Das Modell berücksichtigt die Beschränkung des Marktpreises durch den Interventionspreis. Der modellierte Marktpreis fällt in dem Modell niemals unterhalb des Interventionspreises. Auf den deutschen Butter- und Magermilchpulvermärkten liegen allerdings Preisbeobachtungen unterhalb des jeweiligen Interventionspreises vor. Die Gründe dafür sind in der Beschränkung der Intervention hinsichtlich Zeitraum und Menge, aber auch in der zeitlichen Verzögerung der Aufkäufe zu finden. Solche

Situationen kann das Tobit Modell nicht korrekt abbilden.

Der verwendete Zeitraum für die Untersuchung wurde gewählt, um Vergleichbarkeit zu der Studie von Loy und Steinhagen (2009) zu gewährleisten. Für eine zukünftige Untersuchung ist der Untersuchungszeitraum erheblich auszuweiten. Speziell auf dem Buttermarkt ist seit Anfang der 1990er Jahre der Marktpreis höher als der Interventionspreis. Dadurch gehen keine Preisbeobachtungen in die Schätzung ein, deren Entwicklung durch die Interventionspolitik dominiert wurde. Eine zusätzliche Erweiterung für zukünftige Untersuchungen wäre die Verwendung eines multivariaten Tobit Modells, um mögliche gegenseitige Beeinflussungen der beiden Märkte zu berücksichtigen.

### **Einsatz eines Prozessmanagementzyklus zur Optimierung von eLearning: Praxisbeispiel „eLearning QMM“**

Der Beitrag beschäftigt sich mit der Durchführung einer Lehrveranstaltung im Blended Learning Format anhand eines Prozessmanagementzyklus. Der Zyklus wird anhand des Projektes „eLearning QMM“ dargestellt. Die eingesetzten eLearning Inhalte sind nicht neuartig, wohl aber die Managementmethode. Für eine zukünftige Untersuchung ist die konkretere Beschreibung des didaktischen Designs und der jeweiligen Abwandlung zwischen den verschiedenen Jahrgängen der Veranstaltung hilfreich. Anhand einheitlich eingesetzter Messmethoden könnte über die Jahrgänge hinweg eine Validierung der Umsetzung erhoben werden, zur Messung der Qualitätsverbesserung durch den Prozessmanagementzyklus. Dabei sind aber die Effekte der Änderungen im didaktischen Design auf die Bewertung zu berücksichtigen.

### **Evaluation of eLearning - A study of Undergraduate Agricultural Economics course**

Die Untersuchung präsentiert die Evaluation eines eLearning Einsatzes zur Unterstützung einer Lehrveranstaltung, die zusätzlich zur Präsenzvorlesung eine Lernplattform mit verschiedenen eLearning Inhalten verwendet. Die Ergebnisse der Untersuchung zeigen, dass eine hohe Punktzahl in einem Test nicht mit einer hohen Zufriedenheit mit der Lernumgebung einhergeht, dass aber eine hohe

### *Schlussbetrachtung*

Zufriedenheit mit der Lernumgebung hingegen zum einen bei Studenten mit Fachinteresse, zum anderen bei Studenten, die sich als passive Lerntypen einschätzen, zu finden ist. Aufgrund dieser Resultate wird der eLearning Einsatz überarbeitet und neue Ziele des Einsatzes werden definiert.

Die Ergebnisse der Untersuchungen sind vorerst auf das untersuchte Modul und den einen Jahrgang zu beziehen. Für zukünftige Untersuchungen ist es wichtig, die untersuchte Anzahl an Jahrgängen, Fachbereichen und Modulen auszuweiten. Es ist zu klären, ob es sich um allgemeingültige Ergebnisse handelt. Die Verallgemeinerung ist im Moment aufgrund der Beschränkung der untersuchten Stichprobe auf nur eine evaluierte Veranstaltung nicht möglich. Weiterhin ist die Untersuchung dahin auszuweiten, dass die Gruppe der Studenten, die vom eLearning Einsatz profitiert, zu identifizieren ist und das eLearning zielgerichteter gestaltet werden kann.

### **Quantitative Untersuchung zur Effizienz von eLearning<sup>5</sup>**

Die Entwicklung eines praxisrelevanten Ansatzes zur Beurteilung der Effizienz eines eLearning Einsatzes in der Hochschullehre ist das Ziel dieses Beitrags. Es wird ein approximatives Kosten-Nutzen-Modell aus Sicht der Lehrenden aufgestellt und anhand der Betreuung von Projektarbeiten mit und ohne Learn Content Management System (LCMS) quantifiziert. Als Hauptaufwand wird dabei die individuelle Betreuung der Studenten mittels elektronischer Nachrichten gesehen. Der Hauptnutzen ist die Verbesserung der erreichten Noten. Es wird gezeigt, dass es keine statistisch signifikanten Unterschiede zwischen den beiden Szenarien gibt.

Die Quantifizierung von Kosten und Nutzen ist kritisch zu sehen und in Zukunft zu überarbeiten. Die Anzahl der geschriebenen Nachrichten je Student ist zwar eine leicht zu erhebende Größe, birgt aber Schwierigkeiten. Der Aufwand je Nachricht ist vom Einzelfall abhängig, denn je nach gestellter Frage oder auch Problem des Studentens ist die Antwort zwischen kurzer standardisierbarer Nachricht und komplexer umfangreicher Beschreibung zu gestalten. Es ist kritisch zu hinterfragen, ob dieser Aufwand wirklich durch den Einsatz eines LCMS beeinflussbar ist. Auch

---

<sup>5</sup>Kritik und Erweiterungsvorschläge basieren auf der Diskussion des Beitrags im Rahmen der Konferenz *eLearning Baltics 2011*.



auf der Nutzenseite sollte die wiederum leicht zu erhebende Größe Note durch eine andere messbare Größe ersetzt werden, da fraglich ist, inwieweit die Leistung der Studenten durch die Betreuung beeinflussbar ist. Anstelle der Note könnte dies zum Beispiel die Zufriedenheit der Studenten sein.

Die Verallgemeinerung der Untersuchungsergebnisse ist schwierig, da die Untersuchung auf die Betreuung von Projektarbeiten fokussiert ist. Es ist daher eine zukünftige Aufgabe, die Untersuchung auszuweiten auf weitere eLearning Szenarien. Nur dadurch kann eine Aussage bezüglich der Übertragbarkeit der Ergebnisse gemacht werden. Bei der Ausweitung der Untersuchung sollte aber nicht nur der Frage nachgegangen werden, ob der jeweilige Einsatz von eLearning effizient ist, sondern auch, was ein effizientes eLearning Projekt ausmacht. Daraus sollte die Schlussfolgerung erfolgen, wie ein Projekt zu gestalten ist, das bei gleichem Nutzen im Vergleich zu reiner Präsenzlehre durch den zusätzlichen eLearning Einsatz Kosten einsparen kann.



# Literatur

Loy, Jens-Peter und Carsten Steinhagen (2009). »Preissenkung und Kompensation auf dem EU-Milchmarkt«. In: *Agrarwirtschaft - Zeitschrift für Betriebswirtschaft, Marktforschung und Agrarpolitik* 58 (5/6), S. 259–268.



# Kapitel 10

## Summary

In 1961 the second annual convention of the Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaus (GeWiSoLa) resulted in the establishment of econometrics being an inherent part of the agricultural economics in Germany. Econometrics offer methods to empirically inspect theoretical considerations. In case theory differs with empiricism, the theory is falsified. Revision and improvement is needed for it. With it, econometrics supply agricultural economics with major contributions regarding theoretical developments to explain reality.

Since 1961 attempts are made to enhance quantitative methods. Over time agricultural economists have been in charge for various improvements, e.g. the general analysis of the demand, the supply, and the pricing on agricultural markets, the examination of imperfect markets, and policy evaluation. Ex-post evaluation of policies is needed to judge the effects. Results are used to decide on efficient policy planning. Especially in the EU and further international institutions the importance of empirical policy evaluation will increase. Therefore, improvement, extension, and application of quantitative methods will remain an important responsibility for agricultural economics in future.

In chapter 2 to 4 the studies deal with quantitative applications on agricultural and food market. The first study evaluates consumer scanner panel data with regard to measurement error in price records. Second study examines the effect of European intervention politics in the butter market. A heteroscedastic Tobit model is estimated which takes account of world market development. This model specification is also used in the third study. There, it is estimated for the butter market as well as for skimmed milk powder market. Using the estimation results the price decrease due to reduction in intervention prices is predicted. Finally,

## *Summary*

price reduction is compared to executed compensation payments.

In chapter 5 to 7 the studies are concerned with the transfer of econometrical knowledge within an undergraduate agricultural economics course. The focus is on eLearning implementation. The study in chapter 5 describes a process management to plan, implement, evaluate, and optimise eLearning. In the proximate chapter the study addresses the evaluation in detail. Aspects of the efficiency of eLearning are discussed in chapter 7 together with an approximate appraisal of the effort in supervising students' project work.

## **Validation study of consumer scanner and retail scanner data**

The aim of this study is, first, to analyse consumer scanner panel data with regard to measurement error in price records and, second, to classify potential biases. Studies, which use such data, commonly rely on the assumption of unbiased data. In the used consumer scanner data participating households record each bought article, corresponding price, and attended chain after purchase. For validation, retail scanner data are used which contain information about sales on store-level. Matching between both datasets is impossible on store-level. Therefore, only articles with uniform pricing across all stores of a chain and particular calendar week are analysed. With it, the visited store in detail is negligible. Information about the chain is sufficient. In total, I find an additive, normal distributed measurement error in consumer price data, which is correlated with the price level. The assumption of a classical measurement error has to be rejected. Additional, I observe statistically significant effects of age and income on the measurement error. The calculated reliability ratio is statistically significant less than 1 for each product category, except UHT milk. Hence, in linear regression models estimates are biased, whether the mismeasured price is an explanatory variable or it is the dependent variable regressed on age or income. Due to the non-classical measurement error the bias is rather complex. As a result of the restrictive choice of analysed articles, detected measurement error is only a lower limit. In general the measurement error and the resulting bias are expected to be larger in size.

## **Tobit regression to estimate impact of EU market intervention in the German butter market**

This study examines the effect of European intervention politics in the European butter market in the context of market liberalization using the example of Germany. A heteroscedastic Tobit model is estimated using German butter market data from 1979-2012. Model specification is basing on an error correction model. It is done to account for the non-stationary nature and cointegration of the export price series (world market price plus export subsidies) and German butter price series. We get similar results to a study which analysed the U.S. butter market. There is evidence that price support reduces price instability in the German butter market. With the estimated model we predict market development without lowering the intervention price since 2004. It indicates a positive effect towards the expected butter price, even though the mean market price is higher than the intervention level and no intervention purchases take place. Maximum likelihood method is conducted for estimation. For comparative purpose censored least absolute deviations estimation is applied as well. A Hausmann test should be used to test, whether differences are statistically significant.

## **Beurteilung der Kompensation auf dem EU-Milchmarkt auf Basis eines Tobit-Ansatzes**

Due to the reform of the European dairy market, the level of intervention prices has been continuously reduced on butter and skimmed milk powder market since 2004. With it, the EU started compensation payments. The question rises whether the payments are able to sufficiently compensate for the resulting price reductions. To investigate the price reduction a heteroscedastic Tobit model is estimated to explain price development on butter and skimmed milk powder market. The model accounts for the development of the world market price. Using the estimated model the effect of the reduction in the price support is evaluated. Market development without the reform is predicted and an approximate milk price is calculated. The calculated milk price is used to estimate the degree of compensation.

## *Summary*

The results show stabilizing effects of intervention politics. Throughout the period from 2004 to 2008, farmers are not fully compensated. To cover milk price reduction the compensation payments need to increase by 50%. To stabilise the extent of compensation, compensation payments should be variable to adapt world market development.

## **Einsatz eines Prozessmanagementzyklus zur Optimierung von eLearning: Praxisbeispiel "eLearning QMM"**

To create a purposeful eLearning environment, the eLearning content needs to be planned, implemented, evaluated, and improved. Various inputs, e.g. students' and lecturers' requirements, content of the curriculum, materials, hardware, and software, are transformed into the output eLearning environment. A process management supports such a recurring process of transformation. The study develops a process management for an eLearning process. The process management is suitable passing through several cycles of the eLearning process. Each circuit corresponds to a course per semester.

Using the project "eLearning QMM", the practical approach of this management method is shown. The project "eLearning QMM" is part of the programme for sustainable integration of eLearning within classroom education at Christian-Albrechts-University Kiel. It takes place during summer term 2010 and summer term 2011. eLearning supports the module "Quantitative Methoden der Marktanalyse" (QMM). This module introduces theory and application of statistical and econometrical methods.

Altogether, the module QMM stood to benefit from the process management. The cyclical approach assists the reusability of eLearning content. Due to repeating cycles, standardisation takes place and efficiency rises. It helps to prevent recurrent mistakes. Also, the communication benefits through the improved transparency of planning and implementation. It simplifies cooperation, e.g. with technical support. Limitation of the process management is that not every target of eLearning use is quantifiable. Though, it is not within the quantitative control of the process management. Additional, at universities there are problems of correct cost allocations for central service units, e.g. technical support. It causes



difficulties to keep an eye on efficiency with regard to costs.

### **Evaluation of eLearning - A study of Undergraduate Agricultural Economics course**

In this study we evaluate the combination of traditional classroom education with additional content and support provided by an online Learn Content Management System (LCMS). The LCMS offers instruments to support various forms of teaching and learning. This eLearning usage was evaluated by a written survey of the participating students. Additionally, students answered a multiple choice test dealing with the content of the module.

The survey focuses on individual learning characteristics and eLearning aspects. Findings are that not all dimensions of eLearning success meet the expected outcomes. Principal component analysis shows, high score in the test not necessarily accompanies satisfaction with the eLearning environment. Using achieved scores in examinations to measure successful application of eLearning is suitable only to a limited extent. Instead, high score in an exam is attended by extrinsic motivation to achieve a good mark in the final exam and effort to learn continuously. Satisfaction of the LCMS is attended by econometrical interest as well as passive working attitude. eLearning objectives have to be defined with respect to eLearning's target group. Success of eLearning is not only encouraging active learning attitude. Being an obligatory course, success is also to spend satisfaction with the learning environment, although there is only passive learning attitude.

### **Quantitative Untersuchung zur Effizienz von eLearning**

Commonly, in higher education teaching efficiency and effectiveness of eLearning are not called in to question. The assumption, students have got major benefits from eLearning instead of lecturer, is expressed in some publications. To quantify this assumption, the study analysed the effect of eLearning on the effort in supervising students' project work.

Individual supervision is the main driver of costs in supervising students' project work. In the analysed scenario individual supervision mainly takes place via

## *Summary*

electronic messages. Therefore, the approximate appraisal of the effort is done using the number of messages per student written by the supervisor. The scenario with eLearning is compared to a situation without it. Wilcoxon rank sum test is applied to find general group differences. Additionally, the effect of the number of written messages towards the achieved grade is analysed.

Neither for the achieved grades, nor the number of written messages group differences are statistically significant. Also, there is no statistically significant relationship between the number of written messages and the achieved grade. Overall, eLearning do not per se cause a reduction of effort in supervising students' project work.

# Lebenslauf von Meike Wocken

## Persönliche Daten

Geburtsname: Dickel

geb. 20. Juli 1983 in Bielefeld (Nordrhein-Westfalen)

Staatsangehörigkeit: deutsch

## Tätigkeiten

- 10/2011 - heute      Lehrbeauftragte  
FH Bielefeld, Campus Minden  
Modul "Statistik" im Studiengang Wirtschaftsingenieurwesen (B.Sc.)
- 09/2009 - 08/2012    Wissenschaftliche Mitarbeiterin  
Institut für Agrarökonomie, Abteilung Marktlehre  
Christian-Albrechts-Universität zu Kiel
- 11/2008 - 08/2009    Wissenschaftliche Mitarbeiterin  
Institut für Statistik und Ökonometrie  
Georg-August-Universität Göttingen

## Studium

- 04/2010 - 06/2012    Zusatzstudium Mathematik - Methoden und Modelle (M.Sc.)  
Fernuniversität Hagen
- 09/2009              International Summer School  
Computational Aspects in Environmental Statistics  
Bulgarien  
European Science Foundation-Stipendiatin
- 10/2005 - 10/2008    Diplom Volkswirtschaftslehre  
Georg-August-Universität Göttingen
- 10/2003 - 09/2005    Vordiplom Volkswirtschaftslehre  
Universität Bielefeld

Kiel, Januar 2013



# Danksagung

Ich bedanke mich bei Prof. Dr. Jens-Peter Loy für die Betreuung meiner Promotion und die mir gewährten Freiräume. Herrn Prof. Dr. Martin Schellhorn möchte ich für die Übernahme des Zweitgutachtens danken. Auch möchte ich mich bei Prof. Dr. Thomas Kneib für die gute und hilfreiche Zusammenarbeit bedanken. Besonderer Dank gilt Janine Empen, Friedrich Hedtrich, Angela Hoffmann, Thore Holm, Kirsten Kriegel, Fabian Schaper, Karen Schröder, Heike Senkler und Carsten Steinhagen für das angenehme Arbeitsklima am Institut für Agrarökonomie in Kiel. Die guten fachlichen Diskussionen, aber auch gemeinsame sportliche Herausforderungen außerhalb des Büroalltags haben die Zeit in Kiel für mich sehr wertvoll gemacht. Schließlich danke ich meinem Mann Christian Wocken für seine Geduld und Unterstützung während der Fertigstellung dieser Arbeit.

Meike Wocken

Kiel, Januar 2013